Structure Determination

DOI: 10.1002/ange.200600769

Constraints on the Structure of $(CUG)_{97}$ RNA from Magic-Angle-Spinning Solid-State NMR Spectroscopy**

*Kerstin Riedel, Christian Herbst, Sabine Häfner, Jörg Leppert, Oliver Ohlenschläger, Maurice S. Swanson, Matthias Görlach, and Ramadurai Ramachandran**

Expansions of short nucleotide sequence repeats are associated with a number of neuromuscular diseases.[1–3] CTG triplet repeat expansions in the 3′ noncoding region of the myotonic dystrophy protein kinase (DMPK) gene give rise to transcripts harboring CUG triplet repeat expansions on the RNA level. Repeats containing >50 CUG triplets cause in *trans*-dominant fashion the most frequent form of adult-onset muscular dystrophy (DM1). The current model for DM1 pathogenesis strongly suggests that such repeats fold into large stable double-stranded RNA hairpins, which bind and sequester muscleblind (MBNL) proteins that are involved in the alternative splicing of a number of pre-mRNAs. As a consequence, the MBNL proteins are unavailable to the splicing machinery, and a number of important muscular pre-mRNAs, for example, for the chloride channel ClC-1 and the insulin receptor, are aberrantly spliced; this process ultimately leads to the clinical manifestations of DM1.

To decipher the structural basis of DM1, which could potentially permit the development of suitable drugs that would interfere with the sequestration of MBNL proteins by CUG repeats, we have recently initiated, as a first step, a magic-angle-spinning (MAS) solid-state NMR study of a ≈100-kDa RNA composed of 97 CUG repeats ($(CUG)_{97}$). ^{15}N-^{15}N chemical-shift correlation experiments have enabled us to show the presence of canonical GC base pairs in this RNA.[4] In addition, analysis of the observed ^{13}C chemical shifts for the sugar carbon atoms suggested that $(CUG)_{97}$ adopts an A-form helix conformation with a C3′-*endo* sugar pucker and an *anti* conformation of the glycosidic torsion angle χ.[5] Herein, we have explored the possibility of obtaining structural information directly by exploiting the dependence of inter- and intranucleotide ^1H–^1H distances on RNA conformation. Since the extraction of ^1H–^1H distances by direct ^1H observation techniques is still difficult in the solid state, we have utilized in this study the potential of proton–proton dipolar-coupling-mediated chemical-shift correlation spectroscopy of low-gamma nuclei[6–13] for mapping the spatial proximity of the sugar and aromatic protons in $(CUG)_{97}$.

The radio frequency (RF) pulse sequence[13] and a schematic representation of the double-stranded $(CUG)_{97}$ employed are shown in Figure 1. Longitudinal ^1H magnetization exchange mediated by proton–proton dipolar coupling is allowed to take place during the spin diffusion period τ_{mix}. The experiment is carried out with a very short CP contact time (≈100 μs) and a proton spin diffusion mixing time, τ_{mix}, of 100–200 μs. This time regime minimizes relayed magnetization transfers during the CP and τ_{mix} periods and, hence, crosspeaks with appreciable intensities are expected only between proton-attached ^{13}C sites that are connected by ^1H–^1H distances of less than ≈3 Å.[8–13] ^{13}C homonuclear isotropic chemical-shift correlation spectra of a uniformly {^{15}N,^{13}C}-labeled sample of $(CUG)_{97}$ obtained with different short proton spin diffusion mixing times are shown in Figure 2. The resonance assignments indicated are based on our recent studies.[5,14] The expected crosspeaks arising from direct or relayed magnetization transfers originating from sugar proton–sugar proton dipolar interactions and corresponding to intranucleotide proton–proton distances of <3.0 Å are seen. By contrast, crosspeaks between the aromatic and sugar C1′ and C4′ carbon atoms, which would reflect the spatial proximity of the corresponding attached protons, are essentially absent.

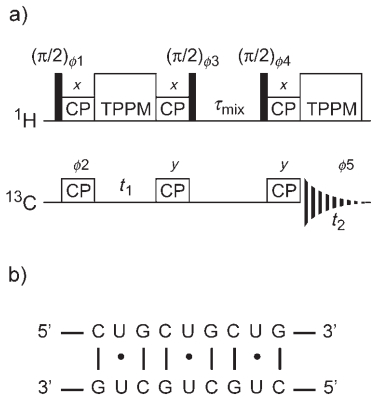

Figure 1. a) CPMAS pulse sequence employed for proton spin diffusion based isotropic chemical-shift correlation of the protonated carbon sites in RNA. The RF pulse and the receiver phases were cycled as $\phi1 = (y, -y)$, $\phi2 = (-y, -y, y, y)$, $\phi3 = (4\times y, 4\times -y)$, $\phi4 = (8\times y, 8\times -y)$, and $\phi5 = (x, -x, -x, x, -x, x, x, -x, -x, x, x, -x, x, -x, -x, x)$. High-power two-pulse phase modulation (TPPM) decoupling was employed in the t_1 and t_2 time periods and phase-sensitive 2D spectra were generated as described by States et al.[18] b) Schematic representation of a section of the double-stranded $(CUG)_{97}$ RNA.

[*] K. Riedel, C. Herbst, S. Häfner, Dr. J. Leppert, Dr. O. Ohlenschläger, Dr. M. Görlach, Dr. R. Ramachandran
Research Group Molecular Biophysics/NMR Spectroscopy
Leibniz Institute for Age Research
Fritz Lipmann Institute
Beutenbergstrasse 11, 07745 Jena (Germany)
Fax: (+49) 3641-656-225
E-mail: raman@fli-leibniz.de
Prof. Dr. M. S. Swanson
Department of Molecular Genetics and Microbiology
Powell Gene Therapy Center
College of Medicine
University of Florida
1600 SW Archer Road, Gainesville, FL 32610-0266 (USA)

[**] This study has been funded in part by a PhD fellowship to Kerstin Riedel from the Stiftung Stipendien-Fonds des Verbandes der Chemischen Industrie e.V. and a grant from the Deutsche Forschungsgemeinschaft (grant no.: GO474/6-1).

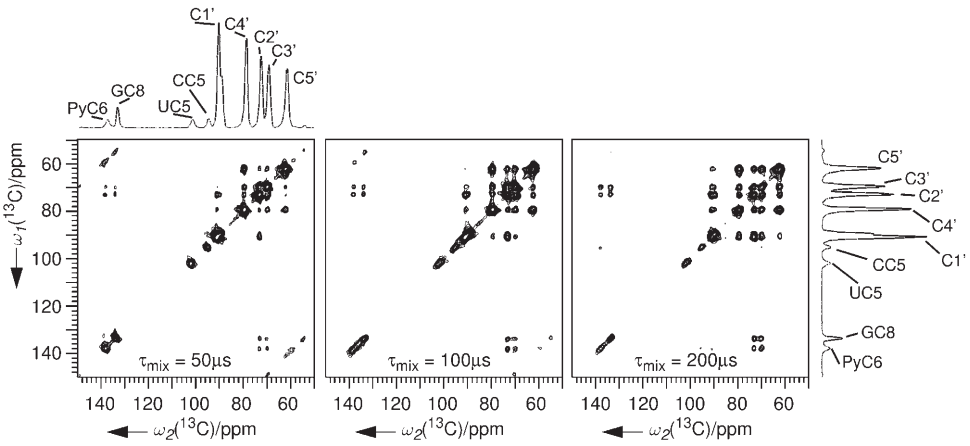

Figure 2. Experimental ^{13}C-^{13}C chemical-shift correlation spectra of the hydrated, uniformly {^{15}N,^{13}C}-labeled $(CUG)_{97}$ RNA. Mixing times were as indicated; spinning speed: 10000 Hz; data acquisition in the direct dimension: 10 ms; CP contact time: 75 µs; recycle time: 1 s; ω_1 spectral width: 30000 Hz; 160 t_1 increments with 64 transients per increment were employed. The RF carrier was kept at the center of the spectral region. Chemical shifts are referenced with respect to external adamantane (δ_{CH} = 29.5 ppm with respect to tetramethylsilane). The assignments of the different resonances are indicated on the spectral projection shown.

It is known that in A-form helical RNA (Figure 3) sequential H1′···H6/H8 distances are larger than 3 Å.[15] In a proton spin diffusion based ^{13}C-^{13}C chemical-shift correlation spectrum, this would be reflected by the absence of the corresponding crosspeaks between the aromatic and sugar C1′ carbon atoms. By contrast, B-form helices are characterized by short internucleotide H1′···H5′/H5″ distances of about 2 Å. Hence, the absence of strong crosspeaks between the C1′ and C5′ carbon atoms with a τ_{mix} period of 50 µs, similar to that between the C1′ and C2′ carbon atoms, is consistent with an A-form helix. By using standard RNA conformations, that is, C3′-*endo* and C2′-*endo* sugar puckers and various torsion angles χ, we have evaluated the intranucleotide distances between aromatic and sugar protons with molecular modeling (Table 1). From these results, it can be deduced that the absence of crosspeaks between the aromatic and sugar C1′ carbon atoms indicates a χ angle, describing the relative orientation of the base and the sugar moiety across the glycosidic bond, in the *anti* range.

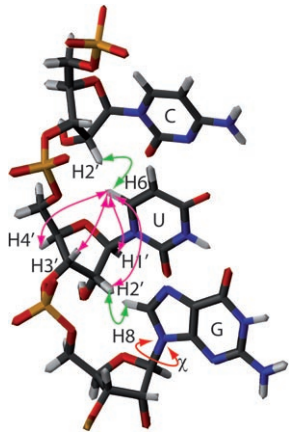

Figure 3. Representative intra- and internucleotide ^{1}H–^{1}H dipolar interactions in a double-stranded A-form helical RNA. For clarity, only one strand of the RNA, modelled with SYBYL software, is shown.

In the spectra shown in Figure 2, crosspeaks arising from dipolar interactions involving aromatic (H6/H8) and sugar (H2′ and H3′) protons can be clearly observed. In a uniformly labeled sample these crosspeaks could, in principle, arise both from intra- and internucleotide ^{1}H–^{1}H dipolar interactions. Since A-form RNA helices are characterized by internucleotide H2′···H8/H6 and H3′···H8/H6 distances in the range of 1.5–2.5 Å and 2.5–4 Å,[15] respectively, the corresponding crosspeaks seen in Figure 2 can arise from sequential interactions. However, it is equally possible that when the χ angle is in the *anti* range, one could also observe crosspeaks arising from intranucleotide correlations (Table 1).

To identify the origin of the sugar–aromatic crosspeaks seen in Figure 2, we have carried out further experiments employing a $(CUG)_{97}$ sample that has been specifically labeled with [^{13}C]U. Figure 4 shows the correlation spectrum for this sample with a τ_{mix} period of 200 µs. In this case all the crosspeaks arising from internucleotide ^{1}H–^{1}H dipolar interactions are not detectable due to specific U labeling, so crosspeaks in the spectrum of such a sample can only arise from intranucleotide ^{1}H–^{1}H dipolar interactions. No such crosspeaks between the aromatic C6 and sugar C2′ carbon atoms are detected (Figure 4). This indicates that the corresponding crosspeak in Figure 2 can only arise from sequential, internucleotide ^{1}H–^{1}H dipolar interactions, as

Table 1: Distance ranges for H6/H8 to sugar protons as extracted from models of riboG and riboU generated with SYBYL software.

χ	for C2′-*endo*				for C3′-*endo*			
	Purine H8/Pyrimidine H6 to							
	H1′	H2′	H3′	H4′	H1′	H2′	H3′	H4′
0°, *syn*	2.7–3.1	3.2–3.5	5.2–5.5	5.6–6.0	2.7–3.1	2.5–2.6	4.3–4.5	5.4–5.7
45°	2.3–2.7	4.1–4.3	5.7–6.1	5.5–5.9	2.3–2.7	3.6–3.7	5.1–5.2	5.4–5.8
90°, high *syn*	2.4–2.8	4.6–4.8	5.9–6.3	5.0–5.4	2.4–2.7	4.3–4.4	5.1–5.3	5.1–5.6
135°	2.8–3.2	4.4–4.7	5.7–6.1	4.4–4.9	2.9–3.2	4.5–4.7	4.5–4.8	4.6–5.1
180°, *anti*	3.4–3.7	3.7–4.1	5.1–5.6	4.1–4.7	3.5–3.7	4.0–4.4	3.4–3.8	4.2–4.6
225°	3.7–4.0	2.7–3.1	4.4–4.9	4.3–4.8	3.7–3.9	3.1–3.6	2.2–2.6	4.2–4.5
270°, high *anti*	3.6–3.9	1.7–2.3	4.0–4.6	4.8–5.3	3.6–3.9	1.9–2.5	2.1–2.3	4.5–4.8
315°	3.3–3.6	2.0–2.5	4.4–4.9	5.4–5.8	3.2–3.6	1.5–2.0	3.1–3.3	5.0–5.3

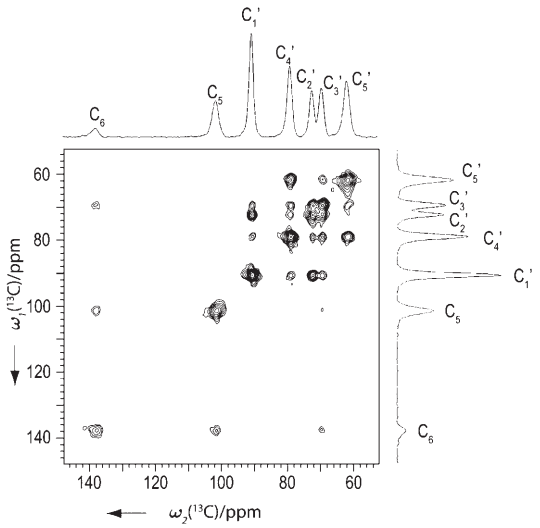

Figure 4. Experimental ^{13}C-^{13}C chemical-shift correlation spectrum of the hydrated [^{13}C]U-labeled $(CUG)_{97}$ RNA sample. Mixing time: 200 μs; CP contact time: 100 μs; 160 t_1 increments with 512 transients per increment were employed. For further details see the legend to Figure 2.

expected for A-form helical RNAs. The crosspeak between the aromatic C6 and sugar C3′ carbon atoms (Figure 4) indicates a C3′-*endo* sugar conformation for U. With the χ angle in the *anti* range and in a C2′-*endo* conformation, the H6···H3′ distance is always larger than that between the H6 and H2′ protons and this would result first in the appearance of the crosspeak between the aromatic C6 and sugar C2′ carbon atoms. This is not the case here, even with lower mixing times (data not shown), and this again suggests a C3′-*endo* sugar conformation for U, as expected for A-form helices.

The data presented herein provide further clear evidence for the conclusions deduced from ^{13}C chemical-shift analysis, namely, that $(CUG)_{97}$ principally adopts an A-form helix conformation with a C3′-*endo* sugar pucker and an *anti* glycosidic torsion angle χ. They also demonstrate that strong dipolar interactions between spatially proximal ^{1}H nuclei can be effectively exploited for extracting RNA conformation constraints from ^{13}C MAS solid-state NMR spectroscopy. On the other hand, and as pointed out earlier by Lange et al.,[8] short-range distance measurements by using low-gamma nuclei typically do not provide structural information. The present approach is also applicable even for the study of uniformly labeled samples. To minimize the effects of conformation averaging/flexibility, it would be advantageous to carry out these experiments at the lowest temperature possible. Further studies will have to be carried out to obtain additional structural constraints for generating a well-defined MAS-NMR-spectroscopically derived 3D structure of $(CUG)_{97}$. However, the A-form helix conformation of $(CUG)_{97}$, as deduced from our earlier studies and as further substantiated herein, has meanwhile been corroborated with an X-ray crystal study on an 18-base-pair RNA containing 6 CUG repeats.[16]

Although to reduce data acquisition time we have carried out only two-dimensional ^{13}C-^{13}C chemical-shift correlation experiments, it should be possible to extend this approach to three dimensions, for example, to obtain $^{15}N/^{13}C$-edited ^{13}C-^{13}C correlation spectra. Such experiments would be useful for resolving spectral overlaps. The approach adopted herein can also be extended to study systems by using different labeling schemes. Conceptually similar to the approach used in structure determinations of RNA molecules by liquid-state NMR spectroscopy,[17] in which a fast classification of the predominant χ conformation is achieved by an analysis of the intensities of the H1–H8/H6 NOE crosspeaks, the analysis of proton spin diffusion induced ^{13}C-^{13}C chemical-shift correlation spectra in the short-mixing-time regime should allow the characterization of the torsion angle χ, the sugar pucker, and the helical regions of large RNA systems through MAS solid-state NMR spectroscopy. While proton–proton dipolar couplings have been successfully employed recently for the extraction of structural constraints from MAS solid-state NMR studies of peptides and proteins, distance estimates involving nonexchangeable protons have been exploited herein for the first time in the study of nucleic acids. Although structural studies of RNAs are typically carried out with solution-state NMR spectroscopy, this study clearly indicates that even large RNA systems are, in principle, amenable for investigations with MAS solid-state NMR spectroscopy.

Experimental Section

RNA samples were prepared by using appropriately labeled ribonucleoside triphosphates as described earlier.[4] A hydrated, undiluted {^{15}N,^{13}C}-labeled RNA sample of $(CUG)_{97}$[4] (8.5 mg) and a hydrated, specifically [^{13}C]U-labeled $(CUG)_{97}$ sample (5 mg) were used. Experiments were carried out at about −15 °C on a 500-MHz wide-bore Varian Unity Inova solid-state NMR spectrometer equipped with a 5-mm DOTY supersonic triple-resonance probe.

Received: February 28, 2006
Revised: May 15, 2006
Published online: July 20, 2006

Keywords: helical structures · magic-angle spinning · NMR spectroscopy · RNA · structure elucidation

[1] K. R. Nykamp, M. S. Swanson, *Prog. Mol. Subcell. Biol.* **2004**, *35*, 57–77.
[2] L. P. W. Ranum, J. W. Day, *Am. J. Hum. Genet.* **2004**, *74*, 793–804.
[3] L. P. W. Ranum, J. W. Day, *Trends Genet.* **2004**, *20*, 506–512.
[4] J. Leppert, C. R. Urbinati, S. Häfner, O. Ohlenschläger, M. S. Swanson, M. Görlach, R. Ramachandran, *Nucleic Acids Res.* **2004**, *32*, 1177–1183.
[5] K. Riedel, J. Leppert, S. Häfner, O. Ohlenschläger, M. Görlach, R. Ramachandran, *J. Biomol. NMR* **2004**, *30*, 389–395.
[6] M. Wilhelm, H. Feng, U. Tracht, H. W. Spiess, *J. Magn. Reson.* **1998**, *134*, 255–260.
[7] I. de Boer, L. Bosman, J. Raap, H. Oschkinat, H. J. M. de Groot, *J. Magn. Reson.* **2002**, *157*, 286–291.
[8] A. Lange, S. Luca, M. Baldus, *J. Am. Chem. Soc.* **2002**, *124*, 9704–9705.

[9] A. Lange, K. Seidel, L. Verdier, S. Luca, M. Baldus, *J. Am. Chem. Soc.* **2003**, *125*, 12640 – 12648.
[10] A. Lange, S. Becker, K. Seidel, L. Verdier, K. Giller, O. Pongs, M. Baldus, *Angew. Chem.* **2005**, *117*, 2125 – 2129; *Angew. Chem. Int. Ed.* **2005**, *44*, 2089 – 2092.
[11] R. Tycko, Y. Ishii, *J. Am. Chem. Soc.* **2003**, *125*, 6606 – 6607.
[12] B. Reif, B. J. van Rossum, F. Castellani, K. Rehbein, A. Diehl, H. Oschkinat, *J. Am. Chem. Soc.* **2003**, *125*, 1488 – 1489.
[13] K. Riedel, J. Leppert, O. Ohlenschläger, M. Görlach, R. Ramachandran, *J. Biomol. NMR* **2005**, *31*, 331 – 336.
[14] K. Riedel, J. Leppert, O. Ohlenschläger, M. Görlach, R. Ramachandran, *J. Biomol. NMR* **2005**, *31*, 49 – 57.
[15] B. Furtig, C. Richter, J. Wöhnert, H. Schwalbe, *ChemBioChem* **2003**, *4*, 936 – 962.
[16] B. H. M. Mooers, J. S. Logue, J. A. Berglund, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16626 – 16631.
[17] P. B. Moore, *Acc. Chem. Res.* **1995**, *28*, 251 – 256.
[18] D. J. States, R. A. Haberkorn, D. J. Ruben, *J. Magn. Reson.* **1982**, *48*, 286 – 292.

