RNA structure determination by solid-state NMR spectroscopy

Alexander Marchanka¹, Bernd Simon¹, Gerhard Althoff-Ospelt² & Teresa Carlomagno¹,³

Knowledge of the RNA three-dimensional structure, either in isolation or as part of RNP complexes, is fundamental to understand the mechanism of numerous cellular processes. Because of its flexibility, RNA represents a challenge for crystallization, while the large size of cellular complexes brings solution-state NMR to its limits. Here, we demonstrate an alternative approach on the basis of solid-state NMR spectroscopy. We develop a suite of experiments and RNA labeling schemes and demonstrate for the first time that ssNMR can yield a RNA structure at high-resolution. This methodology allows structural analysis of segmentally labelled RNA stretches in high-molecular weight cellular machines—indeed of their ability to crystallize— and opens the way to mechanistic studies of currently difficult-to-access RNA-protein assemblies.

¹Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany. ²Bruker BioSpin, Silberstreifen 4, 76287 Rheinstetten, Germany. ³Helmholtz Zentrum für Infektionsforschung, Inhoffenstrasse 7, 38124 Braunschweig, Germany. Correspondence and requests for materials should be addressed to A.M. (email: marchank@embl.de) or to T.C. (email: teresa.carlomagno@embl.de).
In gene expression regulation, stress response and pathogens infection, a multitude of non-coding RNAs and ribonucleoprotein complexes accomplish their function cycling through transient intermolecular contacts and related conformational changes. Taking influence on these processes requires a mechanistic understanding of the intermolecular interactions, which, in turn, necessitates structural information. Both naked RNAs and RNPs represent a challenge for structural biology. The conformational plasticity of the RNA restricts application of X-ray crystallography, while the high-molecular weight of the RNA (or RNP) of interest pushes solution-state NMR to its limits. Lately, solid-state NMR (ssNMR) spectroscopy, which is applicable to macromolecules of any size in non-crystalline form, has emerged as a powerful alternative to study the structure of amyloid fibrils, membrane proteins, and large protein–protein assemblies. Despite these successes, ssNMR has been rarely applied to nucleic acids, and the methodology for RNA structure determination is still lacking. Here, we present the first de novo structure determination of RNA by ssNMR, together with the experimental methods we developed for it. We demonstrate that RNA structure is accessible at high resolution by ssNMR, beyond the differences to crystallographic reference structures expected as a consequence of packing forces and RNA-RNA contacts in the crystals.

Results
The Box C/D RNA bound to L7Ae. In this study, we solve the structure of the 26mer Box C/D RNA from *Pyrococcus Furiosus* (*Pf*) as part of the complex with the protein L7Ae (Fig. 1a and Supplementary Fig. 1). $^{13}$C and $^{15}$N line widths of 0.4–0.5 and 0.6–0.8 p.p.m., respectively, allow determining the structure by ssNMR data with a precision of 0.8 Å. We choose to study this RNA for the following reasons. First, we were unable to crystallize the L7Ae-Box C/D RNA complex with the RNA sequence of Fig. 1a, despite the existence of crystallographic structures of homologues. This demonstrates that the crystallization of RNA-protein complexes can be unexpectedly challenging, in dependence of the RNA sequence. Second, the RNA of Fig. 1a contains the conserved Box C and Box D sequences, which build the so-called k-turn motif. The geometry of the k-turn is measured by an angle $\phi$, which is variable in the free RNA and depends on the concentration of magnesium. Upon protein binding, the k-turn parameter $\phi$ adopts a value close to $23^\circ$ for all k-turn motifs investigated to date, independent of the experimental method, the exact sequence of the RNA or the species it belongs to ref. 12. The conservation of this structural motif offers the opportunity to verify the accuracy of the structure obtained by ssNMR, beyond the differences to crystallographic reference structures expected as a consequence of packing forces and RNA-RNA contacts in the crystals.

Resonance assignment and measurement of distance restraints are the key steps in structure determination by ssNMR. In contrast to proteins, where homonuclear $^{13}$C-$^{13}$C correlations are sufficient for resonance assignment, the poor chemical shift dispersion of ribose resonances in RNA requires additional heteronuclear editing. We find that three-dimensional pulse schemes yield low signal-to-noise within our experimental set-up, while the quality of two-dimensional spectra allows for both assignment and quantification of cross-peaks. Therefore, our strategy does mainly without three-dimensional experiments and
resolves spectral overlaps by selective labeling. To make the method accessible to a broad community, we abstain from using atom-selective labelling and employ only RNAs that can be produced with commercially available building blocks by in vitro transcription. In this study, we designed eight combinations of double-nucleotide-type selective-labelled RNAs (Supplementary Fig. 1), to accomplish both resonance assignment and measurement of structural parameters.

**Sequence-specific assignment.** First, we assigned the spin-systems of individual nucleotides in a non-site-specific manner\(^{13}\). For the 26mer Box C/D RNA, we found six adenosines, seven guanosines, three cytosines and four uridines spin-systems. Analysis of canonical coordinates of ribose shifts\(^{14}\) (Supplementary Fig. 2) suggests that two adenosines, three guanosines and one uridine are not located in regular A-form helices. This allowed us to attribute the uridine spin-system to U20, which was used as starting point for sequential assignment. Seventeen out of the 26 nucleotides were assigned site-specifically using correlations between the C1’ or C6/C8 atoms of nucleotide i to the carbons of neighbouring nucleotides i ± 1, as well as nucleotides of the opposite strand (Fig. 1b, Supplementary Table 1). To improve resolution, before the \(^{13}\)C,\(^{13}\)C transfer, the magnetization of C1’ or C6/C8 was correlated to the respective N1/N9 via TEDOR (Transferred-Echo-DOUBLE-Resonance)\(^{15}\), yielding a two-dimensional (2D) \(^{15}\)N,\(^{13}\)C correlation (Fig. 2a,b, Methods). This allows clear distinction of purine and pyrimidines in double-nucleotide-type selective-labelled samples (for example, G,\(^{13}\)C\(_{lab}\)- or A,\(^{13}\)C\(_{lab}\)-RNAs). For the long-range carbon–carbon transfer, we tested different mixing sequences and finally settled on the PDSD (Proton-Driven-Spin-Diffusion) scheme due to its superior sensitivity\(^{16}\).

\(^{13}\)C,\(^{15}\)N-TEDOR-\(^{13}\)C,\(^{13}\)C-PDSD (Supplementary Fig. 3a) was applied with a mixing time of 700 ms to six selective-labeled RNAs (Supplementary Fig. 1) and yielded several inter-nucleotide contacts up to a distance of 9–10 Å (Fig. 2a,b, Supplementary Figs 4 and 5). As an example of sequential assignment, G,\(^{13}\)C\(_{lab}\)- and A,\(^{13}\)C\(_{lab}\)-RNAs yielded multiple correlations between the C1’,C6 of U20 and both a guanosine and an adenosine spin-system; the latter correlates further with another adenosine of the A,\(^{15}\)N\(_{lab}\)-RNA. These cross-peaks are compatible with either an AAUG or a GUAAG stretch, and identify unambiguously the spin-system A18–A19–U20–G21. This strategy yielded sequential assignment of 17 out of 19 nucleotides in structured regions, excluding the tetra-loop and the terminal ends. The remaining two nucleotides (G14, A15) were assigned by substitution of the tetra-loop sequence GAAG with UUGC. This alleviated the overlap of G14 and A15 with the resonances of the GAAA loop, allowing their assignment, as well as the identification of G10 and of two adenosines of the A11–A13 stretch. The poor intensity of the GAAA tetra-loop resonances is indicative of conformational heterogeneity; likewise, the terminal G1, U25 and C26 spin-systems are not visible in any of the spectra and were not considered in the structure calculation.

Next, we tested the performance of a \(^{13}\)C,\(^{31}\)P correlation, which, with a TEDOR mixing time of 3.2 ms, should provide sequential C2’/P, i+1 and C3’/P, i+1 contacts. The mixing time was optimized for sensitivity of transfer over two to three bonds, ranging up to 4–5 Å distance (Fig. 1c). As expected, the \(^{31}\)P resonances are poorly resolved in helices and the spectrum provided information only for non A-form structural elements (Fig. 2c, Supplementary Fig. 6a,b).

Finally, we could sequentially assign >90% of all carbon resonances of the Pj Box C/D RNA in the stretches 2–10 and 14–24 (81% for both carbons and nitrogens).

---

**Figure 2 | ssNMR spectra for the sequential assignment and measurement of structural restraints. (a-b) Ribose region of 2D \(^{13}\)C,\(^{15}\)N-TEDOR-\(^{13}\)C,\(^{13}\)C-PDSD spectra of (a), G,\(^{13}\)C\(_{lab}\)-RNA and (b) A,\(^{13}\)C\(_{lab}\)-RNA (mixing time, 700 ms). Intra- and inter-nucleotide correlations are labeled in green and red, respectively. Selected sequential correlations are shown. Partially overlapped guanosines G10, G14, G16 are labeled as G*; non-site-specifically assigned adenosines in the tetra-loop (A11–A13) are labeled as A*. (c) 2D \(^{13}\)C,\(^{31}\)P-TEDOR spectrum of G,\(^{13}\)C\(_{lab}\)-RNA. (d) 2D \(^{15}\)N,\(^{15}\)N-RFDR spectrum showing the G-N1/C-N3 correlations for G.C base pairs.**
Structural determination of the RNA by ssNMR. The determination of RNA secondary structure requires the identification of base pairs. To this end, we used a $^{15}$N,$^{15}$N through-space correlation (RFDR, Radio-Frequency-Driven-Recoupling)$^{17,18}$ to reveal the spatial proximity of either A-N1 and U-N3 or G-N1 and C-N3 in Watson–Crick base pairs (Fig. 1d). The presence of three G:C base pairs (Fig. 2d) defined the C-stem. The G24:C2 base pair was not found due to the absence of the G24-N1 resonance in intra-nucleotide correlations, probably as a consequence of conformational heterogeneity at the helix ends.

Secondary structure prediction suggests one U·U and two A·G base pairs (Fig. 1a). Initially, to verify the presence and determine the topology of these non-canonical base pairs, we measured NHHN spectra$^{19,20}$; in this experiment, magnetization is transferred between close-by $^{15}$N nuclei exploiting the spatial proximity of their attached protons (Fig. 1e). This strategy failed, due to severe overlap of the involved nitrogen resonances. Next, we recorded NHHC spectra (Fig. 1e) on three selectively labelled RNA samples (Supplementary Fig. 7); (G,U) lab-RNA yielded weak N2a2z/C1'g2 correlations, while (A,G) lab-RNA yielded correlations were also detected in a 13C-band-selective, $^{15}$N-TEDOR experiment$^{15}$ recorded for samples with 13C-labelling of one nucleotide type and $^{15}$N-labelling of another nucleotide type (Fig. 1f). Our efforts were unsuccessful, due to low signal-to-noise. However, when recording a ($^{13}$C1',$^{13}$C4')-band-selective, $^{15}$N-TEDOR, we obtained six inter-nucleotide cross-peaks from both the k-turn and helical regions ( Supplementary Fig. 6c,d and Supplementary Table 2).

Finally, 2D NHHC and CHHC spectra$^{19,27}$ yielded 17 and 10 inter-nucleotide contacts (Supplementary Figs 7 and 8), respectively, in addition to 21 intra-nucleotide correlations over $\geq$3 bonds ( Supplementary Table 2).

In addition to distance restraints, we obtained dihedral angles from analysis of ribose chemical shifts (Methods and Supplementary Fig. 2) and from CHHC experiments at short mixing times. Similarly to solution-state NMR, the $\gamma$ angle was restrained to $\text{syn}$ in the presence of a strong C1'-C8/C6 crosspeak (short H'1'-H8/H6 distance) and to $\text{anti}$ in the other cases. Only A19 displayed a $\gamma$ angle in the $\text{syn}$ conformation, in agreement with other k-turn RNA structures$^{11}$.

Structure calculations. Distance and dihedral angle restraints, as well as base pair restraints were used in ARIA$^{28}$ to calculate the structure of the Box C/D RNA from ssNMR data. Out of 300 calculated structures, the first 60 converged to a well-defined minimum with precision of 0.9 Å (root-mean-square-deviation (r.m.s.d.) of all heavy atoms of nucleotides 2–9 and 14–24 of the first 20 structures; Fig. 3a and Table 1). The distances derived from the $^{15}$C,$^{15}$N-TEDOR, $^{13}$C,$^{13}$C-PDSD spectra had the highest impact on the precision of the structure, followed by those derived from the CHHC and NHHC experiments (Supplementary Fig. 9). As for structural calculation from solution-state NMR data, the definition of the RNA secondary structure (topology of base pairs) was essential. The stem regions were defined by eight distance and nine angular restraints per residue, while the geometry of the k-turn required 21 distances per nucleotide. The structure determination method was validated by removal of random fractions of restraints. The structures bundles were consistent upon random removal of up to 20% of the total restraints.

Discussion

The 26mer Box C/D RNA used in this study does not crystallize in complex with L7Ae; however, the crystallographic structure of two orthologous complexes from Archaeoglobus fulgidus$^{29}$ (Af, PDB code 1RLG and 4BW0), one orthologous complex from Sulfolobus solfataricus$^{30}$ (Ss, PDB code 3PLA) and another L7Ae-Box C/D RNA from Pf with a different RNA sequence$^{8}$ (PDB code 3NNU) let us evaluate the accuracy of the ssNMR structure in the critical k-turn region (Fig. 3b,c). The $\phi$ angle of 23° that defines the k-turn geometry of the ssNMR structure is in very good agreement with the $\phi$ angles of the reference structures (1RLG, 23°; 4BW0, 22°; 3PLA, 24°; 3NNU, 24°).

Next, we analysed the backbone and glycoside torsion angles of our structures bundle and compared them with the corresponding torsion angles of the four reference structures (Supplementary Fig. 10). We choose to compare torsion angles rather than r.m.s.d. values to better visualize the variability of both the crystallographic structures and our bundle at each nucleotide position. The $\delta$ torsion describes the ribose pucker and is defined by the chemical shift analysis of Supplementary Fig. 2. The values fit nicely to those of the reference structures, with the exception of 5, 19 and 20 of 1RLG, which adopt the C3'-endo conformation. Our NMR data indicate that the conformation of these riboses is...
Figure 3 | ssNMR structure of the Pf Box C/D RNA. (a) Overlay of the 10 lowest energy structures of the Pf Box C/D RNA in complex with L7Ae from ssNMR data. Terminal nucleotides 1 and 25–26 are not shown. Colour code as in Fig. 1a. (b) k-turn of the Pf Box C/D RNA, showing the characteristic geometry. Internal loop, green; NC stem, cyan; C stem, light blue. (c) Comparison of the k-turn geometry of the Pf Box C/D RNA obtained by ssNMR (10 lowest energy structures, gray) with that of the crystallographic structure of the Aβ Box C/D RNA (PDB code 1RLG)\(^7\), red; Pf Box C/D RNA (PDB code 3NMU)\(^8\), blue; Aβ Box C/D RNA (PDB code 4BWO)\(^29\), green; Ss Box C/D RNA (PDB code 3PLA)\(^30\), magenta.

Table 1 | Structural statistics (20 structures out of 300 calculated, PDB code 2n0r).

| Distance restraints         | 208 | 135 |
|----------------------------|-----|-----|
| Total distance restraints   | 73  | 36  |
| Intra-residue               | 96  | 39  |
| Inter-residue               | 34  | 39  |
| Sequential (|j–j| = 1)      | 174 | 102 |
| Non-sequential (|j–j| > 1) | 18  | 54  |
| Hydrogen bonds              | 18  | 54  |
| Total dihedral angle restraints | 18  | 54  |
| Glycosidic angle \(\gamma\) | 18  | 54  |
| Sugar puck | 18  | 54  |
| Backbone                    | 18  | 54  |
| Based on A-form geometry    | 18  | 54  |

| Structure statistics In vacuum Solvent refined |
|-----------------------------------------------|-------|-------|
| Violations (mean and s.d.) | 0.0043 ± 0.0005 | 0.0077 ± 0.003 |
| Distance constraints (Å) | 0.03 ± 0.02 | 0.28 ± 0.07 |
| Dihedral angle constraints (°) | 0.7 | 3.4 |
| Max. distance constraint violation (Å) | 0.7 | 3.4 |
| Max. dihedral angle violation (°) | 0.7 | 3.4 |

| Deviations from idealized geometry | 0.000182 ± 0.000006 | 0.0023 ± 0.0001 |
|------------------------------------|---------------------|-----------------|
| Bond lengths (Å)                   | 0.461 ± 0.002       | 0.64 ± 0.02     |
| Improper torsion (°)               | 0.319 ± 0.001       | 0.41 ± 0.02     |
| Average pairwise r.m.s.d. (Å)      | 0.9 ± 0.2            | 1.0 ± 0.2       |
| All RNA heavy (2–9,14–24)          | 0.8 ± 0.2            | 1.0 ± 0.2       |
| All RNA backbone (2–9,14–24)       | 0.7 ± 0.2            | 0.8 ± 0.2       |
| Kink-turn backbone (4–6,17–22)     | 0.7 ± 0.2            | 0.8 ± 0.2       |

C2'-endo, in agreement with the other three crystallographic structures. Similarly, the \(\epsilon\) and \(\zeta\) angles of the same nucleotides of 1RLG deviate from the values of both our structures bundle and the other three crystallographic structures. The \(\beta\), \(\epsilon\) and \(\zeta\) torsion angles are not directly determined by any NMR parameter, but rather restrained loosely by data base values (see Methods), \(^{31}\)P–\(^{13}\)C and \(^{13}\)C–\(^{13}\)C distances. Nevertheless, the distribution of these angles in the ssNMR bundle is quite narrow and in good agreement with the reference structures. The \(\alpha\) and \(\gamma\) torsion angles are the least well defined by the NMR distance restraints in the stretch 18–21 of the Box C sequence. Interestingly, high variability is observed for these torsion angles among the four reference structures as well, indicating that the k-turn geometry is tolerant to different values (Supplementary Fig. 10). The only clear discrepancy between the ssNMR structures bundle and the four references structures is observed for A5-\(\alpha\). The A5-\(\alpha\) chemical shift value (Supplementary Table 1) does not allow to restrict the A5-\(\alpha\) to the gauche \(\pm\) conformations\(^31\), as observed in the four reference structures. However, despite this local difference, the k-turn geometry of the ssNMR bundle agrees very well with that of the reference conformations, with an average backbone r.m.s.d. for k-turn nucleotides 4–6 and 17–22 of 1.3 Å to the four crystallographic structures (Fig. 3c). Finally, the glycosidic torsion angles \(\gamma\) are determined from the intensities of the C1'-C8/C6 cross-peaks and nicely agree with those of the reference structures.

In summary, we demonstrate that the structure of RNA is accessible by ssNMR with excellent precision and accuracy, despite the difficulties caused by broad line widths and resonance overlap. We present a straightforward, manageable strategy that uses easy-to-produce nucleotide-type selective-labelled RNAs and sensitive magnetization transfer schemes. Our results make the folding of short RNAs and selectively labelled RNA stretches, as well as their interaction with proteins, accessible at high resolution in the context of large RNAs and RNP particles. We anticipate that our method will have a considerable impact in various fields of RNA processing and small RNA regulation (siRNA, miRNA, piRNA), where the dynamic nature of the molecular complexes represents an obstacle to crystallization.

Methods
Sample preparation. The L7Ae–Box C/D RNA complex was assembled from protein and RNA in 1:1 ratio and purified by size exclusion chromatography. L7Ae was expressed in *Escherichia coli* (LB medium) and purified over a Ni-Nta column. Nucleotide-type \(^{13}\)C, \(^{15}\)N selective-labelled Box C/D RNA was prepared by in vitro transcription with T7 polymerase produced in house. Labeling patterns of the RNA were obtained using NTP mixtures where only one or two nucleotide types were either \(^{15}\)N or \(^{13}\)C, or double \(^{13}\)C,\(^{15}\)N labeled.

Sequential resonance assignment and measurement of structural restraints used eight samples with different labeling patterns. Six samples consisted of double \(^{13}\)C,\(^{15}\)N nucleotide-type selective-labelled RNAs: A,\(^{13}\)C,\(^{15}\)N heavy, A,\(^{15}\)N light, A,\(^{13}\)C,\(^{13}\)C light, A,\(^{15}\)N heavy, A,\(^{15}\)N light, A,\(^{13}\)C,\(^{15}\)N heavy, A,\(^{15}\)N light, A,\(^{13}\)C,\(^{15}\)N light, A,\(^{15}\)N heavy.

Nucleotide-type 13C, 15N selective-labeled Box C/D RNA was prepared by

Nucleotide-type \(^{13}\)C, \(^{15}\)N selective-labelled Box C/D RNA was prepared by in vitro transcription with T7 polymerase produced in house. Labeling patterns of the RNA were obtained using NTP mixtures where only one or two nucleotide types were either \(^{15}\)N or \(^{13}\)C, or double \(^{13}\)C,\(^{15}\)N labeled. Sequential resonance assignment and measurement of structural restraints used eight samples with different labeling patterns. Six samples consisted of double \(^{13}\)C,\(^{15}\)N nucleotide-type selective-labelled RNAs: A,\(^{13}\)C,\(^{15}\)N heavy, A,\(^{13}\)C,\(^{13}\)C light, A,\(^{15}\)N heavy, A,\(^{13}\)C,\(^{15}\)N light, A,\(^{15}\)N heavy, A,\(^{15}\)N light, A,\(^{13}\)C,\(^{15}\)N light, A,\(^{15}\)N heavy.

© 2015 Macmillan Publishers Limited. All rights reserved.
with band-selective 13C inversion pulses; the long mixing allows transferring magnetization between carbons and nitrogens as far as 5–6 Å. The 15N and 13C 15N,15N-RFDR 13C-band-selective, 15N-TEDOR 13C. In 15N,15N-RFDR experiments were performed at 16 kHz MAS, while proton diffusion-based CHHC and NHHC experiments were performed at 13 kHz MAS. NMR spectroscopy. Solid-state NMR experiments were performed on a 700 MHz SB Bruker Avance III spectrometer equipped with 3.2 mm MAS H1/H15N/13C probehead. 13C,13C TEDOR experiments were acquired at 600 MHz with a WB Bruker Avance III spectrometer equipped with a tunable H1/H15N probehead at Bruker Biospin in Rheinstetten. The temperature of all experiments was 260 K. 13C,15N-TEDOR-13C,13C-PDSD, 13C,13C-TEDOR, 13C,15N-TEDOR and 15N,15N-RFDR experiments were performed at 16 kHz MAS, while proton diffusion-based CHHC and NHHC experiments were performed at 13 kHz MAS.

13C,15N-TEDOR-13C,13C-PDSD experiment (Supplementary Fig. 3a) 13C magnetization was prepared by short 1H-13C polarization transfer mixing, 200–700 ms. Then, 13C frequency was recorded during PDSD mixing. Alternatively, we recorded the frequency of nitrogens close to carbons, as for example that of N1/N9 directly bound to C1 and C8/C6. The following, long 13C,15N-PDSD step (mixing time, 200–700 ms) transferred the 15N-chemical shift labelled 13C magnetization to nearby carbons. Finally, 13C magnetization was detected during t2. The ambiguity on the carbon from which the magnetization originates in the PDSD step, either C1 or C8/C6, was lifted in three-dimensional experiments, where the 13C frequency was recorded before the PDSD mixing. Alternatively, we evaluated the efficiency of the 15N/13C 13C transfer, which in several instances was found to be better towards the C1 than towards C8/C6. As a third alternative, a 13C-band-selective TEDOR transfer, with selectivity either on C1 or on C6/C8, can be used to resolve the ambiguity. Cross-peaks were evaluated and translated into distance restraints. Distance ranges (d) were applied for inter-nucleotide restraints as 3.5–<d<9 Å, according to previous studies, the ranges for intra-nucleotide base–ribose restraints, 3<d<6 Å, and intra-nucleotide ribose–ribose restraints, 3<d<4 Å, were determined from the nucleotides’ geometry.

13C,15N-PDSD. In the 13C,15N-PDSD experiment (Supplementary Fig. 3b), after initial preparation of 13C magnetization, the 13C,15N dipolar coupling was introduced in a TEDOR mixing time of 3.2 ms; the frequency of 15N was monitored in t1, while 13C magnetization was detected during t2. Optionally, a short 13C,13C-PDSD step (mixing time, 200–700 ms) was introduced to further carbon spins, such as C1’. This experiment was useful to identify the ribose spin systems through the better-resolved C1 chemical shift. 13C,15N-PDSD spectra were recorded for A10-RNA and G10-U10-RNA; due to the limited signal-to-noise, only one TEDOR mixing time was recorded (3.2 ms). The spectra yielded 13C-TEDOR restraints, which were classified as 3<d<5 Å, as appropriate for a mixing time of 3.2 ms.

13C-band-selective, 15N-TEDOR experiment (Supplementary Fig. 3c), after initial preparation of 15N magnetization, the 13C,15N dipolar coupling was reintroduced in a TEDOR mixing time of 6–15 ms with band-selective 13C inversion pulses; the long mixing allows transferring magnetization between carbons and nitrogens as far as 5–6 Å. The 15N and 13C frequencies were recorded during t1 and t2, respectively. With a 13C,15N-TEDOR that was selective for C1’ and C4’, we obtained four G,C1’/C4’/A,N6,N9 cross-peaks from the 1H-13C,15N-1Hm-NMR experiment. Also in this case, we did not acquire multiple TEDOR mixing times, due to limited signal-to-noise. Distance ranges 3<d<5 Å and 3<d<7 Å were attributed to the strong and weak peaks, respectively, at a mixing time of 12 ms. 15N,15N-RFDR. In the 15N,15N-RFDR experiment (Supplementary Fig. 3d), 15N magnetization was prepared through a 300 μs cross-polarization step and its frequency was recorded during t1; subsequently, the magnetization was transferred to nearby nitrogen atoms via an RFDR mixing step of 20 ms and finally detected during t2.

CHHC and NHHC. In the CHHC and NHHC proton spin diffusion-based experiments (Supplementary Fig. 3e,f, respectively), 13C or 15N magnetization was prepared through a short cross-polarization mixing time of 100–200 μs, followed by t1 evolution on either 13C (CHHC) or 15N (NHHC). Next, the magnetization was transferred back to protons, from where, after a short proton mixing of 100–200 μs, it was transferred to nearby carbons with a 100 μs cross-polarization step; finally the frequency of 13C was recorded in t2. Inter-proton distances 2<d<4 and 2<d<5 Å were attributed to the strong and weak signals, respectively, following previous studies. In all the experiments, protons were decoupled in the indirect and direct acquisition times using high-power SPINAL-64 (ref. 37) decoupling at 85–95 kHz. Chemical shifts were referenced as described by Morcombe and Zilm. The spectra were processed with NMRPipe and visualized with NMRView.

Structural calculation protocol. Structures were calculated using the Aria 1.2/ CNS 1.1 set-up following a similar protocol as for structural calculations of RNA by solution-state NMR data. Both canonical and non-canonical base-pairs were incorporated in the structure calculation as distance restraints. Planarity was conferred through weak planar restraints (5 kcal mol−1 Å−2) for canonical base pairs and non-canonical base pair U3–U23. Flexible planarity was introduced for the base pairs A5–G21 and G4–A22 by defining the plane that involves one atom of the acceptor and four atoms of the donor base to allow for propeller twist and tilt, as described in ref. 44. The ribose conformation was restrained through the analysis of ribose chemical shifts (Supplementary Fig. 2). The riboses of nucleotides G4, A5, A19, U20, G21 were given an S-type conformation, while the remaining nucleotides, except for G6 and A11-A13, were restrained to the N-type conformation. The dihedral angles ω, β, ε and ζ were restrained to the range typical for A-form helix (300°±30°, 180°±30°, 300°±30°, respectively). Dihedral angles ω and ζ of nucleotides G4, G6, A18, U20 and G21 were additionally restrained to 0°±120° backbone chemical shifts and 1H chemical shifts were restrained to the gauche + configuration for nucleotides involved in base pairs. Three hundred structures were calculated in one iteration without the automated assignment or the distance calibration options of Aria 1.2 using an assigned distance list. Before minimization, we randomized all backbone dihedral angles. The minimization protocol used three force-field RNA-DNA-allatom-hi-opls.top and the following parameters in the four steps of simulated annealing (SA), together with the PROLSQ nonbonded parameters: (i) the SA protocol started with a high-temperature torsion angle simulated annealing phase of 100,000 steps at 20,000 K (time step of 22.5 fs); (ii) this was followed by a torsion angle dynamic cooling phase from 20,000 to 1,000 K in 100,000 steps and by two cartesian dynamic cooling phases with a time step of 2.5 fs (iii) from 2,000 to 1,000 K in 100,000 steps and (iv) from 1,000 to 50 K in 80,000 steps, respectively. Finally, 20 low energy structures were refined in water (TIP3P) with OPLS nonbonded parameters. Standard ARIA force constants were used for the different restraint types (for example, distances—50 kcal mol−1, and dihedrals—200 kcal mol−1, in the final cooling step).

The final structures were analysed using Molmol and Chimera. Figures were prepared with Chimera.

Reference

Petkova, A. T. et al. A structural model for Alzheimer’s beta-amyloid fibrils based on experimental constraints from solid state NMR. Proc. Natl Acad. Sci. USA 99, 16742–16747 (2002).

Jaroniec, C. P. et al. High-resolution molecular structure of a peptide in an amyloid fibril determined by angle spinning NMR spectroscopy. Proc. Natl Acad. Sci. USA 101, 711–716 (2004).

de Groot, H. J. M. Solid-state NMR spectroscopy applied to membrane proteins. Curr. Opin. Struct. Biol. 10, 593–600 (2000).

Opella, S. J. & Marassi, F. M. Structure determination of membrane proteins by NMR spectroscopy. Chem. Rev. 104, 3587–3606 (2004).

Shi, L. C., Wang, S., et al. Conformational analysis of a seven-helical transmembrane photosensor in the lipid environment. Angew. Chem. Int. Ed. 50, 1302–1305 (2011).

Marchanka, A. & Carlo magnano, T. Solid-state NMR and RNA structure: a new partnership? eMagRes 3, 119–128 (2014).

© 2015 Macmillan Publishers Limited. All rights reserved.
Falb, M.

Jaroniec, C. P., Filip, C. & Griffin, R. G. 3D TEDOR NMR experiments

Goody, T. A., Melcher, S. E., Norman, D. G. & Lilley, D. M. J. The kink-turn motif in RNA is dimorphic, and metal ion-dependent. RNA 10, 254–264 (2004).

Fall, M. et al. Structure of the K-turn U4 RNA: a combined NMR and SANS study. Nucleic Acids Res. 38, 6274–6285 (2010).

Schoener, K. T., McPhee, S. A., Ouellet, J. & Lilley, D. M. A structural database for k-turn motifs in RNA. RNA 16, 1463–1468 (2010).

Marchang, A., Simon, R. & Carlonmagno, T. A suite of solid-state NMR experiments for RNA resonance assignment in a 21 kDa protein-RNA complex. Angew. Chem. Int. Ed. 52, 2345–2349 (2013).

Cherepanov, A. V., Gluzhitz, C. & Schwab, W. High-resolution studies of uniformly C-13,N-15-labeled RNA by solid-state NMR spectroscopy. Angew. Chem. Int. Ed. 49, 4747–4750 (2010).

Jaroniec, C. P., Filip, C. & Griffin, R. G. 3D TEDOR NMR experiments for the simultaneous measurement of multiple carbon-nitrogen distances in uniformly C-13,N-15-labeled RNA. J. Am. Chem. Soc. 124, 10728–10742 (2002).

Szeverenyi, N. M., Sullivan, M. J. & Maciel, G. E. Observation of spin exchange by two-dimensional Fourier-Transform C-13 cross polarization-magic-angle spinning. J. Magn. Reson. 47, 462–475 (1982).

Bennett, A. E., Ok, J. H., Griffin, R. G. & Vega, S. Chemical-shift correlation spectroscopy in rotating solids—radio frequency-driven dipolar recoupling and longitudinal relaxation. J. Chem. Phys. 96, 8624–8627 (1992).

Leppert, J. et al. Identification of NH...N hydrogen bonds by magic angle spinning solid state NMR spectroscopy. Nucleic Acids Res. 32, 1177–1183 (2004).

Lange, A., Luca, S. & Baldus, M. Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. J. Am. Chem. Soc. 124, 9704–9705 (2002).

Riedel, K. et al. Characterisation of hydrogen bonding networks in RNAs via magic angle spinning solid state NMR spectroscopy. J. Biomol. NMR 31, 331–336 (2005).

Farès, C., Amata, I. & Carlonmagno, T. C-13-Detection in RNA bases: revealing structure-chemical shift relationships. J. Am. Chem. Soc. 129, 15814–15823 (2007).

Oehlerlänger, O. et al. The structure of the stemloop D subdomain of cossackievirus B3 cloverleaf RNA and its interaction with the protein3C. Structure 12, 237–248 (2004).

Lewandowski, J. R., De Paep, G. & Griffin, R. G. Proton assisted insensitive nuclei cross polarization. J. Am. Chem. Soc. 129, 728–729 (2007).

De Paep, G. et al. Proton assisted recoupling and protein structure determination. J. Chem. Phys. 129, 245101 (2008).

Loquet, A. et al. 3D structure determination of the Ccb protein from highly ambiguous solid-state NMR restraints. J. Am. Chem. Soc. 130, 3579–3589 (2008).

Manolikas, T., Herrmann, T. & Meier, B. H. Protein structure determination from C-13 spin-diffusion solid-state NMR spectroscopy. J. Am. Chem. Soc. 130, 3959–3966 (2008).

Lange, A. et al. A concept for rapid protein-structure determination by solid-state NMR spectroscopy. Angew. Chem. Int. Ed. 44, 2089–2092 (2005).

Linge, J. P., Habec, M., Rieping, W. & Nilges, M. ARIA: automated NOE assignment and NMR structure calculation. Bioinformatics 19, 315–316 (2003).

Huang, L. & Lilley, D. M. The molecular recognition of kink-turn structure by the L7Ae class of proteins. RNA 19, 1703–1710 (2013).

Lin, J. et al. Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. Nature 469, 559–563 (2011).

Varani, G., Abou-abl, F. & Allain, F. H. T. NMR investigation of RNA structure. Prog. Nuc. Magn. Reson. Spectrosc. 29, 51–127 (1996).

Jehle, S. et al. Intermolecular protein–RNA interactions revealed by 2D P-31-N-15 magic angle spinning solid-state NMR spectroscopy. J. Am. Chem. Soc. 132, 3842–3846 (2010).

Asami, S., Rakwaska-Bange, M., Carlonmagno, T. & Reif, B. Protein–RNA interfaces probed by 1H-detected MAS solid-state NMR spectroscopy. Angew. Chem. Int. Ed. 52, 2345–2349 (2013).

Wasmer, C. et al. Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. Science 319, 1523–1526 (2008).

Berti, I. et al. High-resolution solid-state NMR structure of a 17.6 kDa protein. J. Am. Chem. Soc. 132, 1032–1040 (2010).

Demers, J. P. et al. High-resolution structure of the Shigella type-III secretion needle by solid-state NMR and cryo-electron microscopy. Nat. Commun. 5, 4976 (2014).

Fung, B. M., Khitirin, A. K. & Ermolaev, K. An improved broadband decoupling sequence for liquid crystals and solids. J. Magn. Reson. 142, 97–101 (2000).

Morcombe, C. R. & Zilm, K. W. Chemical shift referencing in MAS solid state NMR. J. Magn. Reson. 162, 479–486 (2003).

Delaglio, F. et al. Nmrpipe—A multidimensional spectral processing system based on Unix pipes. J. Biomol. NMR 6, 277–293 (1995).

Johnson, B. A. & Blevins, R. A. NMRView: a computer program for the visualization and analysis of NMR data. J. Biomol. NMR 4, 603–614 (1994).

Brunger, A. T. et al. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. Sect. D 54, 905–921 (1998).

Carlonmagno, T. et al. Structural principles of RNA catalysis in a 2’-5’ lariat-forming ribozyme. J. Am. Chem. Soc. 135, 4403–4411 (2013).

Nozovinov, S. et al. High-resolution NMR structure of an RNA model system: the 14-mer eUUCGc tetraloop hairpin RNA. Nucleic Acids Res. 38, 683–694 (2010).

Zanier, K. et al. Structure of the histone mRNA hairpin required for cell cycle regulation of histone gene expression. RNA 8, 29–46 (2002).

Linge, J. P. et al. Refinement of protein structures in explicit solvent. Proteins 50, 496–506 (2003).

Koradi, R., Billette, M. & Wuthrich, K. MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graphics 14, 51–55 (1996).

Pettersen, E. F. et al. UCSF chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612 (2004).

Acknowledgements

This work has been supported by the BMBF through grant 0315870B to T.C.

Author contributions

A.M. produced samples, developed and performed both NMR experiments and structure calculations, interpreted the results and wrote the manuscript together with T.C. B.S. contributed to the development of pulse programs and structural calculation protocols. G.A. helped in the acquisition of NMR data. T.C. designed the project, contributed to the development of pulse programs and structural calculation protocols, interpreted the results and wrote the manuscript together with A.M. All the authors discussed the results reported in the manuscript.

Additional information

Accession codes: The atomic coordinates of 10 lowest energy structures have been deposited in the Protein Data Bank under accession number 2NNR. The NMR chemical shifts have been deposited in the Biological Magnetic Resonance Data Bank, entry 25534.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Marchanka, A. et al. RNA structure determination by solid-state NMR spectroscopy. Nat. Commun. 6:7024 doi: 10.1038/ncomms8024 (2015). This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/