Deleterious Effects of Carbon-Carbon Dipolar Coupling On RNA NMR Dynamics

Hyeyeon Nam[1], Owen Becette[1], Regan M. LeBlanc[1,2], Daniel Oh[1], David A. Case[3], Theodore K. Dayie[1]

[1]Center for Biomolecular Structure and Organization, Department of Chemistry and Biochemistry, University of Maryland, College Park, MD, 20742 USA

[2]Current address: Basic Research Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD, 21702 USA

[3]Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ, 08854 USA

*Corresponding author Tel: 301-405-3165

E-mail: dayie@umd.edu

Abstract

Many regulatory RNAs undergo dynamic exchanges that are crucial for their biological functions and NMR spectroscopy is a versatile tool for monitoring dynamic motions of biomolecules. Meaningful information on biomolecular dynamics requires an accurate measurement of relaxation parameters such as longitudinal (R_1) rates, transverse (R_2) rates and heteronuclear Overhauser effect (hNOE). However, earlier studies have shown that the large $^{13}C - {}^{13}C$ interactions complicate analysis of the carbon relaxation parameters. To investigate the effect of ^{13}C-^{13}C interactions on RNA dynamic studies, we performed relaxation measurements on various RNA samples with different labeling patterns and compared these measurements with the computational simulations. For uniformly labeled samples, contributions of the neighboring carbon to R_1 measurements were observed. These contributions increased with increasing magnetic field and overall correlation time (τ_C) for R_1 rates, necessitating more careful analysis for uniformly labeled large RNAs. In addition, the hNOE measurements were also affected by the adjacent carbon nuclei. Unlike R_1 rates, $R_{1\rho}$ rates showed relatively good agreement between uniformly- and site-selectively labeled samples, suggesting no dramatic effect from their attached carbon, in agreement with previous observations. Overall, having more accurate rate measurements avoids complex analysis and will be a key for interpreting ^{13}C relaxation rates for molecular motion that can provide valuable insights into cellular molecular recognition events.

Keywords

RNA dynamics, Dipolar coupling, Longitudinal (R_1) relaxation, On-resonance $R_{1\rho}$, Heteronuclear Overhauser Effect (hNOE), Chemical enzymatic labeling

Introduction

The biological function of many regulatory RNAs depends on both the structure and dynamics of the molecule. Studies continue to show that dynamic exchanges of RNA molecules occur over a wide range of timescales that are important for their biological activity.[1-3] Riboswitches, for example, undergo dynamic conformational changes in response to ligand binding to regulate gene expression at the level of transcription, splicing, or translation.[4,5] Similarly, interhelical motions of tRNA can significantly change the relative orientation of helical domains and affect the formation of tRNA-protein, tRNA-ribosome complexes.[6] Furthermore, active sites of ribozymes are conformationally flexible and can undergo dynamic change during catalytic cycles to form reactive structures.[7,8]

To better understand these dynamic properties of RNA molecules, different biophysical tools such as cryo-EM[9-12], small molecule FRET[13], and time-resolved X-ray crystallography[14] have been applied. Among these tools, NMR spectroscopy enables the direct monitoring and quantification of dynamic motions at atomic resolution. The two commonly measured dynamic parameters comprise longitudinal (R_1) and transverse (R_2) relaxation rates.[15,16] The R_1 relaxation rate measures the return of the longitudinal magnetization to thermal equilibrium whereas the R_2 relaxation rate measures the decay of x- and y- magnetization (loss of coherence as individual spins experience slightly different resonance frequencies). The relaxation rates R_1 and R_2 are directly related to spectral density function, $J(\omega)$, which describes the probability of finding dynamic motions at a given frequency (ω). In addition to R_1 and R_2, hNOE measures the change in heteronuclear spin magnetization in response to saturating protons spins and can provide meaningful information on rotational diffusion, correlation time, internal motions, and flexibility.[17,18]

In proteins, dynamic motions are typically measured using ^{15}N nuclei in the amide backbone.[15,19] For RNAs, ^{15}N nuclei are only present in the bases and due to the water exchange properties of imino protons ^{15}N nuclei are suitable reporters of hydrogen-bonding and dynamics of base-paired guanine and uracil bases[20]; solvent exposed imino regions are usually broadened beyond detection. Nonetheless, ^{15}N chemical shift is very sensitive to RNA secondary structure and makes imino ^{15}N relaxation measurements a useful complement for characterizing the secondary structure of RNA. The limited availability of suitable imino nitrogen probes has necessitated the use of primarily protonated carbons as alternative relaxation probes. These carbon sites are found in all the ribose sugar moiety (C1', C2', C3', C4', and C5') and all 4 bases (C2-Ade, C8-Ade, C8-Gua, C5-Cyt, C6-Cyt, C5-Ura, C6-Ura). Thus, ^{13}C

relaxation measurements can potentially provide a more complete coverage of the structure necessary to obtain further insights into dynamic properties of these RNAs.

Despite the greater number of detectable ^{13}C nuclei in RNA compared to ^{15}N nuclei, there are significant challenges that complicate measurements and analysis of ^{13}C relaxation. First in uniformly labeled RNA, ^{13}C spins do not approximate an isolated 2-spin system. These carbon sites are not only linked by intricate multi-bond couplings to other ^{15}N, ^{13}C, and ^{1}H nuclei but also positioned within 2.5 Å of these nuclei. These extensive dipolar couplings between adjacent carbons complicate analysis of the carbon relaxation rates.[21-23] The relative contribution to the relaxation of a ^{13}C nucleus from covalently attached carbon and proton spins can be determined according to the following approximation:[21]

$$\frac{\rho_{CC}}{\rho_{HC}} \approx \frac{(\omega_C \tau_C)^2}{3}\left(\frac{{\gamma_C}^2}{{\gamma_H}^2}\right)\left(\frac{r_{HC}}{r_{CC}}\right)^6 \quad (1)$$

where ρ_{ic} is the contribution to the selective relaxation rate of the ^{13}C nucleus from an attached proton or carbon, γ_j is the gyromagnetic ratio of spin j, r_{ij} is the distance between spins i and j, ω_C is the Larmor frequency of carbon, and τ_C is the overall molecular tumbling time. As shown in eq 1, the contribution of the attached carbon relative to the attached proton on the ^{13}C relaxation rate is proportional to the square of the overall correlation time.

Previous works have shown that uniform isotopic labeling negatively impacts longitudinal relaxation rate measurements in proteins and nucleic acids; discrepant R_1 measurements increase with increased τ_C as captured by eq 1.[21,23-29] Comparison of uniformly ^{13}C labeled and alternate site ^{13}C labeled adenine ribonucleotide monophosphate (rAMP) showed greater difference in relaxation rate with larger τ_C, again in agreement with eq 1.[23] These network of carbon-carbon couplings hinder accurate measurements and complicate extraction of relaxation rate constants. These prior works suggest that the interpretation of relaxation measurements to determine dynamic models of macromolecules can be erroneously impacted especially for biopolymers of large size (τ_C > 7 ns). Therefore, the contribution of ^{13}C – ^{13}C interaction should be taken into account during data analysis when using uniformly labeled sample for accurate analysis of RNA dynamics.[25,26] Nonetheless a prevailing view in the field is that ^{13}C-^{13}C dipolar coupling are only an issue for atoms such as C5 and C6 in pyrimidines or atoms in the sugar ring (C1'-C5') but not for spin pairs such C8 sites in purine rings. The latter are considered fairly isolated from neighboring carbons and thus

immune to these dipolar coupling effects. We show with simulations and experimental data that C8 also suffers from the same errors, and these get progressively worse with increasing molecular weight and field strength.

Here we, therefore, extend those earlier studies and address the significant effect of adjacent ^{13}C on RNA dynamics measurements for not just longitudinal but also transverse and cross-relaxation measurements in RNA. We transcribe various RNAs *in vitro* using uniform or site selective labeled ribonucleoside-5'-triphosphates (rNTPs). The site-selective labeled rNTPs, synthesized in house using chemo-enzymatic methods (Fig. 1a), are strategically enriched with ^{13}C-isotopes at the ribose C1', purine C8, and pyrimidine C6 carbon sites (Fig. 1b).[30,31] TROSY-detected pulse sequences[26] are used to measure carbon R_1, on-resonance $R_{1\rho}$ rates and hNOE for RNAs with sizes varying between 14 to 61 nt. We also compare our experimental results with the computational simulation of relaxation rates. Our simulation includes the effect of dipolar coupling of nearby nuclei calculated using X-ray crystal coordinates derived distances and CSA effect of each nuclei using values derived from density functional theory (DFT) calculation. Such an approach allows us to compare theoretical predictions with experimentally measured values.

Materials and Methods

Chemicals and oligonucleotides. Uniformly labeled rNTPs were purchased from Cambridge Isotope Laboratories, Inc (MA), and site-selectively labeled rNTPs were synthesized in-house using previously reported chemo-enzymatic methods[30,31] An outline of the synthesis is described in Fig. 1. Synthetic DNA templates and the RNA polymerase promoter sequences (5′-CTA ATA CGA CTC ACT ATA G-3′) were purchased from Integrated DNA Technologies, Inc (CA). Two nucleotides at the 5' end of synthetic DNA templates were modified with 2'-O-methyl group to reduce transcription heterigenety.[32] Modified T7 RNA polymerases[33] used for *in vitro* transcription were expressed in *Escherichia coli* BL21(DE3) and purified with a Ni-chelating Sepharose column (GE Healthcare).

RNA sample preparation. RNA samples were synthesized by *in vitro* transcription using T7 RNA polymerases. Sequences are provided in the supplementary materials. The transcription conditions used buffer containing 40 mM Tris-HCl (pH 8.0), 1 mM spermidine, 10 mM dithiothreitol (DTT), 0.01 % Triton X-100, 300 nM DNA template, 300 nM promoter sequence, 2 units/ml of thermostable inorganic pyrophosphatase (New England

BioLabs, MA) and 0.1 mg/ml T7 polymerases for 10 ml reaction volume. Concentration of rNTPs and Mg^{2+}, ranging between 5 – 20 mM, were optimized for each RNAs to obtain maximum yields. After 3-4 hr of incubation at 37 °C, RNAs were purified by phenol:choloform extraction, ethanol precipitation, and denaturing polyacrylamide gel electrophoresis followed by electroelution. Eluted samples were solvent exchanged and concentrated into an appropriate buffer (buffer conditions provided in supplementary materials) using 3K MWCO tube (Fisher Scientific, NH). The samples were then folded by heating up to 95 °C for 3 min and snap-cooled on ice for 15 min. Finally, the RNA samples were lyophilized and re-suspended in 300 µl of 99.8 % D_2O resulting in a final concentration of 0.5 – 1 mM.

NMR measurements. NMR experiments were performed at 298.15 K on a 800 MHz Avance III Bruker spectrometer equipped with a triple resonance cryogenic probe and collected data were analyzed using Bruker TopSpin 3.2. TROSY-detected pulse sequences for the relaxation measurements were adapted from earlier works.[21,26] For R_1 and $R_{1\rho}$ experiments for uniform labeled samples, selective pulses were applied for decoupling as described in the previous work.[26] Shape pulses used for the on-resonance inversion, on-resonance refocusing, and off-resonance inversion were Q3, RSNOB, and Iburp2, respectively. The pulse length for each pulses were 937.5 µs, 1000 µs, and 450 µs, respectively. The offset for Iburp2 was -8000 ppm. For site-selective labeled samples, hard 180° pulse was applied instead during the second INEPT period. Applying hard pulse instead of the selective pulse did not affect the relaxation rates (Data not shown). For both uniform and site-selective samples, carrier was set to 89.2 ppm for C1' with a sweep width of 6.4 ppm. For C6/C8, the carrier was set to 137.8 ppm with a sweep width of 10.4 ppm. For all experiments, a recycle delay of 2.5 s was used. Two-dimensional experiments were performed with 10 delay points including one repeated point for error estimation by Monte-Carlo simulation. RELAXFIT MATLAB program[34] was used to fit the data and calculate the rates and errors. For steady-state hNOE measurements, 8 s period of proton saturation was applied. Two-dimensional TROSY spectra were collected with and without the proton saturation. Each experiments were performed twice for error calculation. Peak positions and intensities were analyzed using TopSpin 3.2 (Bruker) and NOE was calculated ($\eta = (I_{sat} - I_{ref})/I_{ref}$) using the intensities determined from each spectrum. Uncertainties in NOE values were calculated by propagating the error in the peak intensities.

Results and Discussion

^{13}C Longitudinal Relaxation (R_1) Measurement. Longitudinal relaxation (R_1) is one of the most commonly measured relaxation parameters to obtain dynamics information on biomolecules.[15,19,26] The longitudinal relaxation rate of the ^{13}C spin is given by the equation[21,35,36]:

$$\frac{dC_z{}^K}{dt} = -R_{1,C_K}\left(C_z{}^K - C_z{}^{K,0}\right) - \sum_i \sigma_{C_K,C_i}\left(C_z{}^i - C_z{}^{i,0}\right) - \sum_j \sigma_{C_K,H_j}\left(I_z{}^j - I_z{}^{j,0}\right) \quad (2)$$

where $C_z{}^K$ is the longitudinal magnetization of carbon K, $C_z{}^{K,0}$ is the equilibrium longitudinal magnetization, $C_z{}^i$ and $I_z{}^j$ are longitudinal magnetizations of directly attached carbon i and proton j, respectively. R_{1,C_K} is auto-relaxation rates and σ is cross-relaxation rates given by following equations[21,35]:

$$R_{1,C_K} = \sum_i R_{1,C_K,H_i} + \sum_j R_{1,C_K,C_j} + \sum_j R_{1,C_K,N_j} + R_{1,CSA} \quad (3)$$

$$\sigma_{C_K,i} = \frac{1}{10} d_{C_K,i}{}^2 \left[6J\left(\omega_{C_K} + \omega_i\right) - J\left(\omega_{C_K} - \omega_i\right)\right] \quad (4)$$

in which $R_{1,C_K,i}$, $R_{1,CSA}$, and $d_{C_K,i}$ are[21,35]:

$$R_{1,C_K,i} = \frac{1}{10} d_{C_K,i}{}^2 \left[3J\left(\omega_{C_K}\right) + 6J\left(\omega_{C_K} + \omega_i\right) + J\left(\omega_{C_K} - \omega_i\right)\right] \quad (5)$$

$$R_{1,CSA} = \frac{1}{10} C_{C_K}{}^2 \left[3J\left(\omega_{C_K}\right)\right] \quad (6)$$

$$d_{C_K,i} = \frac{\mu_0 \gamma_{C_K} \gamma_i h}{8\pi^2 r^3} \quad (7)$$

Eq 2 shows that the longitudinal relaxation rates have contribution from both auto-relaxation rate (R_{1,C_K}) and cross-relaxation rate ($\sigma_{C_K,i}$) between neighboring spins. $\sigma_{C,H}$ can be neglected since it is generally 2 – 3 orders of magnitude smaller than auto-relaxation rate and C – C cross-relaxation is experimentally suppressed as previously described.[21,26] Therefore, the major contribution to the longitudinal relaxation rates comes from auto-relaxation rates. As shown in eq 3, auto-relaxation rate (R_{1,C_K}) incorporates R_{1,C_K,H_i}, R_{1,C_K,C_j}, and R_{1,C_K,N_j} terms: these represent dipolar interaction of carbon K with its directly attached proton, other ^{13}C nuclei, and ^{15}N nuclei respectively. In addition R_{1,C_K} includes $R_{1,CSA}$, representing CSA relaxation mechanism of carbon K. C_{C_K} is a CSA constant for carbon K, defined as $\omega_{C_K}\Delta\sigma_{C_K}/\sqrt{3}$, where $\Delta\sigma_{C_K} = \sqrt{\sigma_x{}^2 + \sigma_x{}^2 - \sigma_x\sigma_y}$, $\sigma_x = \sigma_{33} - \sigma_{11}$, and $\sigma_y = \sigma_{33} - \sigma_{22}$. σ_{11}, σ_{22}, and σ_{33} are the

principal components of the chemical shielding tensor.[37-39] $J(\omega)$ is a spectral density function which is assumed to be a Lorentzian $J(\omega) = \frac{\tau_C}{1+(\omega\tau_C)^2}$, γ_i is the gyromagnetic ratio of spin i, r is the distance between the two spins and h is Plank's constant.

We used these equations to analyze and compare the relaxation rates of uniform and site selective labeled NTPs. Unlike the auto-relaxation rate of C6 on selectively labeled pyrimidines ($R_{1,C_6}{}^{s}$), that of C6 on uniformly labeled pyrimidines ($R_{1,C_6}{}^{u}$) includes dipolar contributions from ^{13}C labeled C2, C4 and C5 (R_{1,C_6,C_i}), as well as the dipolar contributions from ^{15}N labeled N1 and N3 (R_{1,C_6,N_i}). Of great interest the $J(\omega_{C_K} - \omega_i)$ term in eq 5 is proportional to $J(0)$ when spin i is carbon, which increases with increasing τ_C (Supplementary Fig. S1). Thus, increased dipolar contributions from the neighboring carbons are expected with greater τ_C. Among all the R_{1,C_6,C_i} terms, R_{1,C_6,C_5} is a major contributor to the difference between $R_{1,C_6}{}^{s}$ and $R_{1,C_6}{}^{u}$, as C5 is spatially the closest carbon to C6 and, thus, d_{C_6,C_5} is greater than other d_{C_6,C_i}. This predicts C5 having the largest dipolar contributions to the $R_{1,C_6}{}^{u}$. Similarly, $R_{1,C_1'}{}^{u}$ includes $R_{1,C_1',C_2'}$ term in contrast to $R_{1,C_1'}{}^{s}$, which leads to a greater deviation between $R_{1,C_1'}{}^{u}$ and $R_{1,C_1'}{}^{s}$ as τ_C increases. On the other hand, because C8 in purines do not have directly bonded carbon neighbors, dipolar contributions for C8 are usually assumed to be relatively small compared to CH-groups in pyrimidines or ribose moieties. The simulation of dipolar contributions from C6, N7, and N9 show the contributions being less than ~ 2%, indicating negligible effects on the R_1 rates even for the large RNAs (Supplementary Fig. S2) However, the contributions from neighboring 2-bond carbons, such as C4 and C5 carbons, increase up to ~ 20 -30 % for large RNAs (Supplementary Fig. S2).

Simulation of auto-relaxation rates for uniform and site selective labeled C6 in UTP (Fig. 2a) and C8 in ATP (Fig. 2b) predict discrepant rates between them. For these simulations we assume an isotropic tumbling as it is the simplest and clearest way to monitor the effect of the neighboring carbons. However, further studies are needed to systematically analyze the effect of anisotropic tumbling. The internuclear distances used for the simulation of dipolar effect are shown in Supplementary Table S1. CSA values, derived from DFT calculation, were taken into account for the simulation as well (details are described in supplementary materials). For UTP, no notable difference exists between $R_{1,C_6}{}^{u}$ and $R_{1,C_6}{}^{s}$ when τ_C is less than ~ 2 ns, (Fig. 2a). However, this difference becomes significant with increasing τ_C due to increased contribution from R_{1,C_6,C_5} (Fig. 2a). Similarly, the calculated difference between $R_{1,C_8}{}^{u}$

and $R_{1,C_8}{}^{s}$ values are negligibly small for τ_C < 5 ns with the differences below ~ 10 % (Fig. 2b, solid lines). This is expected since C8 has no directly bonded carbon-carbon coupling partners. When τ_C > 20 ns, the simulation predicts even greater differences for both UTP and ATP, reaching ~ 90% and ~ 60% at 40 ns for C6 and C8, respectively (Fig. 2a and b, dotted lines). It is noteworthy that purine C8, with no directly attached carbon, also have increased dipolar contributions from neighboring ^{13}C labeled carbons. This indicates the R_{1,C_8,C_i} term, though negligible for τ_C < 5ns, leads to larger difference between $R_{1,C_8}{}^{u}$ and $R_{1,C_8}{}^{s}$ for τ_C > 20ns. In addition, the simulations of R_1 at different magnetic fields indicates greater % differences at higher fields for both UTP (Fig. 2c) and ATP (Fig. 2d). This implies that while NMR experiments at higher frequencies may provide results with better resolution and sensitivity, these higher fields may paradoxically hinder accurate measurements of R_1 for uniform labeled RNA samples.

To see how the contribution of neighboring nuclei on R_1 rates bear out experimentally, uniformly labeled RNAs and site-selectively labeled (Fig. 1b) RNA samples of three different sizes (14, 29 and 61 nt) were prepared to examine how the contribution on R_1 rates varies with the size of the macromolecules. The measurements were performed at 25 °C and 800 MHz for all RNAs. In agreement with the simulations, there were no notable differences in R_1 of uniformly labeled ($R1_u$) and site-selectively labeled ($R1_s$) samples for the 14 nt RNA, indicating insignificant carbon-carbon coupling effects for both base and ribose carbon atoms (Fig. 3a). The average difference in R_1 rates were < 5 % (Table 1) suggesting little to no dipolar coupling effect on R_1 relaxation for small RNAs. However, R_{1s} was slightly greater than R_{1u} for some residues (Fig. 3a). The source of this curious observation, also reported in an earlier work[23], is currently unclear and needs further investigation in future studies. R_{1u} and R_{1s} for C8 position on purine for 29 nt RNA were also comparable resulting in an average difference of ~2 % (Fig. 3b, Table 1). The C6 position of pyrimidine and C1' position of ribose, however, have directly bonded carbon coupling partners and this effect can be observed by the significant differences on R_{1u} and R_{1s} (Fig. 3b). The average difference was ~13 % for C6 and ~ 16 % for C1' showing increased contribution of the attached carbon compared to the 14 nt RNA (Table 1). For 61 nt RNA, the % differences were even greater (Fig. 3c, Table 1) indicating increased dipolar coupling effects. These results show increasing contribution of the adjacent carbon with increasing size of RNA supporting the prediction from eq 1 and simulations (Fig. 2). The τ_C values range between 0 - 20 ns for the RNA sizes (14, 29, 61 nt) used for the measurements. In this τ_C range, the simulated differences between $R_{1,C_6}{}^{u}$ and $R_{1,C_6}{}^{s}$ were within 0 – 25 % (Fig. 2a, solid lines). This % difference is in agreement with the experimentally observed % differences (Table 1). For

the same τ_C range, calculated differences between ${R_{1,C_8}}^u$ and ${R_{1,C_8}}^s$ are between 0 – 15 % (Fig. 2b, solid lines). Again, these calculated values are in agreement with the observed results in Figure 3 and Table 1.

These results are also consistent with the observation that for small nucleic acids (τ_C < 2 ns) the contribution of carbon-carbon dipolar coupling interaction to the relaxation is negligible compared to the proton-carbon interaction (ρ_{CC}/ρ_{CH} < 0. 06) at 800 MHz proton frequency. However, for larger nucleic acids the ^{13}C-^{13}C contribution is no longer negligible (ρ_{CC}/ρ_{CH} = 0.83 for τ_C = 10 ns) (Supplementary Fig. S3). Overall, the observed results and simulations suggest necessity of more careful analysis of R_1 rates for the uniform labeled RNAs, especially when the size is large and experiments are conducted at very high fields.

Steady-state ^1H-^{13}C hNOE Measurement. Accurate measurements of heteronuclear Overhauser effect (hNOE) parameters can also provide meaningful dynamics information on biomolecules when combined with R_1 and R_2.[17] A value of NOE enhancement, η (= NOE-1), is defined as the difference between the peak intensity in the irradiated spectrum (I_{sat}) and the reference non-irradiated spectrum (I_{ref}) scaled by the signal intensity from the reference experiment ($\eta = (I_{sat} - I_{ref})/I_{ref}$). The steady-state NOE enhancement (η_{SS}) of the spin K is given by[35,36]:

$$\eta_{SS} = \frac{\gamma_i \sigma_{i,K}}{\gamma_X R_{1,K}} \tag{8}$$

where $\sigma_{i,K}$ is cross-relaxation rate between spin i and K, $R_{1,K}$ is auto-relaxation rate of spin K and $\gamma_{i(H)}$ is gyromagnetic ratio of spin i (K). For this experiment, proton spins were saturated using hard pulses and the NOE enhancement of C1' on ribose is monitored. As shown in eq 8, auto-relaxation rate ($R_{1,C_1'}$) of C1' and cross relaxation ($\sigma_{C_1',H_1'}$) between C1' and H1' determines the NOE enhancement (η_{SS}) of the C1'. Table 2 shows the results from the steady-state ^1H-^{13}C hNOE measurements on ribose C1' of a 29 nt RNA. As shown in the Table 2, uniformly labeled RNA (NOE_u) had higher NOE values compared to site-selectively labeled RNA (NOE_s). These results are in agreement with the earlier protein work of Kay and co-workers[21] which showed increased NOE for uniformly labeled alanine compared to singly labeled alanine when τ_C is greater than ~ 6 ns. Similarly, Engelke and Rüterjans showed the contribution of carbon-carbon interaction in protein hNOE measurement and speculated greater effect with larger

rotational correlation time.[40] Our simulations also suggest that discrepancy between hNOE for uniform and selective labels will scale as a function of overall correlation time and magnetic field strength (Supplementary Fig. S4)

Unlike site-selectively labeled ribose, in which only the C1' is isotopically enriched, uniformly labeled ribose has both C1' and C2' labeled (Fig. 1). Since both C1' and C2' has a proton attached, saturation of proton spins enhances the longitudinal ^{13}C-magnetization of both carbon positions. Enhanced intensity of the C2' can then be transferred to the C1' spin and further increase the intensity of C1' as a result of cross relaxation. This effect becomes more significant with greater τ_C as cross-relaxation rate ($\sigma_{C_1',C_2'}$) for C1' – C2' increases with increasing correlation time. In contrast, C2' in site-selectively labeled ribose is unlabeled resulting in smaller NOE_s compared to NOE_u. These NOE results suggest that the analysis of steady-state NOE values for the uniformly labeled RNA should be done with more care due to the cross-relaxation effect from neighboring carbon spins for C1', pyrimidine C5 and C6 carbons. Until now such measurements were infrequent for nucleic acids and preferentially omitted for larger RNAs[23,25-27]. Our selective labels[31,41] make these readily accessible and will be useful for mapping the spectral density functions for RNA systems.[18]

Additionally, we applied selective proton saturation pulse for hNOE measurements to test whether such pulses can reduce the contribution from carbon-carbon interaction for uniformly labeled sample. For site-selectively labeled sample, NOE values with the selective proton saturation pulse and the 180° hard pulse were comparable indicating negligible effect of selective pulse when there is no adjacent carbon (Supplementary Table. S2). However, when the same selective pulse was applied to the uniformly labeled sample, the NOE values were reduced to values close to that of NOE_s (Supplementary Table. S2). This result suggests that the selective proton saturation pulse can be applied as an alternative method to minimize the effect of carbon-carbon coupling on hNOE measurement when uniformly labeled sample is used.

^{13}C $R_{1\rho}$ Relaxation Measurement. The transverse relaxation rate constant (R_2) is a measure of the rate at which transverse magnetization decays to zero. Unlike R_1, R_2 provides information on the chemical exchange process in biomolecules. Typically, Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence is used to measure R_2 for proteins.[22] However, this is problematic for nucleic acids due to echo-modulation caused by the large J_{CC} (38-68 Hz) couplings during the ^{13}C relaxation delay.[21,42] Therefore, CPMG experiment has been applied for the R_2 measurements of site-

selectively labeled RNA samples.[43,44] To avoid this problem, $R_{1\rho}$ is used as an alternative method of measuring R_2 of biomolecules.[22] In an $R_{1\rho}$ experiment, a radio frequency field pulse spin-locks the magnetization in the rotating frame and relaxation rate constant along the effective field is measured. The observed $R_{1\rho}$ rates contain contribution from both longitudinal and transverse relaxation. But the information on R_2 is easily extracted according to $R_{1\rho} = R_1 cos^2\theta + R_2 sin^2\theta$ in which $\theta = tan^{-1}\left(\frac{\omega_1}{\Omega}\right)$, ω_1 is the strength of the B_1 spin-lock field (Hz) and Ω is the resonance offset (Hz) from the spinlock carrier frequency. In contrast to R_1 simulation, which predicted significant difference between R_{1u} and R_{1s} (Fig. 2), simulation of R_2 showed negligible difference between uniform (R_{2u}) and site-selective (R_{2s}) labeled samples (Supplementary Fig. S6). The maximum difference predicted was less than 5 % for both purine and pyrimidine at all magnetic field strengths. Therefore, we expect minimum differences in experimentally determined R_{2u} and R_{2s} as well.

For $R_{1\rho}$ measurements, the potential errors caused by carbon – carbon interaction may come from the Hartmann – Hahn transfer during the spin-lock period for spins that have comparable chemical shifts. Work by Kay and co-workers indicated reasonable agreement of $R_{1\rho}$ rates for singly- and uniformly-labeled alanine indicating Hartmann – Hahn effects are negligible, except for Ser and Thr.[21] In addition, Hoogstraten and co-workers compared $R_{1\rho}$ rates of uniform and alternate-site labeled rAMP and observed agreement between the measurements under some combinations of resonance, spin-lock power and τ_C.[23]

In order to reduce the Hartmann – Hahn transfer between scalar coupled spins, the $R_{1\rho}$ pulse sequences[26] used for the experiment include selective spin-locks and therefore, minimum discrepancy was expected between $R_{1\rho}$ rate of uniformly labeled sample and site-selectively labeled sample. To investigate the effect of neighboring carbon, $R_{1\rho}$ experiments were performed at 800 MHz with 1.9 kHz B_1 field on-resonance for a 29 nt RNA and the R_2 rates were calculated as described earlier. The observed results showed good agreement between R_{2u} and R_{2s} for C8 positions giving average difference of ~ 4% (Fig. 4). For C6 on pyrimidines and C1' on ribose, the average differences were ~7 % which is slightly higher than that of C8 (Fig. 4). Compared to the differences observed between R_{1u} and R_{1s} (Fig. 2, Table 1), R_2 measurements showed relatively good agreement between uniformly- and site-selectively labeled samples (Fig. 4). Comparable results were also observed for the R_2 rates of 61 nt RNA indicating that the increasing τ_C has no significant effect on the R_2 measurement (Supplementary Fig. S5). These results are compatible with the R_2 simulation showing small difference between R_{2u} and R_{2s} (Supplementary Fig. S6). The simulation also shows

increasing τ_C has no notable effect on the % differences between R_{2u} and R_{2s}. In addition, the simulation predicts smaller % differences with higher frequencies, suggesting more accurate R_2 measurement at higher magnetic field.

The observed results also suggest minimum Hartmann–Hahn transfer effect in $R_{1\rho}$ measurement for uniformly labeled samples. The maximum amount of net Hartmann–Hahn transfer between two spins 1 and 2 can be given by following equation[45]:

$$R_{1,2}^{max} = \frac{[{J_{1,2}}^2\{1+\cos(\theta_1-\theta_2)\}^2]}{\left[4\left(\nu_{1,eff}-\nu_{2,eff}\right)^2+{J_{1,2}}^2\{1+\cos(\theta_1-\theta_2)\}^2\right]} \quad (9)$$

where $J_{1,2}$ is the coupling constants between two spins, θ_i is the spin-lock angle of spin i defined as described above and $\nu_{i,eff}$ is the effective field-strength on spin i given by $\sqrt{\omega^2 + {\Omega_i}^2}$. Eq 9 predicts low transfer efficiency when the two spins have large difference in effective field strengths and relatively small coupling constants. Calculating maximum Hartmann–Hahn transfer for C1' and C2' using 43 Hz for a coupling constant, 1.9 kHz for a spin-lock field strength and 1.7 kHz for a difference in effective field results in maximum transfer efficiency < 1 ‰. Similar outcomes are observed when calculated for C6 and C5 supporting the experimental results showing minimum effect from adjacent carbon for uniformly labeled samples. Together with earlier works[21,23,26], these results suggest that the effect from the neighboring carbon can be safely ignored for the $R_{1\rho}$ measurements.

Conclusion

Although ^{13}C relaxation measurements provide a more complete coverage of RNA structural elements than ^{15}N measurements, a number of significant challenges render measurement and analysis of ^{13}C relaxation problematic. Foremost being the dipolar coupling network of interactions present in uniformly labeled RNA samples. We have therefore carried out dynamic studies, via simulations and experimental NMR measurements on RNAs of varying sizes with different labeling patterns, to investigate the effect of adjacent carbon on the relaxation measurements. Not unexpectedly, simulations and measurements of R_1 of pyrimidine C6 and ribose C1' positions indicate large discrepancies exists between rates extracted using uniformly- and site-selectively labeled samples. Predictions, confirmed experimentally, suggest the carbon-carbon coupling effects from directly attached carbon hinder accurate dynamics measurements in uniformly labeled RNA. Furthermore, the contribution of the neighboring carbon on R_1

rates increase with increasing τ_C, necessitating more careful analysis of relaxation data with larger size of RNAs when the sample is uniformly labeled. Unexpectedly, simulations and measurements of R_1 on purine C8 position show larger discrepancy between R_{1u} and R_{1s} with increasing τ_C, indicating greater carbon-carbon coupling effects in large RNAs. This contrasts with prevailing assumptions that purine C8 forms an isolated spin system even within the network of carbon couplings found in uniformly labeled nucleobases. Moreover, the simulations suggest greater contribution on R_1 rates at higher magnetic field strengths. On the other hand, R_{2u} and R_{2s} of C6 and C1' extracted from $R_{1\rho}$ measurements showed relatively good agreement suggesting no dramatic effect from their attached carbon on $R_{1\rho}$ measurement.

Figure Caption

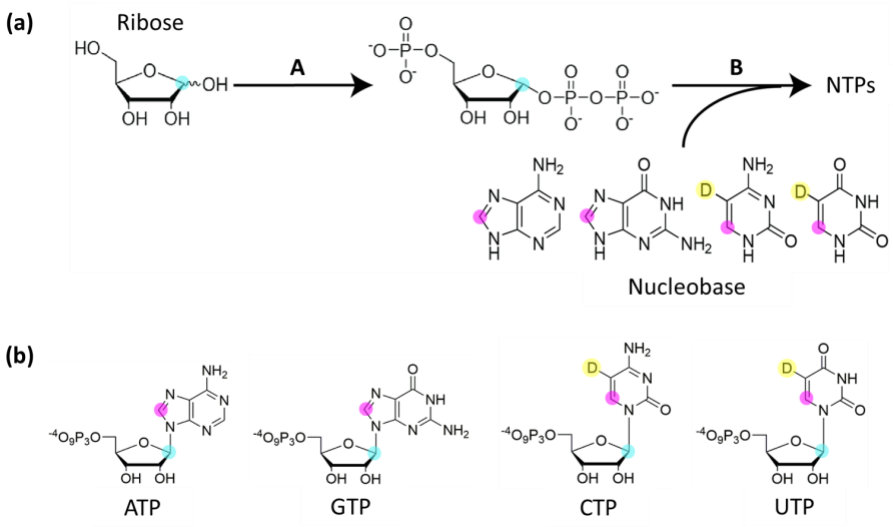

Fig. 1 a Outline of NTP synthesis by chemo-enzymatic methods.[30,31] Enzymes in step A includes RK and PRPPS. Enzymes in Step B includes MK and CK for ATP synthesis, GK and CK for GTP synthesis, UMPK and CK for UTP synthesis. UTP is converted to CTP using CTPS. For all steps MK and CK is added for dATP regeneration. **b** Scheme of four nucleotides used for the study. C1' on ribose (cyan) was labeled for all nucleotides. C8 position on base was labeled (magenta) for ATP and GTP. For pyrimidines, C6 position on base was labeled (magenta) and H5 position was deuterated (yellow) Enzyme abbreviation RK: ribokinase (E.C. 2.7.1.15), PRPPS: ribose-phosphate diphosphokinase (E.C. 2.7.6.1), MK: myokinase (E.C. 2.7.4.3), CK: creatine kinase (E.C. 2.7.3.2), GK: guanylate kinase (E.C. 2.7.4.8), UMPK: UMP kinase (E.C. 2.7.4.22), CTPS: CTP synthetase (E.C. 6.3.4.2)

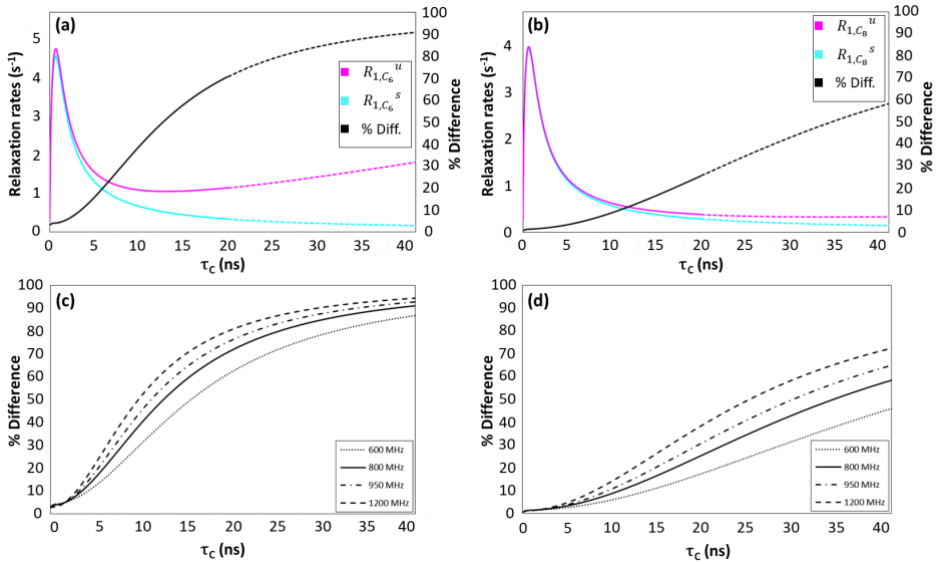

Fig. 2 a Simulation of R_{1,C_6} for U-labeled RNAs at 800 MHz. ${R_{1,C_6}}^u$, ${R_{1,C_6}}^s$ and % differences correspond to magenta, cyan and black lines, respectively. When τ_C is less than ~ 2 ns, there is no notable difference between ${R_{1,C_6}}^u$ and ${R_{1,C_6}}^s$. However, this difference becomes significant with increasing τ_C due to increasing contribution from R_{1,C_6,C_5}. **b** Simulation of R_{1,C_8} for A-labeled RNAs at 800 MHz. ${R_{1,C_8}}^u$, ${R_{1,C_8}}^s$ and % differences correspond to magenta, cyan and black lines, respectively. ${R_{1,C_8}}^u$ and ${R_{1,C_8}}^s$ values are comparable for τ_C < 5 ns with the differences below ~ 10 % (solid lines). When τ_C > 20 ns, the simulation predicts greater differences reaching ~ 60% at 40 ns (dotted lines). **c** Simulation of % differences between ${R_{1,C_6}}^u$ and ${R_{1,C_6}}^s$ for U-labeled RNAs at different frequencies. **d** Simulation of % differences between ${R_{1,C_8}}^u$ and ${R_{1,C_8}}^s$ for A-labeled RNAs at different frequencies. For both R_{1,C_6} and R_{1,C_8}, greater % differences are predicted with higher frequencies

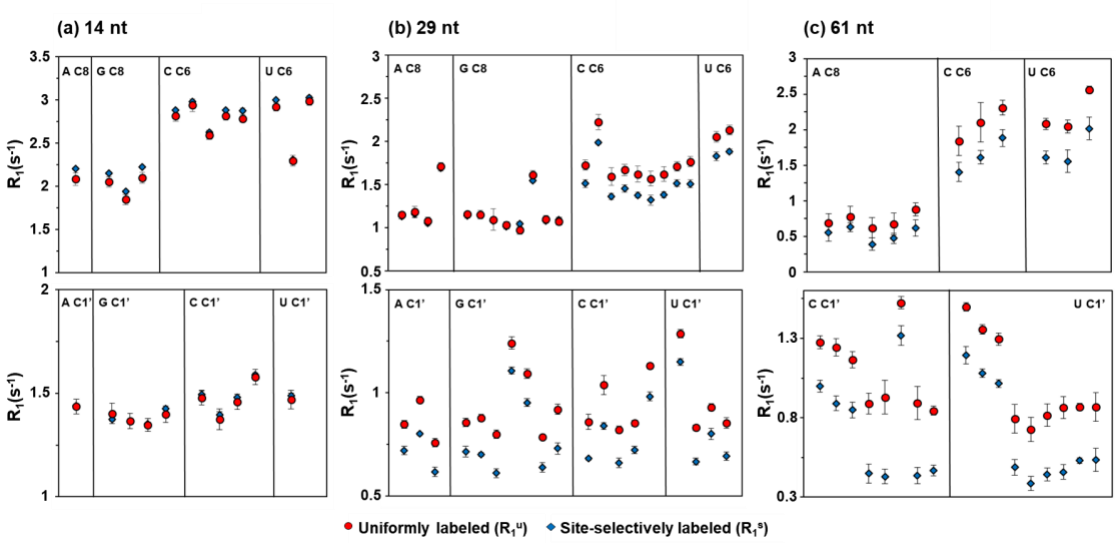

Fig. 3 R_1 rates of uniformly labeled sample (R_{1u}) and site-selectively labeled sample (R_{1s}) at 800 MHz. **a** For 14 nt RNA, there were no notable differences between R_{1u} and R_{1s} indicating no significant carbon-carbon coupling effects for both base and ribose. **b** R_{1u} and R_{1s} for C8 position on purine for 29 nt RNA were comparable which was expected since C8 is fairly isolated with no adjacent carbon-carbon coupling partners. The C6 position of pyrimidine and C1' position of ribose, however, have adjacent carbon-carbon coupling partners and this effect is reflected in the significant differences on R_{1u} and R_{1s}. **c** For 61 nt RNA, the discrepancies between R_{1u} and R_{1s} were even greater for C8, C6 and C1' indicating increased coupling effects

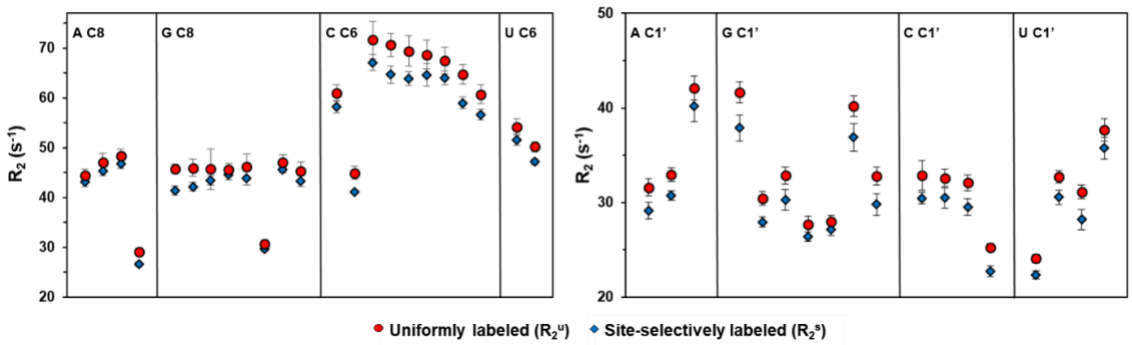

Fig. 4 R_2 rates of uniformly labeled sample (R_{2u}) and site-selectively labeled sample (R_{2s}) at 800 MHz. $R_{1\rho}$ experiments were performed at 1.9 kHz B_1 field on-resonance for a 29 nt RNA and the R_2 rates were extracted according to $R_{1\rho} = R_1 cos^2\theta + R_2 sin^2\theta$. The observed results showed a good agreement between R_{2u} and R_{2s} for C8 positions giving average difference of ~ 4%. For C6 on pyrimidines and C1' on ribose, the average differences were ~7 % which is slightly higher than that of C8. Compared to the discrepancies observed between R_{1u} and R_{1s}, R_2 measurements showed relatively good agreement between uniformly- and site-selectively labeled samples

RNA Size (nt)	^{13}C Position	% Difference in R_1		
		Maximum	Minimum	Average
14	C8	7.5	4.9	6.3
	C6	3.6	0.3	2.0
	C1'	3.9	0.1	1.5
29	C8	6.8	0.1	2.2
	C6	15.8	10.7	13.1
	C1'	23.6	5.4	15.7
61	C8	29.4	18.4	24.3
	C6	23.9	18.4	22.3
	C1'	54.0	13.6	37.8

Table 1 % Differences between R_{1u} and R_{1s} for varying size of RNAs at 800 MHz.

% differences ($= \frac{|{R_1}^u - {R_1}^s|}{{R_1}^u} \times 100$ %) were calculated for multiple peaks and averaged. Maximum and minimum values shows the range of observed values.

Peak	NOE^u	NOE^s
1	1.40 ± 0.02	1.33 ± 0.01
2	1.32 ± 0.01	1.25 ± 0.01
3	1.27 ± 0.02	1.19 ± 0.01
4	1.26 ± 0.02	1.17 ± 0.03
5	1.28 ± 0.01	1.20 ± 0.01
6	1.38 ± 0.02	1.29 ± 0.03

Table 2 Steady-state ^{1}H-^{13}C hNOE measurements on ribose C1' of a 29 nt RNA.

NOE_u and NOE_s represent observed NOE values for uniformly- and site-selectively labeled sample, respectively. This table includes the hNOE measurements for six peaks. The full results for other peaks are included in Supplementary Table S2.

Funding The funding was supported by National Science Foundation 601 (Grant No. 1808705).

Reference

1. Ganser, L.R., Kelly, M.L., Herschlag, D. & Al-Hashimi, H.M. The roles of structural dynamics in the cellular functions of RNAs. *Nat Rev Mol Cell Biol* **20**, 474-489 (2019).
2. Marusic, M., Schlagnitweit, J. & Petzold, K. RNA Dynamics by NMR Spectroscopy. *Chembiochem* **20**, 1-27 (2019).
3. Alphonse, S. & Ghose, R. Cystoviral RNA-directed RNA polymerases: Regulation of RNA synthesis on multiple time and length scales. *Virus Res* **234**, 135-152 (2017).
4. Montange, R.K. & Batey, R.T. Riboswitches: emerging themes in RNA structure and function. *Annu Rev Biophys* **37**, 117-33 (2008).
5. Furtig, B., Nozinovic, S., Reining, A. & Schwalbe, H. Multiple conformational states of riboswitches fine-tune gene regulation. *Curr Opin Struct Biol* **30**, 112-124 (2015).
6. Alexander, R.W., Eargle, J. & Luthey-Schulten, Z. Experimental and computational determination of tRNA dynamics. *FEBS Lett* **584**, 376-86 (2010).
7. Doherty, E.A. & Doudna, J.A. Ribozyme structures and mechanisms. *Annu Rev Biophys Biomol Struct* **30**, 457-75 (2001).
8. Dagenais, P., Girard, N., Bonneau, E. & Legault, P. Insights into RNA structure and dynamics from recent NMR and X-ray studies of the Neurospora Varkud satellite ribozyme. *Wiley Interdiscip Rev RNA* **8**(2017).
9. Murata, K. & Wolf, M. Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochim Biophys Acta Gen Subj* **1862**, 324-334 (2018).

10. Zhang, K. et al. Structure of the 30 kDa HIV-1 RNA Dimerization Signal by a Hybrid Cryo-EM, NMR, and Molecular Dynamics Approach. *Structure* **26**, 490-498.e3 (2018).
11. Zhang, K. et al. Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. *Nat Commun* **10**, 5511 (2019).
12. Larsen, K.P., Choi, J., Prabhakar, A., Puglisi, E.V. & Puglisi, J.D. Relating Structure and Dynamics in RNA Biology. *Cold Spring Harb Perspect Biol* **11**(2019).
13. Ray, S., Widom, J.R. & Walter, N.G. Life under the Microscope: Single-Molecule Fluorescence Highlights the RNA World. *Chem Rev* **118**, 4120-4155 (2018).
14. Stagno, J.R. et al. Structures of riboswitch RNA reaction states by mix-and-inject XFEL serial crystallography. *Nature* **541**, 242-246 (2017).
15. Palmer, A.G., 3rd. NMR characterization of the dynamics of biomacromolecules. *Chem Rev* **104**, 3623-40 (2004).
16. Barnwal, R.P., Yang, F. & Varani, G. Applications of NMR to structure determination of RNAs large and small. *Arch Biochem Biophys* **628**, 42-56 (2017).
17. Eldho, N.V. & Dayie, K.T. Internal bulge and tetraloop of the catalytic domain 5 of a group II intron ribozyme are flexible: implications for catalysis. *J Mol Biol* **365**, 930-44 (2007).
18. Berlin, K., Longhini, A., Dayie, T.K. & Fushman, D. Deriving quantitative dynamics information for proteins and RNAs using ROTDIF with a graphical user interface. *J Biomol NMR* **57**, 333-52 (2013).
19. Wagner, G. NMR relaxation and protein mobility. *Curr Opin Struct Biol* **3**, 748-754 (1993).
20. Xue, Y., Gracia, B., Herschlag, D., Russell, R. & Al-Hashimi, H.M. Visualizing the formation of an RNA folding intermediate through a fast highly modular secondary structure switch. *Nat Commun* **7**, ncomms11768 (2016).
21. Yamazaki, T., Muhandiram, R. & Kay, L.E. NMR experiments for the measurement of carbon relaxation properties in highly enriched, uniformly ^{13}C, ^{15}N-labeled proteins: application to $^{13}C^{\alpha}$ carbons. *J Am Chem Soc* **116**, 8266-8278 (1994).
22. Shajani, Z. & Varani, G. NMR studies of dynamics in RNA and DNA by ^{13}C relaxation. *Biopolymers* **86**, 348-59 (2007).
23. Johnson, J.E., Jr., Julien, K.R. & Hoogstraten, C.G. Alternate-site isotopic labeling of ribonucleotides for NMR studies of ribose conformational dynamics in RNA. *J Biomol NMR* **35**, 261-74 (2006).
24. Dayie, K.T. & Wagner, G. Carbonyl Carbon Probe of Local Mobility in ^{13}C, ^{15}N - Enriched Proteins Using High-Resolution Nuclear Magnetic Resonance. *J Am Chem Soc* **119**, 7797-7806 (1997).
25. Boisbouvier, J., Wu, Z., Ono, A., Kainosho, M. & Bax, A. Rotational diffusion tensor of nucleic acids from ^{13}C NMR relaxation. *J Biomol NMR* **27**, 133-42 (2003).
26. Hansen, A.L. & Al-Hashimi, H.M. Dynamics of large elongated RNA by NMR carbon relaxation. *J Am Chem Soc* **129**, 16072-82 (2007).
27. Johnson, J.E., Jr. & Hoogstraten, C.G. Extensive backbone dynamics in the GCAA RNA tetraloop analyzed using ^{13}C NMR spin relaxation and specific isotope labeling. *J Am Chem Soc* **130**, 16757-69 (2008).
28. Thakur, C.S. & Dayie, T.K. Asymmetry of ^{13}C labeled 3-pyruvate affords improved site specific labeling of RNA for NMR spectroscopy. *J Biomol NMR* **52**, 65-77 (2012).
29. Thakur, C.S., Luo, Y., Chen, B., Eldho, N.V. & Dayie, T.K. Biomass production of site selective $^{13}C/^{15}N$ nucleotides using wild type and a transketolase E. coli mutant for labeling RNA for high resolution NMR. *J Biomol NMR* **52**, 103-14 (2012).
30. Alvarado, L.J. et al. Chemo-enzymatic synthesis of selectively $^{13}C/^{15}N$-labeled RNA for NMR structural and dynamics studies. *Methods Enzymol* **549**, 133-62 (2014).

31. Longhini, A.P. et al. Chemo-enzymatic synthesis of site-specific isotopically labeled nucleotides for use in NMR resonance assignment, dynamics and structural characterizations. *Nucleic Acids Res* **44**, e52 (2016).
32. Kao, C., Zheng, M. & Rudisser, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **5**, 1268-72 (1999).
33. Guillerez, J., Lopez, P.J., Proux, F., Launay, H. & Dreyfus, M. A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc Natl Acad Sci U S A* **102**, 5958-63 (2005).
34. Fushman, D., Cahill, S. & Cowburn, D. The main-chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: analysis of ^{15}N relaxation with monomer/dimer equilibration. *J Mol Biol* **266**, 173-94 (1997).
35. Keeler, J. *Understanding NMR Spectroscopy*, (WILEY, Cambridge, UK, 2010).
36. Cavanagh, J., Fairbrother, W.J., Palmer, A.G., 3rd, Rance, M. & Skelton, N.J. *Protein NMR Spectroscopy*, (Academic Press, 2006).
37. Fiala, R., Czernek, J. & Sklenar, V. Transverse relaxation optimized triple-resonance NMR experiments for nucleic acids. *J Biomol NMR* **16**, 291-302 (2000).
38. Brutscher, B., Boisbouvier, J., Pardi, A., Marion, D. & Simorre, J. Improved Sensitivity and Resolution in ^{1}H-^{13}C NMR Experiments of RNA. *J Am Chem Soc* **120**, 11845-11851 (1998).
39. Goldman, M. Interference Effects in the Relaxation of a Pair of Unlike Spin - 1/2 Nuclei. *J Magn Reson* **60**, 437-452 (1984).
40. Engelke, J. & Ruterjans, H. Determination of ^{13} C $^{\alpha}$ relaxation times in uniformly ^{13} C/ ^{15} N - enriched proteins. *J Biomol NMR* **5**, 173-82 (1995).
41. Alvarado, L.J. et al. Regio-selective chemical-enzymatic synthesis of pyrimidine nucleotides facilitates RNA structure and dynamics studies. *Chembiochem* **15**, 1573-7 (2014).
42. Peng, J.W. & Wagner, G. 2D Heteronuclear NMR Measurements of Spin-Lattice Relaxation Times in the Rotating Frame of X Nuclei in Heteronuclear HX Spin Systems. *J Magn Reson* **94**, 82-100 (1991).
43. Kloiber, K., Spitzer, R., Tollinger, M., Konrat, R. & Kreutz, C. Probing RNA dynamics via longitudinal exchange and CPMG relaxation dispersion NMR spectroscopy using a sensitive ^{13}C-methyl label. *Nucleic Acids Res* **39**, 4340-51 (2011).
44. Strebitzer, E., Nussbaumer, F., Kremser, J., Tollinger, M. & Kreutz, C. Studying sparsely populated conformational states in RNA combining chemical synthesis and solution NMR spectroscopy. *Methods* **148**, 39-47 (2018).
45. Duchardt, E. & Schwalbe, H. Residue specific ribose and nucleobase dynamics of the cUUCGg RNA tetraloop motif by MNMR ^{13}C relaxation. *J Biomol NMR* **32**, 295-308 (2005).

Supplementary Materials

Deleterious Effects of Carbon-Carbon Dipolar Coupling On RNA NMR Dynamics

Hyeyeon Nam[1], Owen Becette[1], Regan M. LeBlanc[1,2], Daniel Oh[1], David A. Case[3], Theodore K. Dayie[1]

[1]Center for Biomolecular Structure and Organization, Department of Chemistry and Biochemistry, University of Maryland, College Park, MD, 20742 USA

[2]Current address: Basic Research Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD, 21702 USA

[3] Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ, 08854 USA

*Corresponding author Tel: 301-405-3165

E-mail: dayie@umd.edu

Electronic structure calculations

Calculations were carried out on 1-methyl-uracil and 9-methyl adenine using geometries optimized at the MP2/ccpVTZ level of theory. Chemical shielding tensors were computed (with GIAO orbitals) at the Hartree-Fock and MP2 levels, as well as with a "pure" density functional (OLYP)[1] and a "hybrid" model (PBE0)[2] that includes 25% of Hartree-Fock exchange. The pcSseg-2 basis, which is optimized for chemical shielding calculations[3], was used for calculations of the shielding tensors. To test basis set convergence, calculations other than MP2 were repeated at the pcSseg-3 basis set level. All calculations used the Gaussian-16 program.[4] The calculated CSA parameters varied slightly from what has been reported from Bax and co-workers[5]. Simulated relaxation rates showed marginal differences when calculated with the CSA parameter from Bax and co-workers[5]. However, it did not affect the % difference of relaxation rates between uniform and selectively labeled samples.

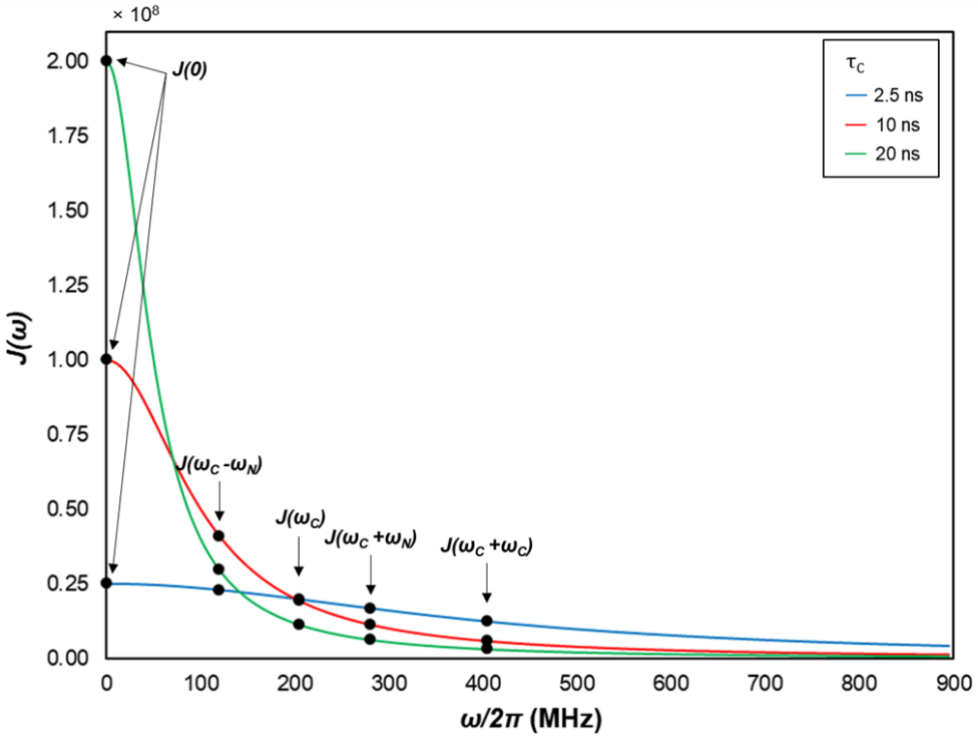

Fig. S1 A simulation of spectral density function at 800 MHz shows dramatic increase of $J(0)$ term with increasing τ_C. Compared to the change in the $J(0)$ term, there is no significant changes in other terms with different τ_C.

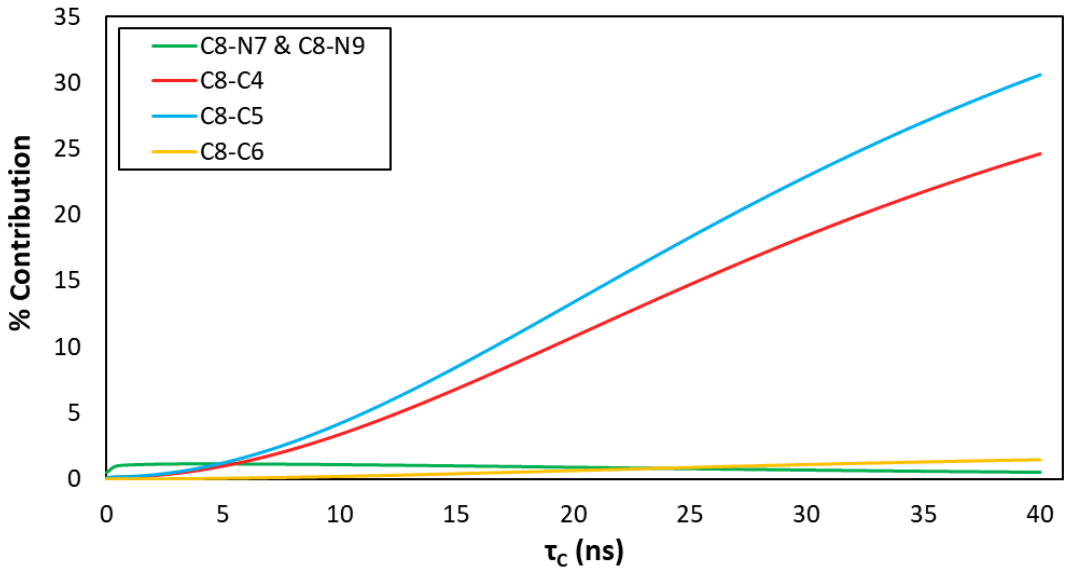

Fig. S2 Dipolar contributions of different nuclei on C8 for ATP. N7, N9 and C6 have a negligible effect (< 2%) on the R_1 rates, whereas C4 and C5 have increasing contribution with increasing size of RNAs.

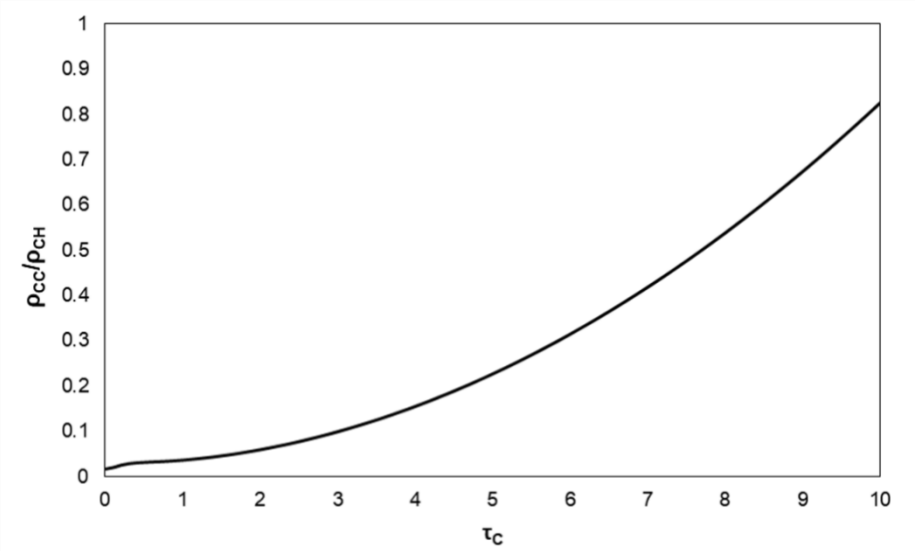

Fig. S3 Simulation of ρCC/ρCH as a function of correlation time, τ_C, at 800 MHz proton frequency. The contribution of carbon-carbon dipolar interaction to the relaxation is negligible compared to the proton-carbon interaction (ρCC/ρCH < 0. 06) when τ_C < 2 ns. However, the ^{13}C-^{13}C contribution is no longer negligible when τ_C > 2 ns. For instance, ρCC/ρCH = 0.83 for τ_C = 10 ns.

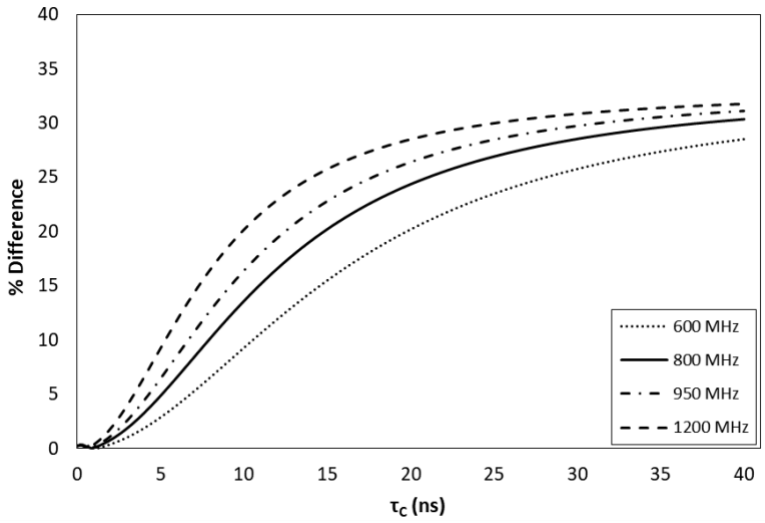

Fig. S4 Simulations of % differences between NOE_u and NOE_s for C1′ on the ribose at various magnetic field strengths. The simulation was performed as described in earlier work using the bond lengths that has been previously reported.[6,7]

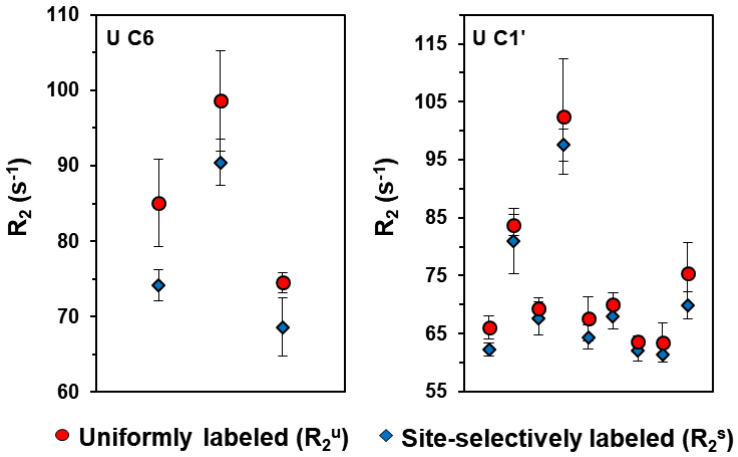

Fig. S5 R_2 rates of uniformly labeled sample (R_{2u}) and site-selectively labeled sample (R_{2s}) for a 61 nt RNA at 800 MHz.

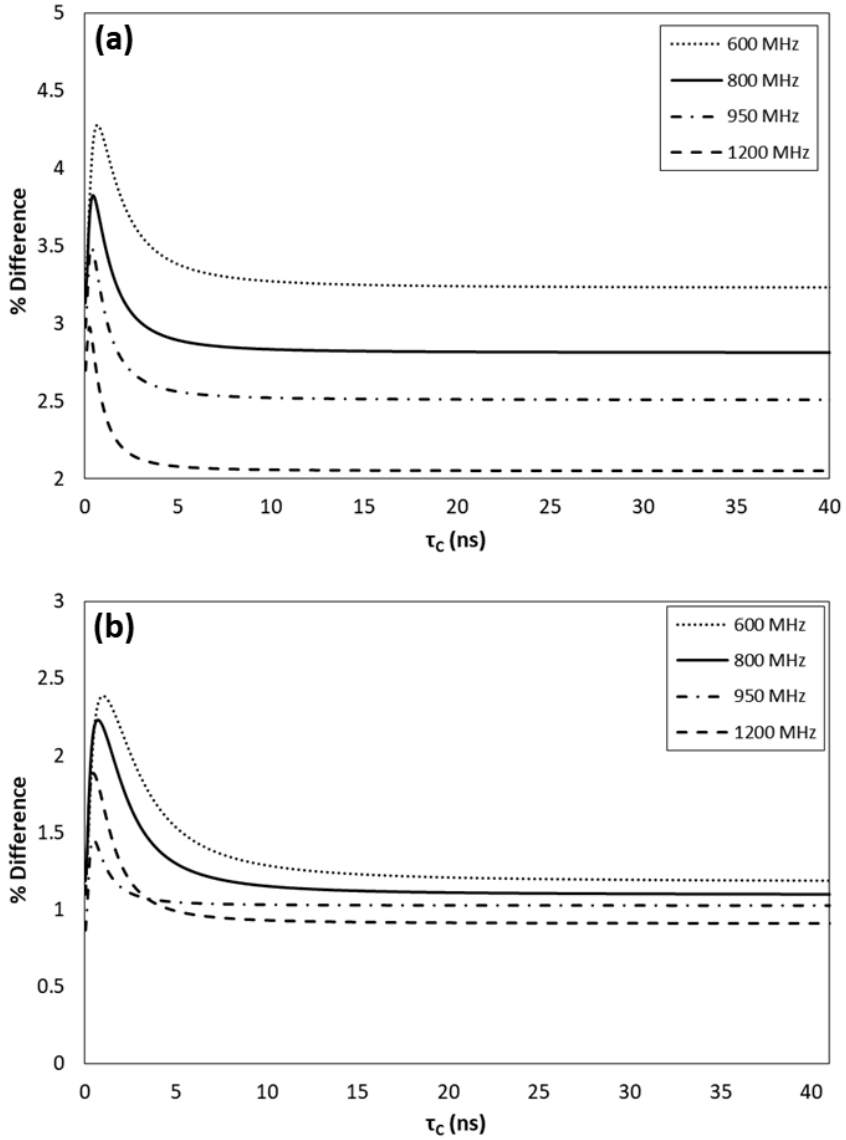

Fig. S6 Simulations of % differences between R_{2u} and R_{2s} for **a** UTP and **b** ATP. At all frequencies, the differences were less than ~ 5%. The internuclear distances used for simulation were same as the values used for the simulation in figure 2.

R_2 relaxation rates and $R_{2,CSA}$ were given by[8]:

$$R_{2,C_K,i} = \frac{1}{20} {D_{C_K,i}}^2 \left[3J(\omega_{C_K}) + 6J(\omega_{C_K} + \omega_i) + J(\omega_{C_K} - \omega_i) + 6J(\omega_i) + 4J(0)\right]$$

$$R_{2,CSA} = \frac{1}{20} {C_{C_K}}^2 \left[3J(\omega_{C_K}) + 4J(0)\right]$$

% differences were defined as: $\frac{|{R_2}^u - {R_2}^s|}{{R_2}^u} \times 100\ \%$

Adenine	Bond length (Å)	Uracil	Bond length (Å)
C8 – H8	1.079	H6 – C6	1.081
C8 – C4	2.197	C6 – C2	2.408
C8 – C5	2.118	C6 – C4	2.419
C8 – C6	3.514	C6 – C5	1.349
C8 – N7	1.329	C6 – N1	1.371
C8 – N9	1.372	C6 – N3	2.658

Table S1. The internuclear distances used for the simulation of dipolar effect.

	Selectively labeled sample			Uniformly labeled sample	
Carbon inversion pulse	Hard	Hard	Selective	Selective	Selective
Proton saturation pulse	Selective	Hard	Hard	Selective	Hard
Peak	NOE	NOE	NOE	NOE	NOE
1	1.32 ± 0.03	1.33 ± 0.01	1.35 ± 0.02	1.27 ± 0.01	1.40 ± 0.02
2	1.23 ± 0.02	1.25 ± 0.01	1.25 ± 0.01	1.22 ± 0.01	1.32 ± 0.01
3	1.18 ± 0.02	1.19 ± 0.01	1.20 ± 0.02	1.18 ± 0.02	1.27 ± 0.02
4	1.18 ± 0.01	1.17 ± 0.03	1.19 ± 0.01	1.18 ± 0.04	1.26 ± 0.02
5	1.21 ± 0.01	1.20 ± 0.01	1.22 ± 0.02	1.21 ± 0.01	1.28 ± 0.01
6	1.28 ± 0.01	1.29 ± 0.03	1.28 ± 0.01	1.26 ± 0.06	1.38 ± 0.02
7	1.28 ± 0.01	1.32 ± 0.04	1.29 ± 0.01	1.20 ± 0.04	1.35 ± 0.06
8	1.18 ± 0.02	1.19 ± 0.01	1.20 ± 0.01	1.16 ± 0.03	1.24 ± 0.02
9	1.20 ± 0.01	1.20 ± 0.02	1.21 ± 0.01	1.20 ± 0.01	1.26 ± 0.01
10	1.17 ± 0.05	1.20 ± 0.03	1.21 ± 0.02	1.17 ± 0.01	1.27 ± 0.03
11	1.14 ± 0.04	1.18 ± 0.01	1.19 ± 0.02	1.16 ± 0.01	1.24 ± 0.01
12	1.16 ± 0.05	1.18 ± 0.03	1.20 ± 0.02	1.13 ± 0.04	1.24 ± 0.02
13	1.17 ± 0.02	1.19 ± 0.01	1.18 ± 0.02	1.17 ± 0.01	1.25 ± 0.02
14	1.31 ± 0.02	1.33 ± 0.02	1.34 ± 0.01	1.25 ± 0.01	1.39 ± 0.02
15	1.14 ± 0.01	1.17 ± 0.03	1.17 ± 0.01	1.15 ± 0.01	1.22 ± 0.01
16	1.20 ± 0.01	1.21 ± 0.03	1.22 ± 0.01	1.19 ± 0.05	1.26 ± 0.01
17	1.23 ± 0.02	1.23 ± 0.03	1.24 ± 0.01	1.22 ± 0.04	1.29 ± 0.01

Table S2. NOE measurements with different pulses applied for carbon inversion and proton saturation for uniformly- and site-selectively labeled RNA.

RNA sequences and buffer conditions

1. 14-nt RNA

Sequence: 5' – GGCACUUCGGUGCC – 3'

Buffer: 20 mM potassium phosphate (pH 6.4), 0.4 mM EDTA

2. 29-nt RNA

Sequence: 5' – GGCAGAUCUGAGCCUGGGAGCUCUCUGCC – 3'

Buffer: 15 mM sodium phosphate (pH 6.4), 0.1 mM EDTA, 25 mM NaCl

3. 61-nt RNA

Sequence:

5' - GGUUCAUGUCCUACUGUUCAAGCCUCCAAGCUGUGCCUUGGGUGGCUUUG GGGCAUGGACC – 3'

Buffer: 10 mM sodium phosphate (pH 6.5), 0.1 mM EDTA

Reference

1. Handy, N.C. & Cohen, A.J. Left-right correlation energy. *Mol. Phys.* **99**, 403-412 (2001).
2. Adamo, C. & Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **110**, 6158-69 (1999).
3. Jensen, F. Segmented contracted basis sets optimized for nuclear magnetic shielding. *J Chem Theory Comput* **11**, 132-8 (2015).
4. Frisch, M.J. et al. Gaussian 16, Revision A.03. *Gaussian, Inc., Wallingford CT* (2016).
5. Ying, J., Grishaev, A., Bryce, D.L. & Bax, A. Chemical shift tensors of protonated base carbons in helical RNA and DNA from NMR relaxation and liquid crystal measurements. *J Am Chem Soc* **128**, 11443-54 (2006).
6. Yamazaki, T., Muhandiram, R. & Kay, L.E. NMR experiments for the measurement of carbon relaxation properties in highly enriched, uniformly ^{13}C, ^{15}N-labeled proteins: application to $^{13}C^{\alpha}$ carbons. *J Am Chem Soc* **116**, 8266-8278 (1994).
7. Gelbin, A. et al. Geometric Parameters in Nucleic Acids: Sugar and Phosphate Constituents. *J Am Chem Soc* **118**, 519-529 (1996).
8. Dayie, T.K. & Wagner, G. Carbonyl Carbon Probe of Local Mobility in ^{13}C,^{15}N-Enriched Proteins Using High-Resolution Nuclear Magnetic Resonance. *J Am Chem Soc* **119**, 7797-7806 (1997).