BIOPHYSICS

Solution NMR readily reveals distinct structural folds and interactions in doubly ^{13}C- and ^{19}F-labeled RNAs

Owen B. Becette[1], Guanghui Zong[1], Bin Chen[1], Kehinde M. Taiwo[1], David A. Case[2], T. Kwaku Dayie[1]*

RNAs form critical components of biological processes implicated in human diseases, making them attractive for small-molecule therapeutics. Expanding the sites accessible to nuclear magnetic resonance (NMR) spectroscopy will provide atomic-level insights into RNA interactions. Here, we present an efficient strategy to introduce ^{19}F-^{13}C spin pairs into RNA by using a 5-fluorouridine-5′-triphosphate and T7 RNA polymerase–based in vitro transcription. Incorporating the ^{19}F-^{13}C label in two model RNAs produces linewidths that are twice as sharp as the commonly used ^{1}H-^{13}C spin pair. Furthermore, the high sensitivity of the ^{19}F nucleus allows for clear delineation of helical and nonhelical regions as well as GU wobble and Watson-Crick base pairs. Last, the ^{19}F-^{13}C label enables rapid identification of a small-molecule binding pocket within human hepatitis B virus encapsidation signal epsilon (hHBV ε) RNA. We anticipate that the methods described herein will expand the size limitations of RNA NMR and aid with RNA-drug discovery efforts.

INTRODUCTION

RNAs form essential regulators of biological processes and are implicated in human diseases, making them attractive therapeutic targets (*1*, *2*). This extensive functional diversity of RNA derives from its ability to fold into complex three-dimensional (3D) structures. Yet, the number of noncoding RNA sequences far outstrips the number of solved RNA structures deposited in the Protein Data Bank (PDB) necessary for understanding RNA function (*3*, *4*). In comparison to x-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy provides high-resolution structural and dynamic information in solution, making it an ideal biophysical technique to characterize the interactions between target RNAs and small drug-like molecules. Nonetheless, NMR studies of RNA suffer from poor spectral resolution and sensitivity, both of which worsen with increasing molecular weight. In contrast with proteins, which are made up of 20 unique amino acid building blocks, RNAs are composed of only four aromatic residues. These four resonate over a very narrow chemical shift region. At high magnetic field strengths, sizable transverse relaxation rates (R_2) cause line broadening and thereby decrease both sensitivity and resolution. These problems are further exacerbated with increasing molecular weight. To overcome these limitations of RNA, novel labeling strategies that expand the number of NMR probes beyond the traditional nonradioactive and stable isotope labels such as hydrogen-1 (^{1}H), phosphorus-31 (^{31}P), carbon-13 (^{13}C), hydrogen-2 (^{2}H), and nitrogen-15 (^{15}N) are needed.

Solution NMR of the magnetically active fluorine-19 (^{19}F) isotope offers clear advantages in the study of RNA structure and conformational changes, which occur upon ligand binding. ^{19}F has high NMR sensitivity (0.83 of ^{1}H) due to a large gyromagnetic ratio that is comparable to ^{1}H (0.94 of ^{1}H), a 100% natural abundance, and ~6 wider chemical shift dispersion than ^{1}H (*5*, *6*). In addition, ^{19}F is also sensitive to changes in its local chemical environment (*5*, *6*). In contrast with other commonly used NMR nuclei (^{1}H/^{31}P/^{13}C/^{15}N), ^{19}F is virtually absent in biological systems, thereby rendering ^{19}F NMR background free. Together, ^{19}F is an attractive probe for incorporation into nucleic acids to study their structure, interactions, and dynamics in solution.

Given its attractive spectroscopic properties, ^{19}F was incorporated into RNA for NMR studies in the 1970s (*7*–*9*). Since then, ^{19}F has been successfully incorporated into DNA and RNA oligonucleotides for NMR analysis and used to probe RNA and DNA structure, conformational exchange, and macromolecular interactions (*10*, *11*). Most of these studies were conducted on short oligonucleotides [~30 nucleotides (nt)] prepared by solid-phase synthesis with only a few residues ^{19}F labeled. Even when 2-fluoroadenine (2FA) was incorporated into a 73-nt (~22 kDa) guanine-sensing riboswitch, only 4 of the 16 signals could be assigned. This 2FA study hinted at the limitations of ^{19}F NMR for large RNAs (*12*). Despite its attractiveness, the application of ^{19}F NMR to study RNA has remained limited because the large ^{19}F chemical shift anisotropy (CSA) contributes substantially to line broadening as a function of increasing molecular weight and polarizing magnetic fields.

To circumvent this limitation, Boeszoermenyi *et al.* (*13*) recently showed that direct coupling of ^{19}F to ^{13}C allowed for cancelation of CSA and dipole-dipole (DD) interactions. By incorporating this ^{19}F-^{13}C spin pair into aromatic moieties of proteins and a 16-nt DNA, they showed that a transverse relaxation optimized spectroscopy (TROSY) version of a ^{19}F-^{13}C heteronuclear single-quantum coherence (HSQC) (*13*) provided improved spectroscopic properties. These exciting results hinted that installing ^{13}C-^{19}F pairs in RNA nucleobases should also lead to improved spectroscopic features.

However, there were no facile methods to readily incorporate ^{19}F-^{13}C spin pairs into RNA. To overcome this technical obstacle of incorporating fluorinated aromatic moieties into RNA, we provide here a straightforward chemoenzymatic synthesis of [5-^{19}F, 5-^{13}C]-uridine 5′-triphosphate (5FUTP) for incorporation into RNA (Fig. 1) using phage T7 RNA polymerase–based in vitro transcription. To showcase its versatility, we transcribed two model RNAs using these labels: the 30-nt (~10-kDa) human immunodeficiency type 2 transactivation response (HIV-2 TAR) element (*6*, *14*) and the 61-nt (~20-kDa) human hepatitis B virus encapsidation signal epsilon (hHBV ε) element (Fig. 1) (*15*, *16*).

[1]Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20782, USA. [2]Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854, USA.
*Corresponding author. Email: dayie@umd.edu

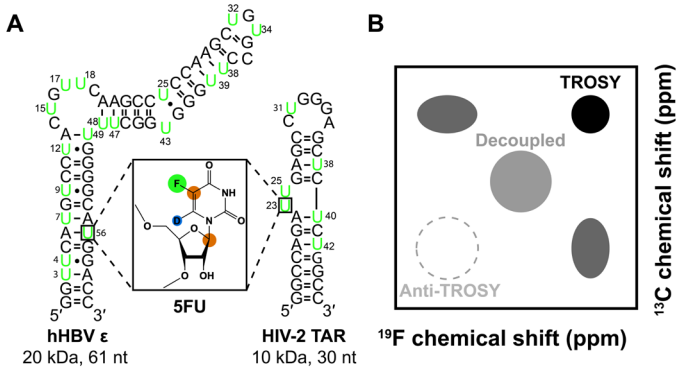

Fig. 1. RNA structures and NMR methods described in this study. (**A**) Model RNA systems: HIV-2 TAR (30 nt, 10 kDa) and hHBV ε (61 nt, 20 kDa). Residues highlighted in green are labeled with ^{19}F-^{13}C–5-fluorouridine (5FU) shown in the box. Green circle, ^{19}F; brown circle, ^{13}C; blue circle, ^{2}H. (**B**) Theoretical ^{19}F,^{13}C spectrum showing the four observable magnetization components of the ^{19}F-^{13}C spin pair as well as the decoupled resonance that has the average chemical shift and linewidths of all four components.

With our new labels, we demonstrate several advantages for RNA NMR studies, including improved resolution and increased sensitivity to ligand binding. We show that a ^{19}F substitution is structurally nonperturbing and has an optimal TROSY effect at readily available magnetic field strength of 600 MHz (^{1}H frequency), in agreement with previous studies (*13*). Unlike C-H spectra, the resolving power of ^{19}F allows for easy identification of RNA structural elements in helical and nonhelical regions, as well as in wobble GU base-paired regions. With protons substituted with deuterium and depending on the molecular weight of the RNA, the TROSY effect in the ^{19}F-^{13}C pair can reduce the ^{13}C linewidth by a factor >2, compared to a ^{13}C-^{1}H pair, and the ^{19}F-^{13}C label enables detection of a small-molecule binding to a 20-kDa RNA. Thus, our ^{19}F-^{13}C label overcomes several of the limitations in sensitivity and resolution facing RNA NMR studies with the potential to extend the application of solution NMR measurements to larger–molecular weight systems in vivo.

RESULTS

Chemical synthesis of 5-fluorouracil aids enzymatic synthesis of 5-fluorouridine-5′-triphosphate

Given the potential utility but unavailability of ^{19}F-^{13}C spin pairs in aromatic moieties of RNA, we first sought to develop a reliable and scalable method that combined chemical synthesis with enzymatic coupling in almost quantitative yields. This chemoenzymatic approach is a versatile method that combines chemical synthesis of atom-specific labeled nucleobases with commercially available selectively labeled ribose using enzymes from the pentose phosphate pathways (PPPs) (*3*, *4*). To this end, we adapted the method of Santalucia *et al.* (*17*) and Kreutz and co-workers (*18*) and first synthesized the uracil base (U) specifically labeled with ^{13}C at the aromatic C5 and ^{15}N at the N1 and N3 positions (Fig. 1). This synthesis is readily accomplished using unlabeled potassium cyanide, ^{13}C-labeled bromoacetic acid, and ^{15}N-labeled urea. The resulting U was converted to 5-fluorouracil (5FU) by direct fluorination with Selectfluor (*19*, *20*). This strategy allows for efficient and cost-effective synthesis of the 5FU base with high yield of ~63%. In addition, to remove unwanted scalar coupling interactions (*14*), we selectively deuterated H6 (~95%) using well-established methods (*21*). Next, using enzymes from the PPP, we coupled 5FU to D-ribose labeled at the C1′ position to give 5FUTP (Fig. 1) (*3*, *22*) with an overall yield of ~50%. This site-specifically labeled 5FUTP was then used for DNA template–directed T7 RNA polymerase–based in vitro transcription with overall yields comparable to those obtained with unmodified nucleotides.

Thermal stabilities of wild-type and 5FU HIV-2 TAR and hHBV ε

Fluorine substitution at uridine C5 is thought to reduce the imino N3 p*K*a values by about 1.7 to 1.8 units with respect to their protonated analogs (*23*), leading to extensive line broadening of imino protons in 5FU RNAs (*24*). To determine if incorporation of 5-fluorouridine alters the folding thermodynamics of our RNAs (Fig. 1), we recorded ultraviolet (UV) thermal melting profiles for both wild-type (WT) and 5FU HIV-2 TAR and hHBV ε (table S1). Both WT and 5FU RNAs showed a single transition in their melting profiles, consistent with unimolecular folding (*25*). WT and 5FU HIV-2 TAR had melting temperatures within ~1 K of each other (WT: T_m = 355.6 ± 0.5 K; 5FU: T_m = 357.4 ± 0.4 K). Similarly, 5FU hHBV ε had a melting temperature of 327.1 ± 0.1 K, which is within the error of the melting temperature of WT. Together, these results suggest that 5FU does not markedly alter the thermodynamic stability of HIV-2 TAR and hHBV ε, in accordance with previous studies of 5FU RNAs (*6*, *7*, *24*).

DD-CSA cross-correlation in ^{19}F-^{13}C 5FU and TROSY effect

The linewidth for aromatic ^{19}F-^{13}C spin pair (Fig. 1B) is expected to become dominated by the CSA mechanism with increasing polarizing magnetic fields (*13*). To estimate this effect for 5FU, we calculated the chemical shielding tensor (CST) for ^{19}F-^{13}C spin pairs using density functional theory (DFT) methods (tables S2 and S3) (*26*–*29*). Using these CST parameters and relaxation theory implemented in the Spinach library (*30*), we computed the TROSY R_2 relaxation rates for the ^{19}F-^{13}C pair of 5FU ($^{13}C_F$ and $^{19}F_C$) and the ^{13}C-^{1}H pair of U ($^{13}C_H$ and $^{1}H_C$) (Fig. 2) assuming isotropic tumbling. The R_2 of fluorinated carbon ($^{13}C_F$) TROSY resonance is ~2 times smaller than that of the protonated carbon ($^{13}C_H$) at their respective minima of ~600 and ~950 MHz, respectively, for all molecular weights greater than 5 ns (Fig. 2A). Compared with the decoupled resonance, the R_2 of the $^{13}C_F$ TROSY resonance is ~3 times smaller than that of protonated carbon for all molecular weights greater than 5 ns (fig. S1). Although the TROSY effect is quite small for ^{19}F nuclei bonded to ^{13}C ($^{19}F_C$) and for ^{1}H nuclei bonded to ^{13}C ($^{1}H_C$), the R_2 of $^{19}F_C$ is three times bigger than that of $^{1}H_C$ (fig. S2). Thus, sensitive, high-resolution NMR spectra for the ^{19}F-^{13}C pair of 5FU in RNAs can be obtained by selective detection of the $^{13}C_F$ TROSY resonance as demonstrated for the ^{19}F-^{13}C pair in aromatic amino acids (*13*).

^{19}F-^{13}C TROSY on 5FU HIV-2 TAR and hHBV ε

To validate these theoretical TROSY predictions experimentally, we adapted the ^{1}H-^{15}N TROSY experiment (*3*, *31*, *32*) to perform a ^{19}F-^{13}C TROSY experiment on a ~10-kDa 5FU HIV-2 TAR and on a ~20-kDa 5FU hHBV ε RNAs (Fig. 3). Because of hardware limitations, we could only run experiments that start with and end on the magnetization of ^{19}F, with the ^{13}C frequency encoded in the indirect dimension. That is, we used the so-called ^{19}F-detected out-and-back method, rather than the more sensitive ^{19}F-excited out-and-stay ^{13}C-detected experiment (*13*). We collected spectra for each of the

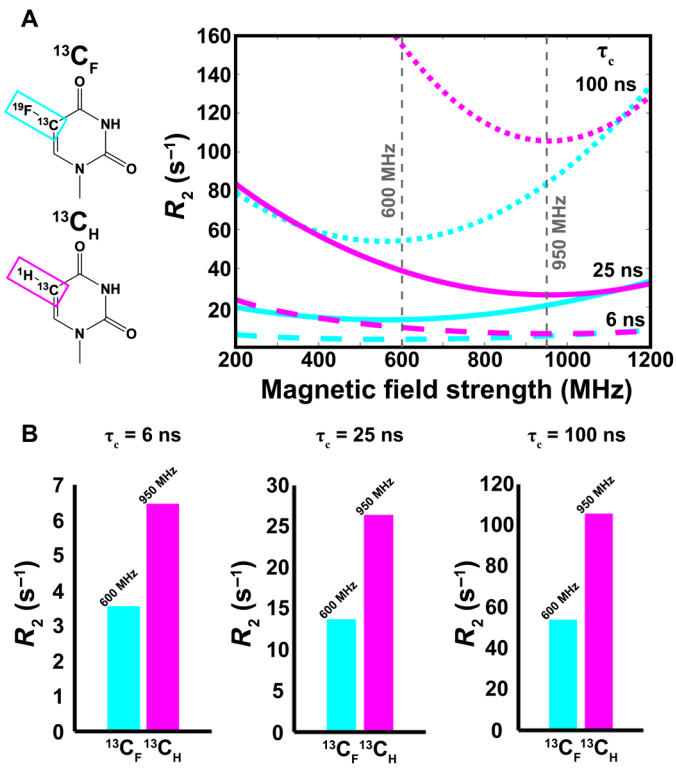

Fig. 2. Theoretical R_2 values for the TROSY components for ^{19}F-^{13}C of 5-fluorouracil ($^{13}C_F$) or ^{1}H-^{13}C of uracil ($^{13}C_H$). (**A**) Theoretical curves showing the expected R_2 values for the TROSY component of $^{13}C_F$ (cyan) and $^{13}C_H$ (magenta) as a function of magnetic field strength (relative to ^{1}H Larmor frequency) for τ_c = 6 ns (dashed line), 25 ns (solid line), and 100 ns (dotted line) at 25°C. (**B**) Theoretical R_2 values taken at the commercially available magnetic field strength closest to the maximum TROSY effect ($^{13}C_H$ = 950 MHz; $^{13}C_F$ = 600 MHz) for τ_c = 6, 25, and 100 ns at 25°C.

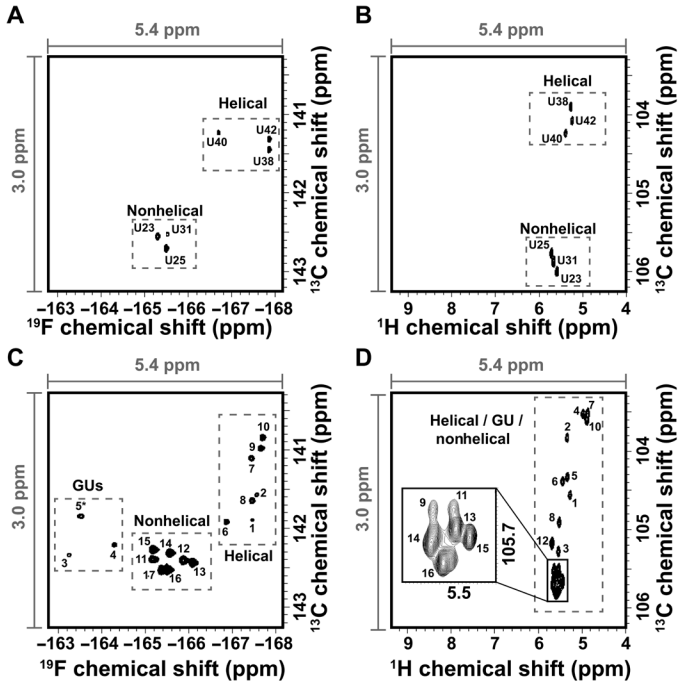

Fig. 3. TROSY spectra of RNA systems. (**A**) ^{19}F-^{13}C TROSY of 5FU HIV-2 TAR. (**B**) ^{1}H-^{13}C TROSY of WT HIV-2 TAR. (**C**) ^{19}F-^{13}C TROSY of 5FU hHBV ε. (**D**) ^{1}H-^{13}C TROSY of WT hHBV ε. The assignments of 5FU and WT TAR-2 are indicated, as well as the arbitrary peak numbers for 5FU and WT hHBV ε. The same window size was used in all four spectra to aid in comparison. Gray dashed boxes indicate signals from helical, GU, and nonhelical regions. For (D), the black box indicates a zoom-in view of poorly resolved signals.

four components (Fig. 1B) of the ^{19}F-^{13}C (^{1}H-^{13}C) correlations for both 5FU (WT) HIV-2 TAR and hHBV ε (figs. S3 to S6).

Both HIV-2 TAR and hHBV ε show ~6-fold improvement in chemical shift dispersion of ^{19}F compared with ^{1}H and similar dispersion in ^{13}C (Fig. 3). All six correlations of HIV-2 TAR are well resolved for both ^{1}H-^{13}C and ^{19}F-^{13}C correlations and are in agreement with previously published ^{1}H-^{19}F and ^{1}H-^{13}C RNA spectra (*6*, *24*, *33*). Nonetheless, even for this small RNA, the ^{19}F-^{13}C spin pair markedly improves the spectral resolution. 5FU HIV-2 TAR shows a chemical shift dispersion of 2.6 parts per million (ppm) in the ^{19}F dimension and only 0.5 ppm in the ^{1}H dimension for WT (Fig. 3, A and B). Replacing ^{1}H with ^{19}F at C5 results in a slight reduction in chemical shift dispersion along the ^{13}C dimension from 2.1 to 1.5 ppm, although this effect is much smaller than the gain in resolution for ^{19}F over ^{1}H (Fig. 3, A and B). Similarly, the ^{19}F resonances of 5FU hHBV ε are spread over 4.5 ppm, whereas the WT ^{1}H signals resonate over a narrow 0.8-ppm window. This represents 5.7 times better dispersion (Fig. 3, C and D). Again, substitution of ^{1}H with ^{19}F at C5 results in a reduction in chemical shift dispersion of 2.3 to 1.7 ppm along the ^{13}C dimension for hHBV ε (Fig. 3, C and D). Of the anticipated 18 signals for hHBV ε, 16 are resolved for WT and 17 for 5FU. Together, these results demonstrate the marked gain in resolution afforded by the ^{19}F-^{13}C spin pair in 5FU RNAs compared with the ^{1}H-^{13}C spin pair in WT.

In addition to this considerable gain in resolution, ^{19}F-^{13}C labeling confers favorable $^{13}C_F$ TROSY linewidths. We compared the relative linewidths for both RNAs, which we assume to be Lorentzian (Figs. 4 and 5). For 5FU HIV-2 TAR, the $^{13}C_F$ TROSY linewidths were 1.5 times sharper on average than the anti-TROSY components, with a range of 1.3 to 1.7 (Fig. 4A). For WT HIV-2 TAR, the $^{13}C_H$ TROSY component was 3.7-fold narrower than the anti-TROSY component (range, 1.6 to 8.7) (Fig. 4B). Similarly, for 5FU HBV ε, the $^{13}C_F$ TROSY linewidths were 2.2-fold narrower than the anti-TROSY ones over a range of 1.5 to 3.3 (Fig. 4C). For WT HBV ε, only 5 of the 16 $^{13}C_H$ anti-TROSY signals were observed and were 2.6 times broader than the TROSY resonances (range, 2.0 to 3.3) (Fig. 4D). As predicted from our simulations (Fig. 2), the $^{13}C_F$ TROSY component relaxes ~2 times slower than the $^{13}C_H$ TROSY component in both HIV-2 TAR and hHBV ε. The $^{19}F_C$ TROSY linewidths for 5FU HIV-2 TAR and 5FU HBV ε were 1.4 (range, 1.3 to 1.6) and 1.6 (range, 1.1 to 2.5) times narrower than the anti-TROSY components, respectively (Fig. 5, A and C). For both WT HIV-2 and WT HBV ε, the ^1H_C TROSY and anti-TROSY linewidths were comparable (Fig. 5, B and D). Consistent with our simulations, the $^{19}F_C$ TROSY linewidth is ~2-fold larger than that of the ^1H_C component for both RNAs (fig. S3). Again, this is in line with the poor performance of ^{19}F NMR experiments due to the large CSA-induced relaxation. Thus, the incorporation of the ^{13}C label mitigates the deleterious relaxation of the ^{19}F nuclei within a ^{19}F-^{13}C spin pair. However, even for medium-sized RNAs ~20 kDa, ^{19}F TROSY detection of the ^{19}F-^{13}C spin pair still outperforms that for a ^{1}H-^{13}C

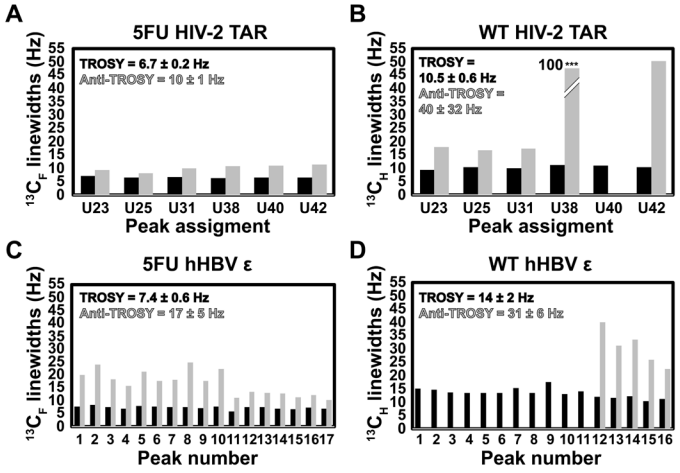

Fig. 4. Measured ^{13}C linewidths for 5FU and WT HIV-2 TAR and hHBV ε. Quantification of TROSY (black) and anti-TROSY (gray) (**A**) $^{13}C_F$ and (**B**) $^{13}C_H$ linewidths for HIV-2 TAR. Note that U40 was not observed in the anti-TROSY spectrum of WT HIV-2 TAR (B). In addition, the anti-TROSY component of U38 in (B) was 97 Hz and truncated to fit in the plot. Quantification of TROSY (black) and anti-TROSY (gray) (**C**) $^{13}C_F$ and (**D**) $^{13}C_H$ linewidths for hHBV ε. Note that peaks 1 through 11 in WT hHBV ε were not observed in the anti-TROSY spectrum (D). The average ± SD in Hz is shown for the TROSY and anti-TROSY components in each plot. Peak numbers and assignments are given in Fig. 3.

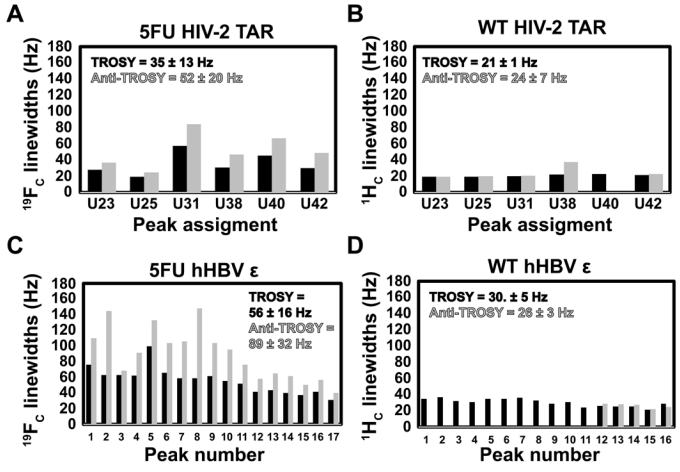

Fig. 5. Measured ^{19}F and ^{1}H linewidths for 5FU and WT HIV-2 TAR and hHBV ε. Quantification of TROSY (black) and anti-TROSY (gray) (**A**) $^{19}F_C$ and (**B**) $^{1}H_C$ linewidths for HIV-2 TAR. Quantification of TROSY (black) and anti-TROSY (gray) (**C**) $^{19}F_C$ and (**D**) $^{1}H_C$ linewidths for hHBV ε. The average ± SD in Hz is shown for the TROSY and anti-TROSY components in each plot. Peak numbers and assignments are given in Fig. 3.

spin pair. Therefore, to reap the maximum benefits of this label, it is advantageous to monitor the ^{13}C nuclei rather than the ^{19}F nuclei. We anticipate that the ^{19}F-^{13}C TROSY effect will continue to scale with molecular weight for RNAs as was seen recently with proteins (*13*) and our simulations.

^{19}F chemical shifts reveal RNA structural equilibria

In addition to these gains in resolution and favorable linewidths, previous work suggested the ^{19}F chemical shifts serve as sensitive markers of RNA secondary structure (*10, 11*). For example, GU wobble base pairs are deshielded and shifted by ~4.5 ppm to lower fields compared with AUs within Watson-Crick geometries (*34*). On the basis of these earlier observations, we hypothesized that ^{19}F-^{13}C correlations of HIV-2 TAR and hHBV ε can be grouped on the basis of whether or not they are in helical, nonhelical, or GU base-paired regions of the RNA. As a positive control, we note that nonhelical U23, U25, and U31 in 5FU HIV-2 TAR resonate around ~−165.5 ppm in ^{19}F and ~142.5 ppm in ^{13}C (Fig. 3A). On the other hand, the helical residues U38, U40, and U42 of 5FU HIV-2 TAR are centered around ~−167.5 ppm in ^{19}F and ~141.5 ppm in ^{13}C in line with previous observations for ^{19}F-^{1}H samples of HIV-2 TAR (*6*) and tRNA (*34*). Comparison of the equivalent ^{1}H-^{13}C spectra shown in Fig. 3B indicates that even though helical residues cannot be distinguished from nonhelical residues in the ^{1}H dimension, nonhelical residues can be differentiated from helical base pairs in the ^{13}C dimension for a ^{1}H-^{13}C spin pair.

The 17 ^{19}F-^{13}C resolved correlations of 5FU hHBV ε show similar clustering as 5FU HIV-2 TAR (Fig. 3C). For instance, the six most intense signals are centered around ~−165.5 ppm in ^{19}F and ~142.5 ppm in ^{13}C where the nonhelical signals of HIV-2 TAR are located. On the basis of the secondary structure of hHBV ε (Fig. 1A), these six intense peaks belong to the six nonhelical uridines (U15, U17, U18, U32, U34, and U43) (Fig. 3C). A seventh peak is also seen in this region, most likely due to U48 or U49, both of which flank the bulge region. The weaker peaks are from the helical portions of hHBV ε because these signals located at ~−167.5 ppm in ^{19}F and ~141.5 ppm in ^{13}C resonate in the same region as the helical signals from 5FU HIV-2 TAR (*6*). HIV-2 TAR contains only Watson-Crick base pairs, and so, signals in this region of the hHBV ε spectrum correspond to AUs (U3, U7, U38, U39, U47, U48, U49, and U56). Of the eight anticipated peaks belonging to helical residues, only seven are observed, further suggesting that U48 or U49 may fray and resonate within the nonhelical region. Unlike HIV-2 TAR, hHBV ε has four noncanonical GU wobble base pairs embedded within helical regions. The three signals resonating in a distinct region centered at ~−163.5 ppm in ^{19}F and ~142.0 ppm in ^{13}C are from the four GUs (U4, U9, U12, and U25). This is in line with previous observations of GU base pairs in tRNA (*34*). Peak 5 (Fig. 3C) is most likely two GUs that are overlapped. Again, comparison of the equivalent ^{1}H-^{13}C spectra shown in Fig. 3D indicates that even though helical residues can be distinguished from nonhelical residues, nonhelical residues cannot be differentiated from GU base pairs for a ^{1}H-^{13}C spin pair. Thus, the spectroscopic discrimination of helical and nonhelical regions as well as GU wobble and Watson-Crick base pairs in RNA structures becomes possible with the high sensitivity of ^{19}F to the local chemical environment of a ^{19}F-^{13}C spin pair. This distinguishing feature is not readily available for a ^{1}H-^{13}C spin pair.

^{19}F chemical shift perturbation enables facile identification of site-specific RNA binders

Ligand-based (*35*) and protein-observed (*36*) ^{19}F NMR screening methods are important for identifying small drug-like molecules that act as protein inhibitors. Although most work to date has focused on proteins, recent work suggests that RNAs also contain specific binding pockets that could be easily distinguished and targeted with small molecules (*1, 2*). hHBV ε is at the center of the viral replication cycle since the first two residues in its internal bulge are used

by the virus to initiate synthesis of the minus-strand DNA. Thus, targeting this RNA structure will notably expand the repertoire of HBV drug targets beyond the current focus on viral proteins (*37*). Given ^{19}F chemical shifts serve as sensitive markers of RNA secondary structure, we reasoned that ^{19}F-^{13}C spectroscopy will likely pinpoint loop over helical region binders. Rather satisfyingly, we found a small molecule that specifically binds a subset of nonhelical residues in 5FU hHBV ε (Fig. 6). Overlay of the full spectra of 5FU hHBV ε with and without the small-molecule shows chemical shift perturbations (CSPs) (*38*) predominantly confined to nonhelical regions (Fig. 6). Within the nonhelical residues, only four of the seven signals shift with the addition of the small molecule, which suggests selectivity for certain nonhelical residues over others (Fig. 6). We propose a model whereby our small molecule binds hHBV ε in the 6-nt bulge formed between C14 and C19, but not anywhere else in the RNA. The minor CSPs seen in the helical portion of the 5FU hHBV spectra are from U residues flanking the 6-nt bulge, specifically U47, U48, and U49. Last, the CSP seen in the GU portion is from U12, which also flanks our proposed binding pocket.

DISCUSSION

^{19}F is an attractive spectroscopic probe to study biomolecular structure, interactions, and dynamics in solution. Nonetheless, a number of obstacles must be overcome for it to become widely useful. First, we must be able to easily install the label into any biopolymer. While incorporation of fluorinated aromatic amino acids and nucleobases into proteins and nucleic acids is usually not a technical challenge, until now, synthesis of carbon-labeled and fluorinated nucleobase to create a ^{19}F-^{13}C spin pair has been problematic for RNA. Here, we present a facile strategy to incorporate ^{19}F-^{13}C 5-fluorouridine into RNA using in vitro transcription for characterization of small-molecule binding interactions by NMR. Our protocol to prepare ^{19}F-^{13}C 5-fluorouridine-5′-triphosphate (5FUTP) involves chemically synthesizing 5FU and then enzymatically coupling it to ^{13}C-labeled D-ribose. Our synthetic strategy can be generalized to selectively place labels in the pyrimidine nucleobase at either ^{15}N1, ^{15}N3, ^{13}C2, ^{13}C4, ^{13}C5, or ^{13}C6 or any combinations thereof, and then enzymatically couple ribose labeled at either ^{13}C1′, ^{13}C2′, ^{13}C3′, ^{13}C4′, or ^{13}C5′ or any of the preceding ribose combinations to the base. The resulting isotopically enriched 5FUTP is then readily incorporated into any desired RNA using DNA template–directed T7 RNA polymerase–based in vitro transcription. This enzymatic approach, unlike solid-phase RNA synthesis, is not limited to RNAs less than 70 nt or to nucleotides made of labeled nucleobase coupled to unlabeled ribose. Although fluorine substitution at C5 in pyrimidines strongly affects the shielding of the nearby H6, it has little effect on the anomeric H1′ chemical shifts (*24*). We therefore anticipate that our unique strategy that combines ribose ^{13}C1′ label with ^{19}F-^{13}C uracil should allow the transfer of assignments from unmodified RNAs to 5-fluoropyrimidine–substituted RNAs made with our labels.

Second, because of van der Waals radii comparable to that of ^{1}H, ^{19}F is considered minimally perturbing when incorporated into biopolymers (*24*). Although fluorine substitution in 5FU RNAs leads to sizeable line broadening of the imino protons, thermal melting analysis indicates that the 5FU RNAs are thermodynamically equivalent to the nonfluorinated RNAs (*6*, *7*, *24*). In future work, it will be important to systematically investigate the effect of fluorine substitution not only on thermodynamic stability but also on folding kinetics of RNAs. Insights derived from solving, at high-resolution, the 3D structures of fluorinated and nonfluorinated RNA could potentially guide the use of these spin pairs to spy on the biological processes within the cell.

Third, despite its huge potential, nucleic acid observed ^{19}F (NOF) NMR has remained underused because the large ^{19}F CSA induces severe line broadening at high molecular weights and magnetic fields. Using DFT calculations of CST parameters, we show that an optimal ^{19}F-^{13}C TROSY enhancement occurs at 600-MHz ^{1}H frequency to enable slow relaxation of ^{13}C bonded to ^{19}F. Our RNAs show an enhanced ^{19}F-^{13}C TROSY effect with increasing molecular weight and ^{13}C linewidths that are twice as sharp as seen with traditional ^{1}H-^{13}C spin pairs. Thus, nucleobase ^{19}F -^{13}C TROSY will expand the applicability of RNA NMR beyond the ~30-nt (~10-kDa) average.

Fourth, the RNA secondary structure is made up of segments of nucleotides that are either base paired or not. The arrangements of base-paired with unpaired regions can leave distinct NMR chemical shift signatures that can provide low-resolution structural information with minimum expenditure of time and cost. For example, the H5 of a pyrimidine is sensitive to the nature of the residue that comes before it within a triplet of canonical Watson-Crick AU and GC base pairs. When the A in a central UA base pair is substituted by a G, the H5 resonance shifts downfield because of the formation of the GU base pair. Yet, an analysis of the commonly used ^{1}H-^{13}C probes fails to unambiguously separate nonhelical residues from helical ones (*39*). In contrast, the ^{19}F-^{13}C labels resonate in distinct chemical shift regions based on their secondary structure. For instance, nonhelical residues resonate in spectral regions distinct from

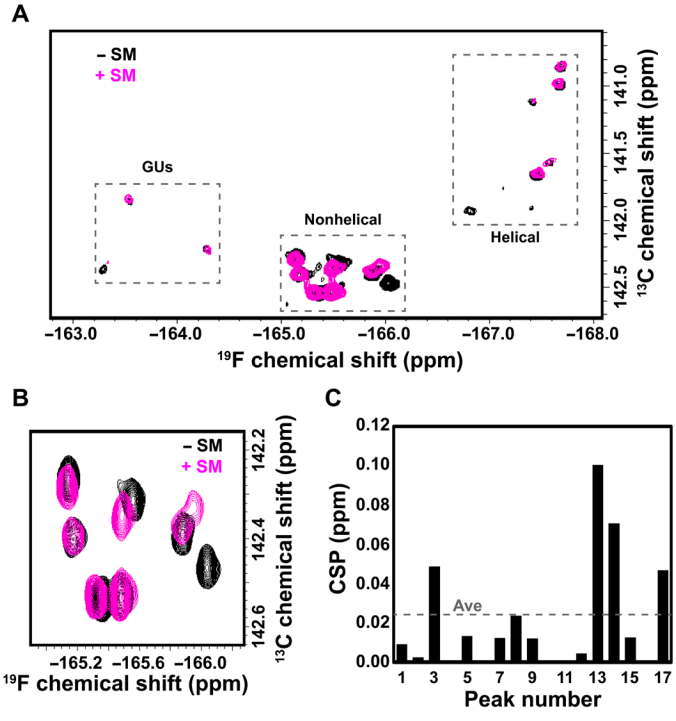

Fig. 6. Small-molecule binding to 5FU hHBV ε. (**A**) Overlay of ^{19}F-^{13}C-TROSY spectra for hHBV ε without (black) and with small molecule (SM, magenta). (**B**) Zoom-in of nonhelical residues showing chemical shift perturbations (CSPs) upon addition of SM. (**C**) Quantification of the CSPs upon addition of SM. The average (Ave) CSP is shown as a dashed line.

helical ones, which are further separated into GU wobble and AU Watson-Crick base-paired regions. The ability to differentiate between different structural features in an RNA simply based on chemical shifts removes the need for the time-consuming and laborious process of resonance assignment.

Given the ubiquity and functional importance of GU wobble base pairs (*40*) in all kingdoms of life (*41*), the ability to easily distinguish GU from canonical GC and AU base pairs has several important implications. For instance, in the minor groove, a GU base pair presents a distinctive exocyclic amino group that is unpaired and the U's C1′ atom rotates counterclockwise compared with the C's C1′ atom in a canonical GC base pair. This region serves as an important site for protein-RNA interactions. Similarly, in the major groove, G N7 and O6 together with U O4 create an area of intense negative electrostatic potential conducive for binding divalent metal ions. Furthermore, all canonical Watson-Crick base pairs are circumscribed by ~10.6-Å diameters formed by a line connecting their C1′-C1′ centers. These ribose-connected centers are superimposable with almost perfect alignment. In contrast, a GU base pair is misaligned counterclockwise by a residual twist of +14°, and an UG base pair is misaligned clockwise by a residual twist of −11° (*42*). That is, the GU base pair is not isosteric with canonical Watson-Crick pairs. Rather, these wobble base pairs either overtwist or undertwist the RNA double helix. ^{19}F-^{13}C labels might aid in elucidating the structural and dynamic basis of these twists depending on the identity of the base pairs neighboring the wobble pair. We, therefore, anticipate that our new label could potentially open up avenues for probing GU wobble pairs in various structural contexts outlined above, such as ^{19}F-^{13}C–labeled RNA-protein interactions and metalloribozyme-ion interactions.

In summary, the labeling technologies presented here open the door for characterizing the structure, dynamics, and interactions of RNA, RNA-RNA, RNA-DNA, RNA-protein, and RNA-drug complexes in vitro and in vivo for complexes as large as 100 kDa or higher with the appropriate fluorine NMR hardware. This ^{19}F -^{13}C labeling approach will also enable correlating chemical shift–structure relationships to aid chemical shift–centered probing of RNA structure, dynamics, and interactions. We envision that the ^{19}F-^{13}C spin pair, by providing a clear demarcation of RNA structural elements, may facilitate the discovery and identification of small drug-like molecules that target RNA binding pockets in vitro and in vivo.

MATERIALS AND METHODS

The full description of Materials and Methods can be found in the Supplementary Materials. A brief summary is provided here.

Chemoenzymatic synthesis of 5FUTPs

[5-^{19}F, 5-^{13}C, 6-^{2}H]– and [5-^{19}F, 5-^{13}C, 6-^{2}H, 1,3-^{15} N$_2$]–5FU were synthesized from unlabeled potassium cyanide, ^{13}C-labeled bromoacetic acid, and ^{15}N-labeled urea as described elsewhere (*3*, *17*, *18*, *24*). The resulting uracil was converted to 5FU by direct fluorination with Selectfluor and deuteration (*19*–*21*). [1′,5-^{13}C$_2$, 5-^{19}F, 6-^{2}H]–5FUTP and [1′,5-^{13}C$_2$, 5-^{19}F, 6-^{2}H, 1,3-^{15} N$_2$]–5FUTP were synthesized using PPP enzymes (*3*, *4*, *6*, *22*, *24*, *43*).

RNA in vitro transcription

All RNAs were prepared by in vitro transcription and purified as previously described (*3*, *4*). RNA concentrations were approximated by UV absorbance using extinction coefficients of 387.5 mM^{-1} cm^{-1} for HIV-2 TAR and 768.3 mM^{-1} cm^{-1} for hHBV ε. All RNA concentrations were >0.5 mM (~0.3 ml) in Shigemi NMR tubes.

Thermal melt analysis

We collected thermal melting profiles for both WT and 5FU-substituted HIV-2 TAR and hHBV ε as previously described (*24*, *25*).

Electronic structure calculations

Calculations were carried out on 1-methyl-uracil and 5-fluoro-1-methyl uracil using optimized geometries (*26*–*28*). All calculations used the Gaussian-16 program (*29*). Details are provided in the Supplementary Materials.

Solution NMR spectroscopy

All ^{19}F-^{13}C TROSY spectra were collected at 298 K using a Bruker 600 MHz Avance III spectrometer equipped with TXI (triple resonance inverse) and BBI (broad band inverse) probes. All data were processed with Bruker's Topspin 4.0.7 software. ^{1}H chemical shifts were internally referenced to DSS (0.00 ppm), with the ^{13}C chemical shifts referenced indirectly using the gyromagnetic ratios of ^{13}C/^{1}H (*44*). The ^{19}F chemical shifts were internally referenced to trifluoroacetic acid (−75.51 ppm) (*45*). Experiments showing each component of the ^{1}H/^{19}F-^{13}C correlations were adapted from a sensitivity- and gradient-enhanced ^{1}H-^{15}N TROSY used for proteins (*31*).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/41/eabc6572/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. K. D. Warner, C. E. Hajdin, K. M. Weeks, Principles for targeting RNA with drug-like small molecules. *Nat. Rev. Drug Discov.* **17**, 547–558 (2018).
2. A. Donlic, A. E. Hargrove, Targeting RNA in Mammalian Systems with Small Molecules. *WIREs RNA* **9**, e1477 (2018).
3. L. J. Alvarado, R. M. LeBlanc, A. P. Longhini, S. C. Keane, N. Jain, Z. F. Yildiz, B. S. Tolbert, V. M. D'Souza, M. F. Summers, C. Kreutz, T. K. Dayie, Regio-selective chemical-enzymatic synthesis of pyrimidine nucleotides facilitates RNA structure and dynamics studies. *Chembiochem* **15**, 1573–1577 (2014).
4. A. P. Longhini, R. M. LeBlanc, O. Becette, C. Salguero, C. H. Wunderlich, B. A. Johnson, V. M. D'Souza, C. Kreutz, T. K. Dayie, 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).
5. F. Rastinejad, C. Evilia, P. Lu, Studies of nucleic acids and their protein interactions by ^{19}F NMR. *Methods Enzymol.* **261**, 560–575 (1995).
6. M. Hennig, L. G. Scott, E. Sperling, W. Bermel, J. R. Williamson, Synthesis of 5-fluoropyrimidine nucleotides as sensitive NMR probes of RNA structure. *J. Am. Chem. Soc.* **129**, 14911–14921 (2007).
7. J. Horowitz, O. Ching-Nan, M. Ishaq, J. Ofengand, J. Bierbaum, Isolation and partial characterization of Escherichia coli valine transfer RNA with uridine and uridine-derived residues replaced by 5-fluorouridine. *J. Mol. Biol.* **88**, 301–312 (1974).
8. J. Horowitz, J. Ofendgan, W. Daniel Jr., M. Cohn, ^{19}F nuclear magnetic resonance of 5-fluorouridine-substituted tRNA$_1^{Val}$ from Escherichia coli. *J. Biol. Chem.* **252**, 4418–4420 (1977).
9. A. G. Marshall, J. L. Smith, Nuclear Spin-Labeled Nucleic Acids. 1. ^{19}F Nuclear Magnetic Resonance of Escherichia coli 5-Fluorouracil-5S-RNA. *J. Am. Chem. Soc.* **99**, 635–636 (1977).
10. S. L. Cobb, C. D. Murphy, ^{19}F NMR applications in chemical biology. *J. Fluor. Chem.* **130**, 132–143 (2009).
11. F. Guo, Q. Li, C. Zhou, Synthesis and biological applications of fluoro-modified nucleic acids. *Org. Biomol. Chem.* **15**, 9552–9565 (2017).
12. F. Sochor, R. Silvers, D. Müller, C. Richter, B. Fürtig, H. Schwalbe, ^{19}F -labeling of the adenine H2-site to study large RNAs by NMR spectroscopy. *J. Biomol. NMR* **64**, 63–74 (2016).

13. A. Boeszoermenyi, S. Chhabra, A. Dubey, D. L. Radeva, N. T. Burdzhiev, C. D. Chanev, O. I. Petrov, V. M. Gelev, M. Zhang, C. Anklin, H. Kovacs, G. Wagner, I. Kuprov, K. Takeuchi, H. Arthanari, Aromatic ^{19}F-^{13}C TROSY: a background-free approach to probe biomolecular structure, function, and dynamics. *Nat. Methods* **16**, 333–340 (2019).
14. M. Hennig, M. L. Munzarová, W. Bermel, L. G. Scott, V. Sklenář, J. R. Williamson, Measurement of Long-Range ^{1}H–^{19}F Scalar Coupling Constants and Their Glycosidic Torsion Dependence in 5-Fluoropyrimidine-Substituted RNA. *J. Am. Chem. Soc.* **128**, 5851–5858 (2006).
15. T. Knaus, M. Nassal, The encapsidation signal on the hepatitis B virus RNA pregenome forms a stem-loop structure that is critical for its function. *Nucleic Acids Res.* **21**, 3967–3975 (1993).
16. S. Flodell, M. Petersen, F. Girard, J. Zdunek, K. Kidd-Ljunggren, J. Schleucher, S. Wijmenga, Solution structure of the apical stem-loop of the human hepatitis B virus encapsidation signal. *Nucleic Acids Res.* **34**, 4449–4457 (2006).
17. J. Santalucia Jr., L. X. Shen, Z. Cai, H. Lewis, I. Tinoco Jr., Synthesis and NMR of RNA with selective isotopic enrichment in the bases. *Nucleic Acids Res.* **23**, 4913–4921 (1995).
18. C. H. Wunderlich, R. Spitzer, T. Santner, K. Fauster, M. Tollinger, C. Kreutz, Synthesis of (6-C-13)Pyrimidine Nucleotides as Spin-Labels for RNA Dynamics. *J. Am. Chem. Soc.* **134**, 7558–7569 (2012).
19. G. S. Lal, W. Pastore, R. Pesaresi, A Convenient Synthesis of 5-Fluoropyrimidines Using 1-(Chloromethyl)-4-fluoro-1,4-diazabicyclo[2.2.2]octane Bis(tetrafluoroborate)-SELECTFLUOR Reagent. *J. Organomet. Chem.* **60**, 7340–7342 (2005).
20. H. S. Rangwala, J. W. Giraldes, V. J. Gurvich, Synthesis and purification of [2-^{13}C]-5-fluorouracil. *J. Label. Compd. Radiopharm.* **54**, 340–343 (2011).
21. R. J. Cushley, S. R. Lipsky, Reactions of 5-fluorouracil derivatives with sodium deuteroxide. *Tetrahedron Lett.* **52**, 5393–5396 (1968).
22. L. J. Alvarado, A. P. Longhini, R. M. LeBlanc, B. Chen, C. Kreutz, T. K. Dayie, Chemo-enzymatic synthesis of selectively ^{13}C/^{15}N-labeled RNA for NMR structural and dynamics studies. *Methods Enzymol.* **549**, 133–162 (2014).
23. P. C. Bevilacqua, T. S. Brown, S. I. Nakano, R. Yajima, Catalytic Roles for Proton Transfer and Protonation in Ribozymes. *Biopolymers* **73**, 90–109 (2004).
24. L. G. Scott, M. Hennig, ^{19}F-site-specific-labeled nucleotides for nucleic acid structural analysis by NMR. *Methods Enzymol.* **566**, 59–87 (2016).
25. J. L. Mergny, L. Lacroix, UV melting of G-quadruplexes. *Curr. Protoc. Nucleic Acid Chem.* **37**, 17.1.1–17.1.15 (2009).
26. N. C. Handy, A. J. Cohen, Left-right correlation energy. *Mol. Phys.* **99**, 403–412 (2001).
27. C. Adamo, V. Barone, Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **110**, 6158–6170 (1999).
28. F. Jensen, Segmented contracted basis sets optimized for nuclear magnetic shielding. *J. Chem. Theory Comput.* **11**, 132–138 (2015).
29. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. Montgomery, J. A., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian 16, Revision A.03 (2016).
30. I. Kuprov, N. Wagner-Rundell, P. J. Hore, Bloch-Redfield-Wangsness theory engine implementation using symbolic processing software. *J. Magn. Reson.* **184**, 196–206 (2007).
31. J. Weigelt, Single scan, sensitivity- and gradient-enhanced TROSY for multidimensional NMR experiments. *J. Am. Chem. Soc.* **120**, 10778–10779 (1998).
32. K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12366–12371 (1997).
33. T. Carlomagno, I. Amata, J. R. Williamson, M. Hennig, NMR assignments of HIV-2 TAR RNA. *Biomol. NMR Assign.* **2**, 167–169 (2008).
34. W. Chu, J. Horowitz, ^{19}F NMR of 5-fluorouracil-substituted transfer RNA transcribed in vitro: resonance assignment of fluorouracil-guanine base pairs. *Nucleic Acids Res.* **17**, 7241–7252 (1989).
35. C. Dalvit, A. Vulpetti, Ligand-Based Fluorine NMR Screening: Principles and Applications in Drug Discovery Projects. *J. Med. Chem.* **62**, 2218–2244 (2019).
36. A. Divakaran, S. E. Kirberger, W. C. K. Pomerantz, SAR by (Protein-Observed) ^{19}F NMR. *Acc. Chem. Res.* **52**, 3407–3418 (2019).
37. U. S. Gill, P. T. F. Kennedy, The impact of currently licensed therapies on viral and immune responses in chronic hepatitis B: Considerations for future novel therapeutics. *J. Viral Hepat.* **26**, 4–15 (2019).
38. M. P. Williamson, Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1–16 (2013).
39. S. Barton, X. Heng, B. A. Johnson, M. F. Summers, Database proton NMR chemical shifts for RNA signal assignment and validation. *J. Biomol. NMR* **55**, 33–46 (2013).
40. F. H. Crick, Codon--anticodon pairing: the wobble hypothesis. *J. Mol. Biol.* **19**, 548–555 (1966).
41. G. Varani, W. H. McClain, The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Rep.* **1**, 18–23 (2000).
42. P. Ananth, G. Goldsmith, N. Yathindra, An innate twist between Crick's wobble and Watson-Crick base pairs. *RNA* **19**, 1038–1053 (2013).
43. P. K. Arthur, L. J. Alvarado, T. K. Dayie, Expression, purification and analysis of the activity of enzymes from the pentose phosphate pathway. *Protein Expr. Purif.* **76**, 229–237 (2011).
44. T. Aeschbacher, M. Schubert, F. H. T. Allain, A procedure to validate and correct the ^{13}C chemical shift calibration of RNA datasets. *J. Biomol. NMR* **52**, 179–190 (2012).
45. C. P. Rosenau, B. J. Jelier, A. D. Gossert, A. Togni, Exposing the Origins of Irreproducibility in Fluorine NMR Spectroscopy. *Angew. Chem. Int. Ed.* **57**, 9528–9533 (2018).
46. D. Rovnyak, D. P. Frueh, M. Sastry, Z. Y. J. Sun, A. S. Stern, J. C. Hoch, G. Wagner, Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction. *J. Magn. Reson.* **170**, 15–21 (2004).

Acknowledgments: We thank P. Deshong, J. Kahn, L.-X. Wang, and P. Y. Zavalij (University of Maryland) and H. Arthanari (Harvard University) for the helpful comments. We thank S. Bentz and D. Oh for help in preparing samples for thermal melt analysis, and M. Svirydava for help in analyzing samples by mass spectrometry. **Funding:** We thank the National Science Foundation (DBI1040158 to T.K.D. for NMR instrumentation) and the NIH (U54AI50470 to T.K.D. and D.A.C.) for support. **Author contributions:** T.K.D.: conceptualization. T.K.D. and O.B.B.: implementation of the project and manuscript preparation. G.Z., B.C., K.M.T., and T.K.D.: synthesis of 5FU. O.B.: synthesis of 5FUTP, RNA synthesis, and thermal melt analysis. T.K.D., K.M.T., B.C., and O.B.B.: TROSY measurements. O.B.B.: small-molecule titration. D.A.C.: DFT calculations. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors

Submitted 6 May 2020
Accepted 18 August 2020
Published 7 October 2020
10.1126/sciadv.abc6572

Citation: O. B. Becette, G. Zong, B. Chen, K. M. Taiwo, D. A. Case, T. K. Dayie, Solution NMR readily reveals distinct structural folds and interactions in doubly ^{13}C- and ^{19}F-labeled RNAs. *Sci. Adv.* **6**, eabc6572 (2020).

ScienceAdvances

Solution NMR readily reveals distinct structural folds and interactions in doubly ^{13}C- and ^{19}F-labeled RNAs

Owen B. Becette, Guanghui Zong, Bin Chen, Kehinde M. Taiwo, David A. Case and T. Kwaku Dayie

Sci Adv **6** (41), eabc6572.
DOI: 10.1126/sciadv.abc6572