# Application of NMR Spectroscopy to Determine the 3D Structure of Small Non-Coding RNAs

Marie-Eve Chagot, Marc Quinternet, Xavier Manival, Isabelle Lebars

# Chapter ?

## Application of NMR spectroscopy to determine the 3D structure of small non-coding RNA

**Marie-Eve Chagot[1], Marc Quinternet[2], Xavier Manival[1*] and Isabelle Lebars[3*]**

[1] CNRS, IMoPA, Université de Lorraine, Nancy, F-54000, France

[2] CNRS, INSERM, IBSLOR, Université de Lorraine, Nancy, F-54000, France

[3] Université de Strasbourg, CNRS, ARN UPR 9002, 2 allée Conrad Roentgen, F-67000 Strasbourg, France

*To whom correspondence should be addressed:

Dr. Isabelle Lebars: i.lebars@ibmc-cnrs.unistra.fr (Tel.: +33 (0) 3 88 41 70 02)

Dr. Xavier Manival: xavier.manival@univ-lorraine.fr (Tel.: +33(0) 3 72 74 66 79)

## Abstract

Many RNA architectures were discovered to be involved in a wide range of essential biological processes in all organisms from carrying genetic information to gene expression regulation. The remarkable ability of RNAs to adopt various architectures depending on their environment enables the achievement of their myriads of biological functions. Nuclear Magnetic Resonance (NMR) is a powerful technique to investigate both their structure and dynamics. NMR is also a key tool for studying interactions between RNAs and their numerous partners such as small molecules, ions, proteins or other nucleic acids.

In this chapter, to illustrate the use of NMR for 3D structure determination of small non-coding RNA, we describe detailed methods that we used for the yeast C/D box small nucleolar RNA U14 from sample preparation to 3D structure calculation.

**Key words**: RNA, NMR, 3D structure, dynamics

# 1. Introduction

In the living world, RNA is the only molecule, which carries genetic information, functions like an enzyme, nucleates and guides the assembly of complex molecular machines and regulates multitude of cellular processes. These manifold competencies are the foundations of an autonomous system based solely on RNA, the RNA world [1]. The catalytic functions of RNA are involved in the three basic cellular processes of life: replication, transcription and translation, where it is found most of the time in association with proteins. The recent new roles of RNA predominantly reside in the regulation of gene expression by the discovery of non-coding RNAs (silencing or siRNA, micro or miRNA, Piwi-interacting or piRNA, long non-coding or lncRNA, small nucleolar or snoRNA, small nuclear or snRNA...) and others should be discovered during coming years [2]. The remarkable ability of RNAs to fold into various architectures enables the achievement of their multiple biological functions. Most RNAs in eukaryotic cells do not act in isolation, but rather exist in association with proteins to form what we call RNPs (RiboNucleoProtein) complexes.

Nuclear Magnetic Resonance (NMR) is the most powerful experimental method to investigate the structure and dynamics of small RNAs (size < 50 kDa) [3–6]. The first three-dimensional (3D) structure of a free RNA in solution determined by NMR dates from the early 90s [7]. The partially ordered 3D structure and conformational plasticity of RNA, which is a hallmark of its function, make it difficult to crystallize. In addition, small RNAs cannot be analyzed by single-particle cryo-electron microscopy (cryo-EM), a technique dedicated to high molecular weight molecules (> 150 kDa). Investigating the 3D structure of RNA at high-resolution and the dynamics of its complex transitions are essential for a complete mechanistic understanding of its numerous cellular regulations. Studying at atomic scale the dynamics and the conformational exchanges of a molecule is

particularly well adapted to solution-state NMR. NMR is also a key tool for following, at the atomic level, interactions between RNAs and their various partners such as small molecules, ions, proteins or other nucleic acids [8] or the transition between multiple RNA conformational states [9]. Recently, the high-resolution structure of RNA by solid-state NMR has emerged as an alternative approach based on nucleotide-type selectively labeled samples [10]. Over the past 10 years, NMR spectroscopy has also become a method of choice for studying post-translational modifications (PTM) in proteins [11]. Similarly, time-resolved NMR measurements provided ways to follow epitranscriptome (chemical modifications on multiple nucleotides in RNA) into cells extracts [12].

NMR structural studies require the preparation of milligram amounts of RNA oligonucleotides. Large quantities of unlabeled and $^{13}C/^{15}N$ labeled RNA are generally synthesized from DNA templates by *in vitro* transcription using T7 RNA polymerase and requires optimizations, in particular for magnesium concentration [13]. RNA transcripts often contain one or more extra nucleotides at their 3'-end [13]. Nevertheless, various methods have been developed to reduce this heterogeneity. The insertion of an autocatalytic ribozyme at the end of the gene allows precise self-cleavage leading to a uniform 3'-end [14–16]. The substitution of the last two nucleotides or the penultimate nucleotide on the DNA template by a 2'-methoxyribose sugar can improve the quality of the transcript [17]. Then, the desired small RNA can be purified either by denaturing polyacrylamide gel electrophoresis (PAGE) or by denaturing high pressure liquid chromatography (HPLC) [18, 19]. After the production of unlabeled and $^{13}C/^{15}N$ labeled RNA samples, NMR spectra are collected and NMR frequencies are assigned for all atoms ($^{1}H$, $^{13}C$, $^{15}N$ and $^{31}P$). The assignment step can be challenging for high molecular weight RNA due to frequencies overlapping and line widths broadening increasing with the size of the RNA. However, biochemical and NMR methods have been developed to

overcome these limitations [4, 6, 19–23]. Assignment of frequencies is essential to interpret chemical shifts values, nuclear overhauser effect (NOE) correlations, coupling constant values and residual dipolar coupling values that form together a set of distance and dihedral angle restraints. All these constraints permit to calculate a set of 3D structure in a simulated annealing protocol. In this chapter, we describe detailed methods that we used for the yeast C/D box snoRNA U14 from sample preparation to 3D structure calculation.

## 2. Materials

All solutions are prepared using ultrapure water (water purified on resin filters followed by deionization to achieve a resistivity of 18.2 MΩ.cm at 25°C). Materials have to be RNase-free [24].

### *2.1. In vitro transcription and purification of RNA NMR sample*

1. DNA templates at 100 μM in water, stored at -20°C (See **Fig.1**).
2. Nucleotides triphosphates (ATP, UTP, CTP, GTP) dissolved individually in water and titrated to pH 7.0 with NaOH, stored at -20°C at a concentration of 100 mM (See **Note 1**).
3. T7 RNA polymerase "in-house" (T7 RNAP) in 20 mM sodium phosphate pH 7.7, 1 mM DTT, 1 mM EDTA, 100 mM NaCl and 50% glycerol; stored in aliquots at -20°C.
4. 50% (w/v) polyethylene glycol (PEG-8000) stored at -20°C in single use aliquots.
5. 1 M $MgCl_2$ stored in aliquots at -20°C.
6. 10X T7 RNAP transcription buffer stored in single use aliquots at -20°C: 400 mM Tris-HCl pH 8.1, 10 mM spermidine, 0.1% (v/v) Triton X-100, 50 mM DTT [13, 25].
7. EDTA 0.5 M, pH 8.0, stored at 4°C.
8. Phenol:Chloroform saturated solution at pH 4.7 stored at 4°C.
9. Ammonium acetate 3 M, pH 5.5, stored at 4°C.
10. 100% ethanol stored at -20°C.
11. 10X TBE running buffer: 0.9 M Tris-borate, 10 mM EDTA, pH 8.3, stored at room temperature.
12. Denaturing acrylamide gel solution: 20% Acrylamide/Bisacrylamide (19:1), 0.5X TBE and 7 M urea.
13. 7 M urea solution.
14. N,N,N,N'-tetramethyl-ethylenediamine (TEMED).

15. 10% (w:v) Ammonium persulfate solution.

16. 6X loading dye: 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol blue and 30% glycerol.

17. PAGE equipment.

18. Electroelution apparatus (Elutrap Electrophoresis Unit, Schleicher & Schuell).

19. Elution buffer: TBE 0.5X.

20. Dialysis tubes (MW cut-off 3 500).

21. NMR buffer: 50 mM sodium phosphate, pH 6.4.

### *2.2 NMR studies*

1. NMR spectrometers equipped with a z-gradient TBI probe and a z-gradient TXI cryoprobe.
2. 3 mm NMR microtubes.
3. Deuterium oxide 100% (Eurisotop).
4. Pf1 filamentous bacteriophages supplied in 10 mM sodium phosphate buffer at pH 6.5 (ASLA Biotech Ltd).
5. Topspin software (Bruker).
6. NMRFAM-SPARKY software [26].

### *2.3 Structure calculations and analysis*

1. CNS software (Crystallography and NMR system) [27].
2. PALES software [28].
3. PyMOL software [29].
4. 3DNA software [30].

## 3. Methods

NMR structural studies require the preparation of milligram amounts of RNA [4, 5]. Most unlabeled and $^{13}C/^{15}N$ labeled RNAs are synthesized from DNA templates by *in vitro* transcription using T7 RNA polymerase [13, 25, 31]. Small RNA samples can also be produced by solid-phase chemical synthesis methods which is more expensive [32, 33]. Here, we detail first the preparation of the U14 snoRNA NMR sample using *in vitro* transcription. Secondly, we describe the strategy for NMR resonances assignment, which is mostly based on the analysis NOE contacts combined to heteronuclear experiments.

### 3.1. *In vitro transcription and purification of RNA NMR samples*

After optimization, the protocol below is described at a 1.0 [$MgCl_2$]/[NTPs] ratio in 5 mL reaction volume.

1. Mix 2.57 mL of water and 1 mL of 50% PEG 8000 in 15 mL Falcon tube.
2. Add 500 µL of 10X T7 RNAP transcription buffer.
3. Add 80 µL of 1M $MgCl_2$.
4. Add 200 µL of each NTP.
5. Shake slowly the reaction mixture.
6. Add 15 µL of each DNA strand (stock: 100 µM).
7. Add 20 µL of T7 RNAP (final concentration: 0.04 mg/mL).
8. Incubate the reaction solution at 37°C for 4 h 30.
9. Quench by addition of 500 µL of 0.5 M EDTA pH 8.0.
10. Mix.
11. Add 2.5 mL of Phenol/Chloroform pH 4.7.
12. Vortex.
13. Centrifuge 15 min at room temperature at 8 063 x *g*.

14. Transfer the aqueous phase in a 50 mL Falcon tube.
15. Add 1 mL of water to the remaining Phenol fraction.
16. Vortex.
17. Centrifuge 15 min at room temperature at 8 063 x *g*.
18. Add the aqueous phase to the first one in the 50 mL Falcon tube.
19. Add 0.1 volume of Ammonium acetate and 2.5 volumes of cold 100% ethanol.
20. Vortex.
21. Place at -20°C overnight to precipitate.
22. Centrifugation at 8 063 x *g* for 30 min at 4°C.
23. Remove the supernatant gently.
24. Resuspend the pellet in 800 µL of a 7 M urea solution supplemented by 160 µL of 6X loading dye.
25. Prepare a 20% polyacrylamide preparative denaturing (7 M urea) gel.
26. Pre-run the gel at a constant power of 40 W in 0.5X TBE running buffer for 30 min before loading.
27. Rinse the wells thoroughly with 0.5X TBE solution.
28. Load the RNA sample.
29. Run the gel at a constant power of 40 W in 0.5X TBE running buffer until the Bromophenol blue reaches to the bottom of the gel.
30. Remove the plates from the gel.
31. Place the gel on a clean cellophane over a silica chromatography plate.
32. Visualize RNA by $UV_{256}$ shadowing.
33. Excise the expected RNA transcript from the gel with a clean scalpel.
34. Elute the RNA from the gel in 0.5X TBE buffer by using an Elutrap Electrophoresis Unit.
35. Recover the RNA in the sample chamber every hour, until the absorbance decreases.

36. Combine the RNA containing fractions.
37. Add 0.1 volume of 3 M Ammonium acetate pH 5.5 and 2.5 volumes of cold 100% ethanol.
38. Precipitate at -20°C overnight.
39. Centrifuge at 8 063 x *g* for 30 min at 4°C.
40. Dialyzed against the NMR buffer.
41. Concentrate the sample by lyophilization.
42. Resuspend in 90/10 $H_2O/D_2O$ in 150 µL for experiments involving exchangeable protons and in 99.9% $D_2O$ for experiments involving non-exchangeable protons.
43. Transfer the sample in a 3 mm NMR tube.

### 3.2. NMR experiments and RNA resonance assignment

All NMR data are processed using Topspin software (Bruker) and analyzed with NMRFAM-SPARKY software [26].

#### 3.2.1. Exchangeable and partial non-exchangeable resonance assignment

NMR experiments described hereafter are recorded in 90/10 $H_2O/D_2O$.

1. 1D spectra over a spectral width of 20 ppm at low temperature (from 4 to 15°C). Imino protons in Watson-Crick base-pair are observable between 12 and 15 ppm whereas imino protons involved non-canonical base-pair are found between 9 and 12 ppm.
2. 2D $^{1}H$-$^{1}H$ NOESY experiments (spectral width of 20 ppm) at low temperature with varying mixing times (See **Note 2**).
3. Assign sequentially the imino protons from U and G residues. U-H3 are discriminated from G-H1 by the strong correlation observed between U-H3 and A-H2 within a Watson-Crick A:U pair. In a G:C pair, two strong NOEs are observable between G-H1 and the cytosine amino protons. In A-helical form, NOE cross-peaks are observable between

imino protons of neighboring base-pairs. Using U-H3 as a starting point, imino protons can be assigned. Assignments are confirmed by 2D $^{1}H$-$^{15}N$ HSQC and 2D $^{1}H$-$^{15}N$ HNN-COSY [35, 36] (See **Fig.2** and **Note 3**).

4. Assign H2 and amino protons of adenine and cytosine, respectively (See **Fig.3**).

### 3.2.2. Non-exchangeable resonance assignment

NMR experiments described hereafter are recorded in 99.9 % $D_2O$ at 25°C.

1. 2D $^{1}H$-$^{1}H$ NOESY experiments (spectral width of 8 ppm) with several mixing times (50, 150, 300 and 400 ms).
2. 2D $^{1}H$-$^{1}H$ TOCSY and 2D $^{1}H$-$^{1}H$ DQF-COSY experiments: sugar puckering and identification of H5-H6 cross-peaks from pyrimidine (See **Note 4**).
3. 2D-HCN experiment: intranucleotide correlation between H1' and H6/H8 protons.
4. 2D $^{1}H$-$^{13}C$ HSQC: aromatic and ribose correlations (See **Note 5**).
5. Assign sequentially H6/H8 and H1' in the 2D NOESY experiment (See **Fig.4**): in an A-type helix, aromatic H6/8 are correlated to their intranucleotide H1' and to the H1' of the preceding nucleotide ($H1'_{n-1}$ <-> $H6/8_n$ <-> $H1'_n$). By combining the analysis of 2D TOCSY experiment, H6 are distinguished from H8. H2 are distinguished from H8 by the analysis of the aromatic 2D $^{1}H$-$^{13}C$ HSQC (See **Note 5**). Assignment of H1' are confirmed by the 2D-HCN experiment (See **Fig.5**).
6. 3D $^{1}H$-$^{13}C$ HCCH TOCSY: identification of individual sugar spin systems H1', H2', H3', H4' and H5'/H5" (See **Fig.6**).
7. Report all assignment in 2D $^{1}H$-$^{1}H$ NOESY experiments.
8. 2D $^{1}H$-$^{31}P$ HP-COSY experiment: derive β angle restraints from $^{1}H_{5'-5"}$ – $^{31}P$ coupling constants (See **Note 6**).
9. 3D $^{1}H$-$^{13}C$-$^{31}P$ HCP experiment: derive ε angle restraints from $^{3}J(P_i\text{-}C4'_i)$, $^{3}J(P_i\text{-}C5'_i)$ and

$^{3}J(P_i\text{-}C2'_{i-1})$ coupling constants (See **Fig.7** and **Note 6**).

### 3.2.3. Residual Dipolar Coupling constants (RDC) measurement

10. Acquire 2D $^{1}H\text{-}^{15}N$ IPAP-HSQC and 2D $^{1}H\text{-}^{13}C$ IPAP-HSQC.
11. Add Pf1 filamentous bacteriophages (See **Note 7**).
12. Acquire the same 2D $^{1}H\text{-}^{15}N$ IPAP-HSQC and 2D $^{1}H\text{-}^{13}C$ IPAP-HSQC experiments in the presence of Pf1.
13. Calculate RDCs as the difference between couplings measured for isotropic and partially aligned samples.

## 3.3. Structure calculations

Restraint files used for 3D structure calculations will be written in the CNS format.

1. Create hydrogen bond restraints file from base-pairs identified in 2D $^{1}H\text{-}^{1}H$ NOESY and 2D $^{1}H\text{-}^{15}N$ HNN-COSY:

```
assign (resid 2 and name N1)    (resid 30 and name N3)     2.90   0.2   0.5
assign (resid 2 and name O6)    (resid 30 and name N4)     2.90   0.2   0.5
assign (resid 2 and name N2)    (resid 30 and name O2)     2.90   0.2   0.5
assign (resid 2 and name H1)    (resid 30 and name N3)     1.90   0.2   0.5
assign (resid 2 and name O6)    (resid 30 and name H42)    1.90   0.2   0.5
assign (resid 2 and name H22)   (resid 30 and name O2)     1.90   0.2   0.5
                                          etc .....
```

2. Classify NOE cross-peaks into four distance bond ranges: strong (1.8 – 3.0 Å), medium (3.0 – 4.5 Å), weak (4.0 – 5.5 Å) and very weak (4.5 – 7.0 Å).

3. Create a distance restraints file:

```
assign (resid 2 and name H8)    (resid 2 and name H1')     3.50   0.5   0.5
assign (resid 2 and name H8)    (resid 2 and name H2')     4.00   0.5   0.5
assign (resid 2 and name H8)    (resid 2 and name H3')     3.00   0.5   0.5
assign (resid 2 and name H8)    (resid 1 and name H1')     4.50   0.5   0.5
assign (resid 2 and name H8)    (resid 1 and name H2')     2.50   0.5   0.5
assign (resid 2 and name H8)    (resid 1 and name H3')     3.00   0.5   0.5
assign (resid 2 and name H8)    (resid 1 and name H8)      4.50   0.5   0.5
etc .....
```

4. Create a dihedral angle restraints file (See **Note 6**):

```
assign (resid 2 and name C4') (resid 2 and name O4') (resid 2 and name C1') (resid 2 and name C2') 1 6 20 2
assign (resid 2 and name O4') (resid 2 and name C1') (resid 2 and name C2') (resid 2 and name C3') 1 -25 20 2
```

```
assign (resid 2 and name C1') (resid 2 and name C2') (resid 2 and name C3') (resid 2 and name C4') 1 37 20 2
assign (resid 2 and name C2') (resid 2 and name C3') (resid 2 and name C4') (resid 2 and name O4') 1 -37 20 2
assign (resid 2 and name C3') (resid 2 and name C4') (resid 2 and name O4') (resid 2 and name C1') 1 21 20 2
assign (resid 2 and name O3') (resid 3 and name P) (resid 3 and name O5') (resid 3 and name C5') 1 -65 20 2
                                        etc …..
```

5. Calculate 100 structures with CNS software including the distance and dihedral angle restraints files created in points 1, 3 and 4 using a torsion angle molecular dynamics (TAMD) protocol optimized for nucleic acids (See **Fig.8**).

6. Create a RDC restraints file:

```
assign (resid 500 and name OO) (resid 500 and name Z) (resid 500 and name X) (resid 500 and name Y) (resid 2
and name C8) (resid 2 and name H8) -9.35 5.0
assign (resid 500 and name OO) (resid 500 and name Z) (resid 500 and name X) (resid 500 and name Y) (resid 4
and name C8) (resid 4 and name H8) 1.50 5.0
assign (resid 500 and name OO) (resid 500 and name Z) (resid 500 and name X) (resid 500 and name Y) (resid 5
and name C6) (resid 5 and name H6) 17.28 5.0
assign (resid 500 and name OO) (resid 500 and name Z) (resid 500 and name X) (resid 500 and name Y) (resid 6
and name C8) (resid 6 and name H8) 11.13 5.0
                                        etc …..
```

7. Using the PALES software, estimate Axial (Da) and Rhombic (R) components of the aligned tensor from the best fit of measured RDCs to the calculated prefolded structures with zero violation on NOE distances and dihedral angles.

8. Run a grid search to determine optimal Da and R [37].

9. Refine the 100 prefolded structures with Da and R optimal values including all the restraint files (distance, dihedral angle and RDC restraints).

10. Select the “ouput” structures presenting no restraints violation and best fits with RDCs and refine them using the nucleic acid force field to include optimized potentials for liquid simulation (opls), charges and non-bonded parameters [38].

11. Calculate quality factors (Q) and root-mean-square deviations (r.m.s) between input and calculated RDCs with PALES software.

12. Sort 3D structures according to best Q, r.m.s. deviations and with no restraints violations, and select the 10 best structures.

13. Analyze converged structures with Pymol and 3DNA software packages.

## Notes

### Note 1

Each NTP is dissolved individually in water and the pH is adjusted to 7.0 using 1 M NaOH, at a concentration of 100 mM. rNTPs are stored at -20°C. $^{13}C/^{15}N$ labeled rNTPs are commercially available at a concentration of 100 mM (SIGMA-ALDRICH USA or Eurisotop Europe).

### Note 2

2D $^{1}H$-$^{1}H$ NOESY recorded in 90/10 $H_2O/D_2O$ allows the visualization of imino-imino correlations between 9 and 15 ppm. The analysis provides the sequential assignment of imino protons.

### Note 3

Uracil and guanine imino protons are easily discriminated in $^{1}H$-$^{15}N$ HSQC as N1(G) chemical shift is found between 145-150 ppm while N3(U) resonates at 160-165 ppm. Hydrogen bonds can be directly observed in $^{1}H$-$^{15}N$ HNN-COSY, which allows the discrimination of imino protons when resonances overlap and/or are involved in non-canonical pairing.

### Note 4

Nucleotides with no cross-peaks in 2D $^{1}H$-$^{1}H$ TOCSY and in 2D $^{1}H$-$^{1}H$ DQF COSY experiments are restrained to a *C3'-endo* conformation with $\nu_0 = 6 \pm 15°$, $\nu_1 = -25 \pm 15°$, $\nu_2 = 37 \pm 15°$, $\nu_3 = -37 \pm 15°$, $\nu_4 = 21 \pm 15°$.

### Note 5

The aromatic 2D $^{1}H$-$^{13}C$ HSQC allows the discrimination of C6 and C8 from C2 that resonate about 15 ppm downfield. The ribose 2D $^{1}H$-$^{13}C$ HSQC allows to distinguish C1' (89-95 ppm), C4' (81-86 ppm) and C5' (62-70 ppm). C2' and C3' resonate at almost the same chemical shifts (70-80 ppm). Nevertheless, H2' and H3' protons can be discriminated by the analysis of 2D NOESY spectra. In A-form helix, strong sequential NOE $H2'_{i-1}$-$H6/8_{i}$ are observable in 2D NOESY spectrum recorded with a short mixing time (50 ms). H3' protons can be identified by the two strong sequential NOEs $H3'_{i-1}$-$H6/8_{i}$-$H3'_{i}$ also observable in NOESY experiments.

**Note 6**

Sequential assignment of adjacent residues can be obtained by the analysis of 3D $^{1}H$-$^{13}C$-$^{31}P$ HCP experiment. $H3'_{i}C3'_{i}$ and $H4'_{i}C4'_{i}$ are correlated to $P_{i+1}$ of the phosphodiester backbone. $H5'H5''_{i}C5'_{i}$ and $H4'_{i}C4'_{i}$ are correlated to $P_{i}$ of the phosphodiester backbone. In addition, β and ε angles restraints can be derived from this NMR experiment for structure calculations. The coupling constant $^{3}J(P_{i}$-$C4'_{i})$ varies from 1 to 11 Hz. Low values (≈ 1 Hz) are associated to a *gauche* - (*g*-) or *gauche* + (*g*-) β angle conformations. Interestingly, when β angle is in *g*- conformation, the $^{3}J(P_{i}$-$H5'_{i})$ or $^{3}J(P_{i}$-$H5''_{i})$ increases to 23 Hz. With a β angle in *trans* conformation, the $^{3}J(P_{i}$-$C4'_{i})$ coupling constant ranges from 7 to 11 Hz. The $^{3}J(P_{i+1}$-$C4'_{i})$ decreases from 11 to 1 Hz when the ε angle changes from ε in *trans* conformation to ε in *gauche* - (*g*-) conformation. Simultaneously, the $^{3}J(P_{i+1}$-$C2'_{i})$ increases from 1 to 11 Hz when ε in *trans* conformation changes to ε in *g*- conformation.

**Note 7**

Partial alignment of the U14 snoRNA was achieved by added Pf1 filamentous bacteriophages

supplied in 10 mM phosphate buffer at pH 6.5 (ASLA Biotech Ltd) inducing a splitting of the deuterium solvent line of 11 Hz.

**References**


1. Joyce GF (2002) The antiquity of RNA-based evolution. Nature 418:214–221
2. Hannon GJ (2002) RNA interference. Nature 418:244–251
3. Allen M, Varani L, Varani G (2001) Nuclear magnetic resonance methods to study structure and dynamics of RNA-protein complexes. Meth Enzymol 339:357–376
4. Latham MP, Brown DJ, McCallum SA, Pardi A (2005) NMR Methods for Studying the Structure and Dynamics of RNA. ChemBioChem 6:1492–1505
5. Fürtig B, Richter C, Wöhnert J, Schwalbe H (2003) NMR Spectroscopy of RNA. ChemBioChem 4:936–962
6. Getz M, Sun X, Casiano-Negroni A, Zhang Q, Al-Hashimi HM (2007) NMR studies of RNA dynamics and structural plasticity using NMR residual dipolar couplings. Biopolymers 86:384–402
7. Cheong C, Moore PB (1992) Solution structure of an unusually stable RNA tetraplex containing G- and U-quartet structures. Biochemistry 31:8406–8414
8. Chagot M-E, Quinternet M, Rothé B, Charpentier B, Coutant J, Manival X, Lebars I (2019) The yeast C/D box snoRNA U14 adopts a "weak" K-turn like conformation recognized by the Snu13 core protein in solution. Biochimie 164:70–82
9. Zhao B, Zhang Q (2015) Characterizing excited conformational states of RNA by NMR spectroscopy. Current Opinion in Structural Biology 30:134–146
10. Marchanka A, Simon B, Althoff-Ospelt G, Carlomagno T (2015) RNA structure determination by solid-state NMR spectroscopy. Nature Communications. https://doi.org/10.1038/ncomms8024
11. Smith MJ, Marshall CB, Theillet F-X, Binolfi A, Selenko P, Ikura M (2015) Real-time NMR monitoring of biological activities in complex physiological environments. Current Opinion in Structural Biology 32:39–47
12. Barraud P, Gato A, Heiss M, Catala M, Kellner S, Tisné C (2019) Time-resolved NMR monitoring of tRNA maturation. Nature Communications. https://doi.org/10.1038/s41467-019-11356-w
13. Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res 15:8783–8798
14. Ferre-D'Amare AR, Doudna JA (1996) Use of Cis- and Trans-Ribozymes to Remove 5' and 3' Heterogeneities From Milligrams of In Vitro Transcribed RNA. Nucleic Acids Research 24:977–978
15. Grosshans CA, Cech TR (1991) A hammerhead ribozyme allows synthesis of a new form of the *Tetrahymena* ribozyme homogeneous in length with a 3' end blocked for transesterification. Nucleic Acids Research 19:3875–3880
16. Avis JM, Conn GL, Walker SC (2013) Cis-Acting Ribozymes for the Production of RNA In Vitro Transcripts with Defined 5′ and 3′ Ends. In: Conn GL (ed) Recombinant and In Vitro RNA Synthesis. Humana Press, Totowa, NJ, pp 83–98
17. Kao C, Rüdisser S, Zheng M (2001) A Simple and Efficient Method to Transcribe RNAs with Reduced 3′ Heterogeneity. Methods 23:201–205
18. Summer H, Grämer R, Dröge P (2009) Denaturing Urea Polyacrylamide Gel Electrophoresis (Urea PAGE). Journal of Visualized Experiments. https://doi.org/10.3791/1485
19. Duss O, Lukavsky PJ, Allain FH-T (2012) Isotope Labeling and Segmental Labeling of Larger RNAs for NMR Structural Studies. In: Atreya HS (ed) Isotope labeling in Biomolecular NMR. Springer Netherlands, Dordrecht, pp 121–144
20. Dominguez C, Schubert M, Duss O, Ravindranathan S, Allain FH-T (2011) Structure

determination and dynamics of protein–RNA complexes by NMR spectroscopy. Progress in Nuclear Magnetic Resonance Spectroscopy 58:1–61
21. Yadav DK, Lukavsky PJ (2016) NMR solution structure determination of large RNA-protein complexes. Progress in Nuclear Magnetic Resonance Spectroscopy 97:57–81
22. Schlundt A, Tants J-N, Sattler M (2017) Integrated structural biology to unravel molecular mechanisms of protein-RNA recognition. Methods 118–119:119–136
23. Barnwal RP, Yang F, Varani G (2017) Applications of NMR to structure determination of RNAs large and small. Archives of Biochemistry and Biophysics 628:42–56
24. Nielsen H (2011) Working with RNA. In: Nielsen H (ed) RNA. Humana Press, Totowa, NJ, pp 15–28
25. Wyatt JR, Chastain M, Puglisi JD (1991) Synthesis and purification of large amounts of RNA oligonucleotides. BioTechniques 11:764–769
26. Lee W, Tonelli M, Markley JL (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics 31:1325–1327
27. Brünger AT, Adams PD, Clore GM, et al (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 54:905–921
28. Zweckstetter M, Hummer G, Bax A (2004) Prediction of charge-induced molecular alignment of biomolecules dissolved in dilute liquid-crystalline phases. Biophys J 86:3444–3460
29. Schrödinger, LLC (2015) The PyMOL Molecular Graphics System, Version 1.8.
30. Colasanti AV, Lu X-J, Olson WK (2013) Analyzing and building nucleic acid structures with 3DNA. J Vis Exp e4401
31. Milligan JF, Uhlenbeck OC (1989) Synthesis of small RNAs using T7 RNA polymerase. Meth Enzymol 180:51–62
32. Usman N, Ogilvie KK, Jiang MY, Cedergren RJ (1987) The automated chemical synthesis of long oligoribuncleotides using 2'-O-silylated ribonucleoside 3'-O-phosphoramidites on a controlled-pore glass support: synthesis of a 43-nucleotide sequence similar to the 3'-half molecule of an Escherichia coli formylmethionine tRNA. Journal of the American Chemical Society 109:7845–7854
33. Ogilvie KK, Usman N, Nicoghosian K, Cedergren RJ (1988) Total chemical synthesis of a 77-nucleotide-long RNA sequence having methionine-acceptance activity. Proceedings of the National Academy of Sciences 85:5764–5768
34. Wincott F, DiRenzo A, Shaffer C, Grimm S, Tracz D, Workman C, Sweedler D, Gonzalez C, Scaringe S, Usman N (1995) Synthesis, deprotection, analysis and purification of RNA and ribozymes. Nucleic Acids Res 23:2677–2684
35. Dingley AJ, Grzesiek S (1998) Direct Observation of Hydrogen Bonds in Nucleic Acid Base Pairs by Internucleotide $^{2}J_{NN}$ Couplings. Journal of the American Chemical Society 120:8293–8297
36. Hennig M, Williamson JR (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
37. Clore GM, Gronenborn AM, Tjandra N (1998) Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. J Magn Reson 131:159–162
38. Jorgensen WL, Tirado-Rives J (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. J Am Chem Soc 110:1657–1666

**Figure Legends**

**Figure 1.**

DNA template sequences for the *in vitro* transcription of the U14 snoRNA.

**Figure 2.**

Imino proton region of one-dimensional $^{1}H$ (up) and 2D $^{1}H$-$^{15}N$ HNN-COSY (bottom) spectra recorded at 25°C. On the right, the secondary structure of the terminal conserved boxes C and D from U14 snoRNA is represented with non-natural nucleotides indicated in outline.

**Figure 3.**

Region of a 2D $^{1}H$-$^{1}H$ NOESY experiment recorded at 10°C with a 50 ms mixing time showing U-H3/A-H2 and G-H1/C-H4 correlations. On the left, the observed NOEs in Watson-Crick base-pairs are indicated with blue arrows.

**Figure 4.**

H6/8-H1' region of a 2D $^{1}H$-$^{1}H$ NOESY spectrum recorded at 25°C in 100% $D_2O$ with a mixing time of 400 ms. Dashed lines in blue and purple represent sequential assignment ($H1'_{n-1}$ <-> $H6/8_n$ <-> $H1'_n$) identified for the U14 snoRNA. H6-H5 correlations are indicated for U11, U17 and C18 residues.

**Figure 5.**

The 2D-HCN experiment correlates H1' and H6/8 protons within a residue. The region of the spectrum corresponding to U nucleotides is shown.

**Figure 6**.

Sugar spin system assignment of residue U24. **(A)** Plane extracted at 3.62 ppm from the 3D $^{1}$H-$^{13}$C HCCH TOCSY recorded at 25°C showing $^{1}$H-$^{13}$C correlations. **(B)** Planes extracted from the 3D $^{1}$H-$^{13}$C HCCH TOCSY showing the sequential assignment of the sugar protons of U24. Each plane is visualized at a proton frequency indicated on the left.

**Figure 7**.

Planes extracted at $\delta^{1}$H = 4.12 ppm and 4.62 ppm from the 3D $^{1}$H-$^{13}$C HCCH TOCSY recorded at 25°C showing $^{1}$H-$^{31}$P correlations for U24 and G25 residues. β angles were restrained to the *trans* conformation for U24 and G25 residues ($^{3}$J(C4i'-Pi)). ε angle was restrained to the *gauche* $^{-}$ conformation for U24 (low value of $^{3}J(C4_{i}'\text{-}P_{i+1})$).

**Figure 8**.

Structure calculations performed with CNS using a torsion angle molecular dynamics (TAMD) protocol for nucleic acids. Structures are generated from an extended strand using distance, dihedral angle, hydrogen bond and RDC restraints.

**Acknowledgments**

This work was supported by CNRS, University of Strasbourg and University of Lorraine. We acknowledge the NMR Core Facility of UMS2008/US40 IBSLor (Université de Lorraine, CNRS, INSERM), F-54000 Nancy, France.

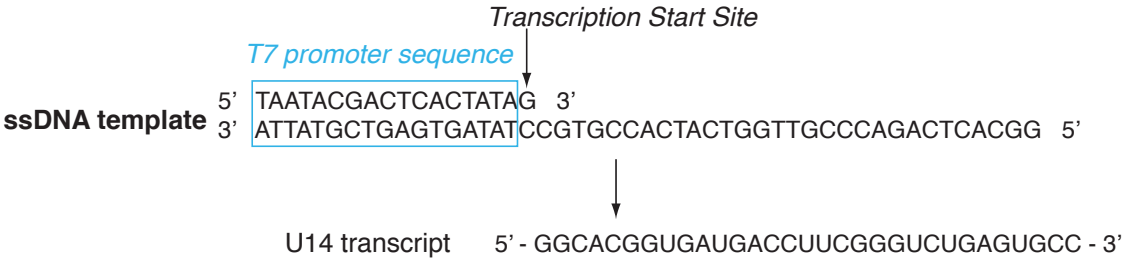

Transcription Start Site
T7 promoter sequence
ssDNA template
5’ TAATACGACTCACTATAG 3’
3’ ATTATGCTGAGTGATATCCGTGCCACTACTGGTTGCCCAGACTCACGG 5’
U14 transcript
5’ - GGCACGGUGAUGACCUUCGGGUCUGAGUGCC - 3’

**Figure 1**

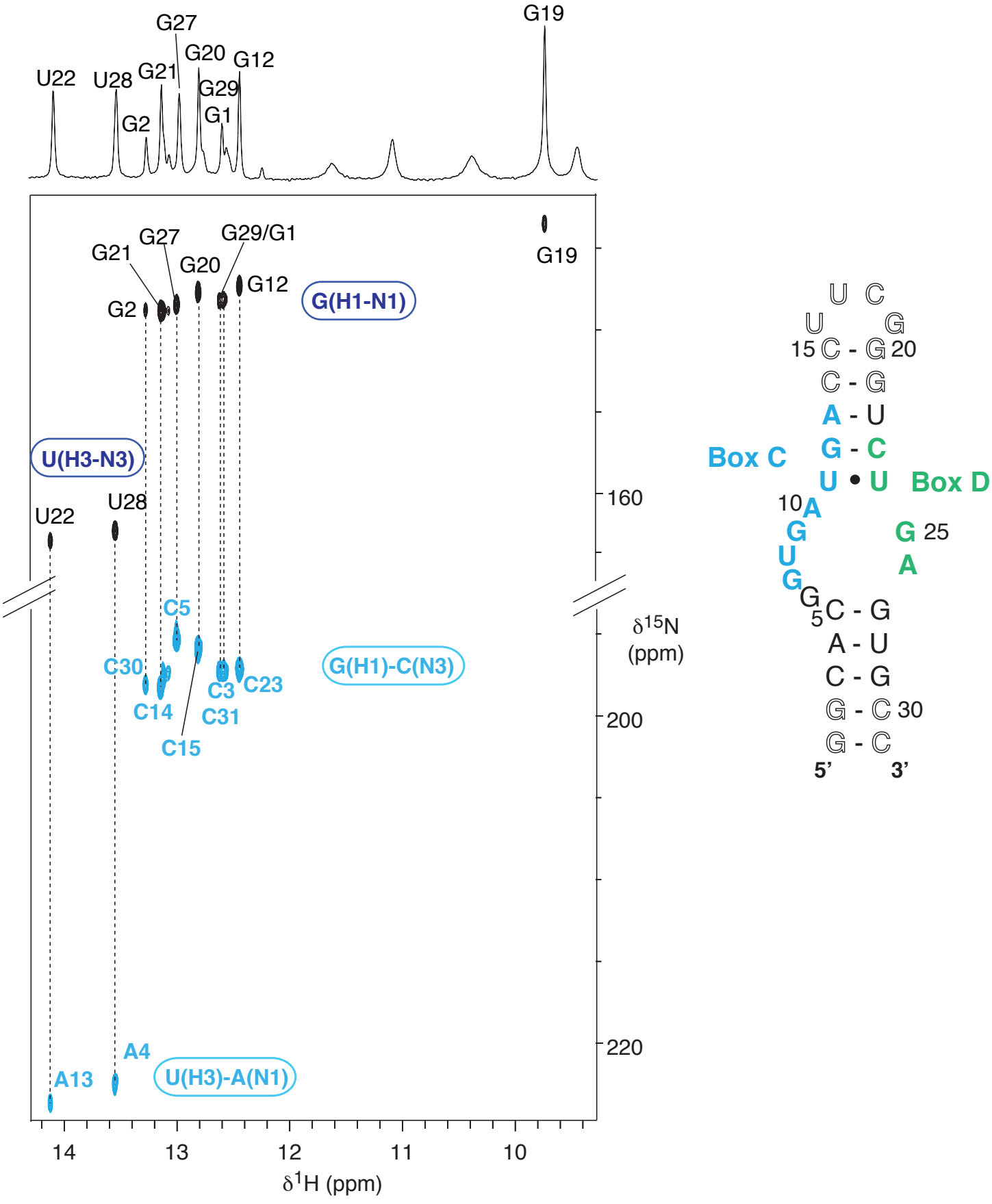

G19
G27
G20
G12
U22
U28
G21
G29
G1
G2
G29/G1
G21
G27
G20
G12
G2
G19
G(H1-N1)
U(H3-N3)
U22
U28
C5
C30
C14
C3
C23
C31
C15
G(H1)-C(N3)
A4
A13
U(H3)-A(N1)
160
200
220
δ15N (ppm)
14
13
12
11
10
δ1H (ppm)
U C
U G
15 C - G 20
C - G
A - U
G - C
U • U
Box C
Box D
10 A
G
G 25
U
A
G
G
5 C - G
A - U
C - G
G - C 30
G - C
5'
3'

**Figure 2**

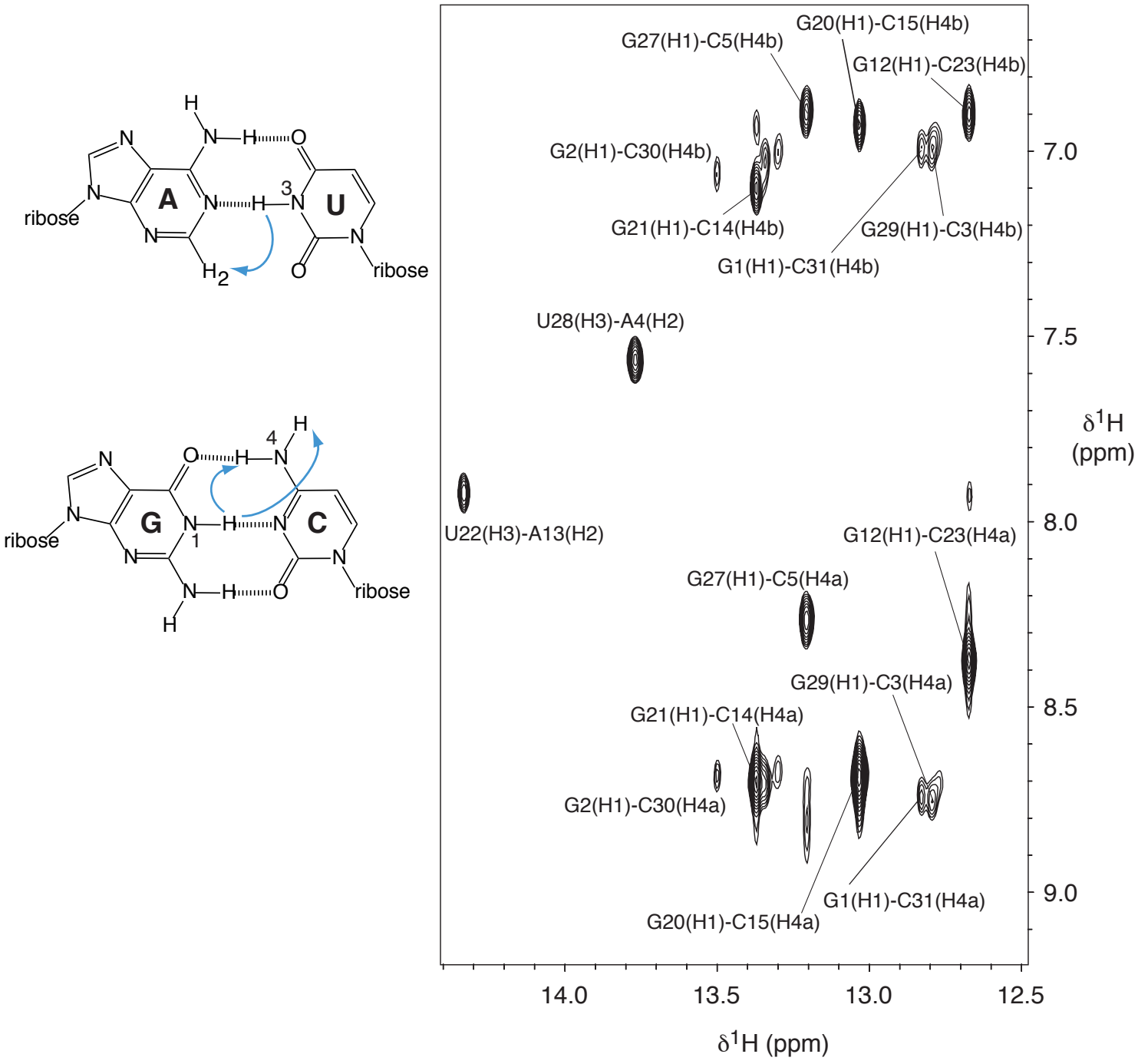

A
U
G
C
ribose
ribose
ribose
ribose
H2
1
3
4
G20(H1)-C15(H4b)
G27(H1)-C5(H4b)
G12(H1)-C23(H4b)
G2(H1)-C30(H4b)
G21(H1)-C14(H4b)
G29(H1)-C3(H4b)
G1(H1)-C31(H4b)
U28(H3)-A4(H2)
U22(H3)-A13(H2)
G12(H1)-C23(H4a)
G27(H1)-C5(H4a)
G29(H1)-C3(H4a)
G21(H1)-C14(H4a)
G2(H1)-C30(H4a)
G1(H1)-C31(H4a)
G20(H1)-C15(H4a)
7.0
7.5
8.0
8.5
9.0
δ1H (ppm)
14.0
13.5
13.0
12.5
δ1H (ppm)

**Figure 3**

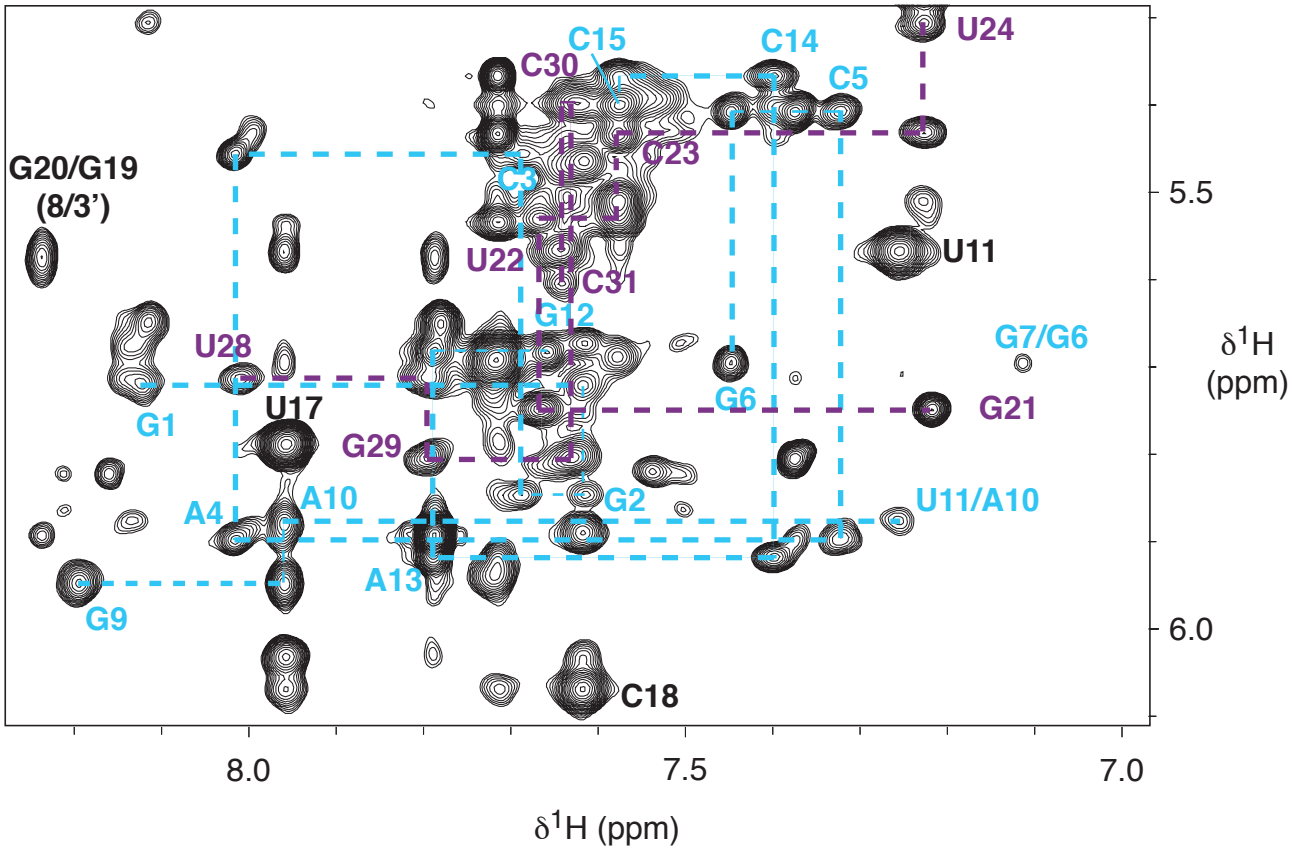

C15
C30
C14
U24
C5
G20/G19
(8/3')
C23
C3
U11
U22
C31
G12
U28
G7/G6
G6
G21
U17
G1
G29
A10
U11/A10
A4
G2
A13
G9
C18
5.5
6.0
δ1H
(ppm)
8.0
7.5
7.0
δ1H (ppm)

**Figure 4**

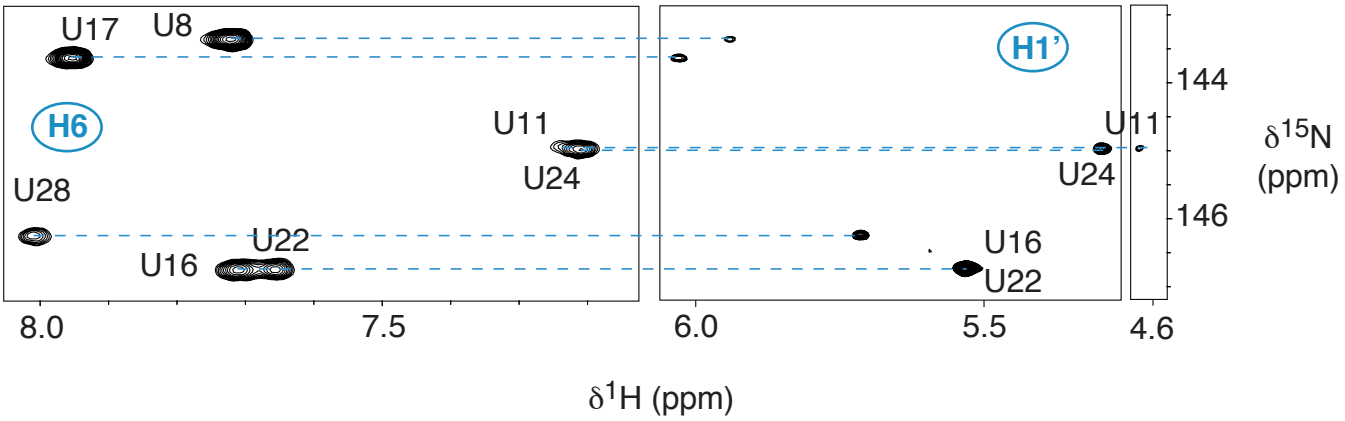
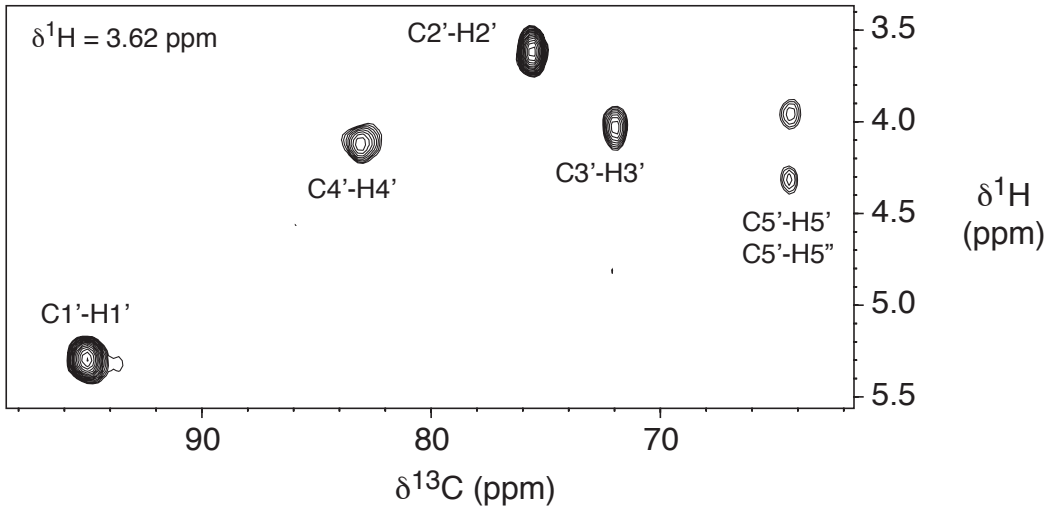

U17
U8
H6
U28
U11
U24
U22
U16
H1'
U11
U24
U16
U22
144
146
δ15N
(ppm)
8.0
7.5
6.0
5.5
4.6
δ1H (ppm)

**Figure 5**

**(A)**


δ1H = 3.62 ppm
C2'-H2'
C4'-H4'
C3'-H3'
C5'-H5'
C5'-H5"
C1'-H1'
3.5
4.0
4.5
5.0
5.5
δ1H (ppm)
90
80
70
δ13C (ppm)

**(B)**

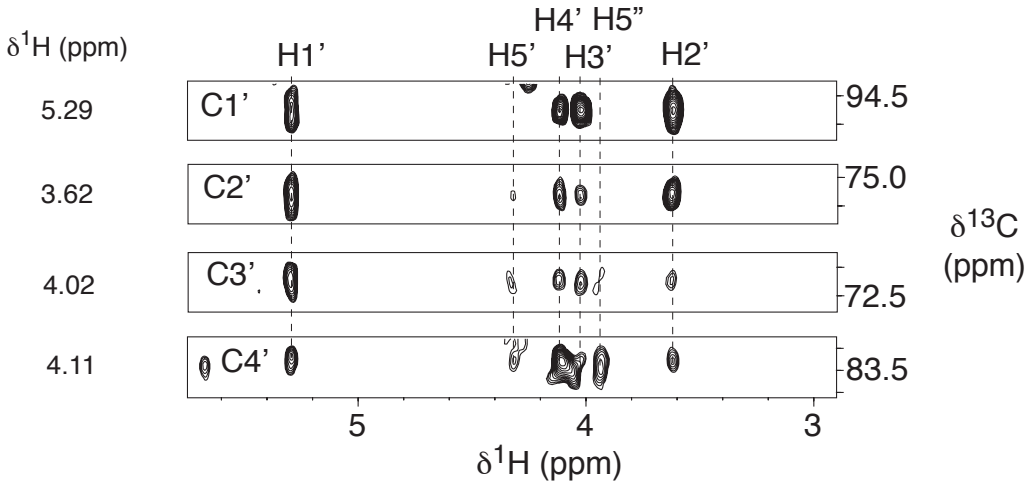

δ1H (ppm)
H4' H5"
H1'
H5'
H3'
H2'
5.29
3.62
4.02
4.11
C1'
C2'
C3'
C4'
94.5
75.0
72.5
83.5
δ13C (ppm)
5
4
3
δ1H (ppm)

**Figure 6**

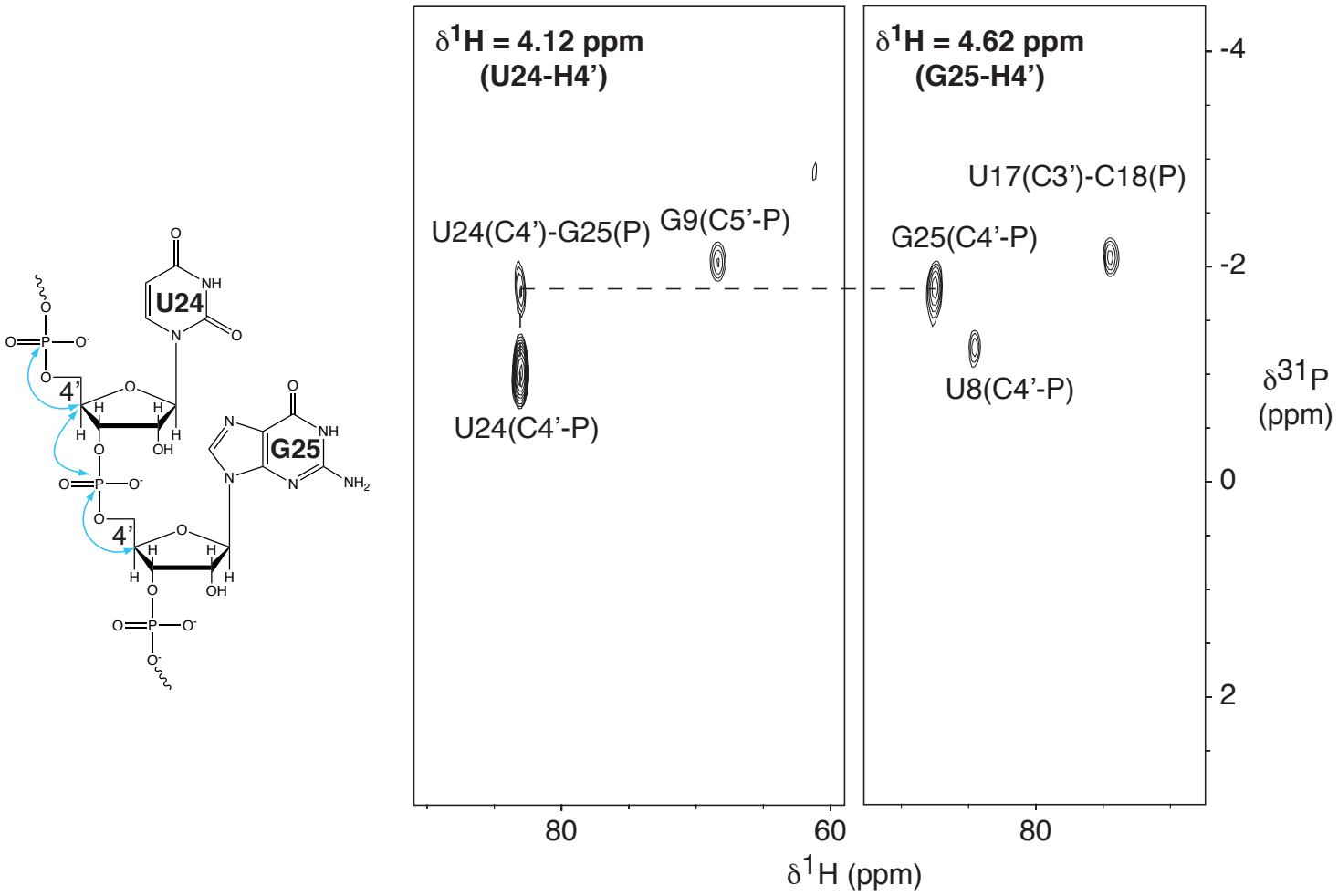

U24
G25
4'
4'
δ1H = 4.12 ppm
(U24-H4')
δ1H = 4.62 ppm
(G25-H4')
U24(C4')-G25(P)
G9(C5'-P)
U24(C4'-P)
U17(C3')-C18(P)
G25(C4'-P)
U8(C4'-P)
-4
-2
0
2
δ31P
(ppm)
80
60
80
δ1H (ppm)

**Figure 7**

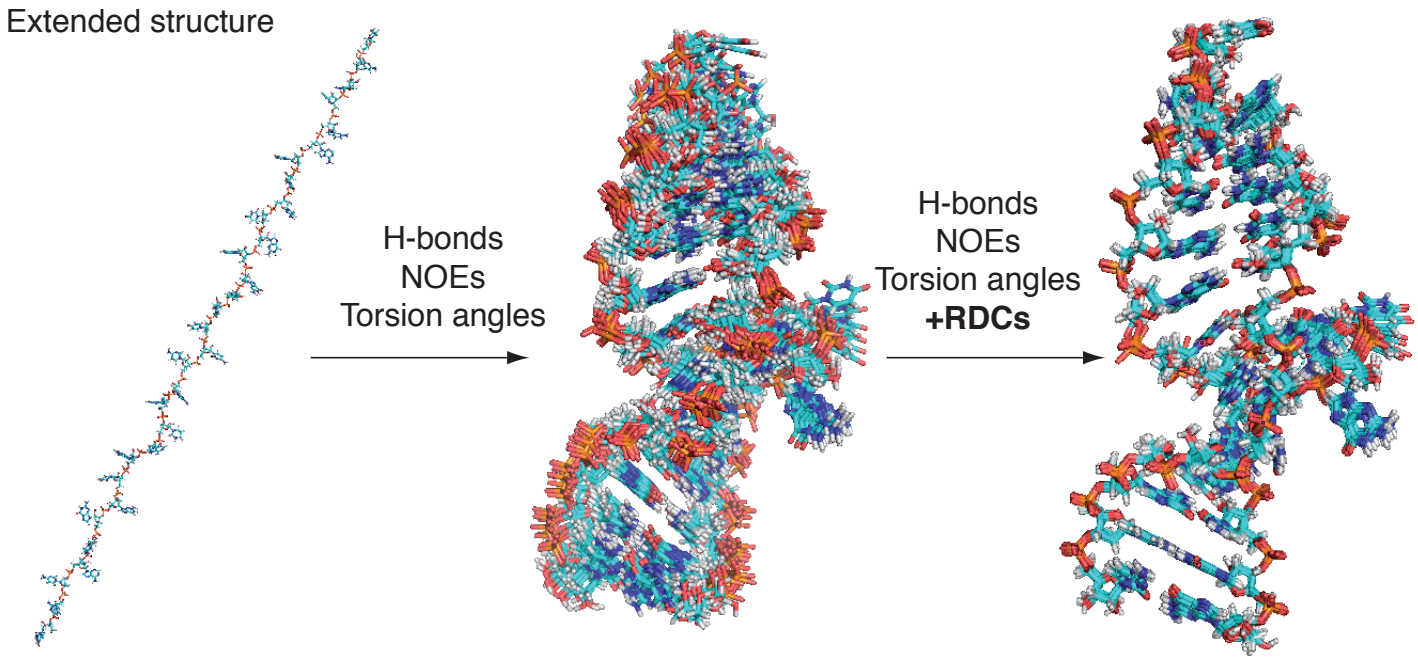

Extended structure
H-bonds
NOEs
Torsion angles
H-bonds
NOEs
Torsion angles
+RDCs

**Figure 8**