The Stereospecific Assignment of H5′ and H5″ in RNA Using the Sign of Two-Bond Carbon-Proton Scalar Couplings

Jennifer V. Hines, Gabriele Varani,† Stacy M. Landry, and Ignacio Tinoco, Jr.*

Department of Chemistry and Laboratory of Chemical Biodynamics, University of California, Berkeley, California 94720

Received June 10, 1993

Stereospecific assignment of H5′ and H5″ protons in the NMR spectra of nucleic acids provides significant improvement in the use of NOE distance and torsion angle constraints for structure determination.[1] With stereospecific assignments, the torsion angles β (P5′–O5′–C5′–C4′) and γ (O5′–C5′–C4′–C3′) can be determined on the basis of ^1H–^1H and ^{31}P–^1H couplings.[2] In addition, stereospecific assignment allows use of NOE distance constraints for H5′ and H5″. Earlier approaches to stereospecific assignment have included chemical shifts,[1] NOEs combined with small homonuclear couplings,[3] a combination of $^3J_{HH}$, $^3J_{HP}$ and $^3J_{HC}$ couplings,[1,4] and stereoselective deuteration.[5] For the first method it is assumed that the H5′ proton has the more downfield chemical shift in A-form RNA.[1] However, in nonhelical structures and in the presence of ligands (e.g., proteins), this assumption is not necessarily true. The second method is unreliable because of spin diffusion and line-width interference at large molecular weight, as well as conformational flexibility that may skew the NOE intensities significantly. The third method relies on comparing the absolute magnitudes of small ^3J couplings (0–10 Hz) whose measurement is limited by the line widths of large RNA molecules. The fourth method requires the synthesis of an NMR sample specifically for the purpose of stereoassignment. We report a novel, independent method for the stereospecific assignment of the H5′ (*pro-S*) and H5″ (*pro-R*)[6] protons which is based only on the sign of the carbon–proton two-bond scalar couplings and will not be greatly limited by line width. The sign of $^2J_{CH}$ can also be used to determine torsion angle γ and the sugar conformation.

Three- and four-dimensional NMR are now quite common in protein studies.[7] With the ability to prepare isotopically-labeled RNA in quantities sufficient for NMR,[8] 3D and 4D NMR have been done on RNA.[9] All of the studies thus far have involved 100% ^{13}C-labeled RNA. We have prepared[10] a lower enrichment of 30% to study $^2J_{CH}$ and $^3J_{CH}$ scalar couplings without the interference of extra ^{13}C splittings due to neighboring ^{13}C atoms.[11] The **UUCG** RNA hairpin (Figure 1), whose structure has been determined in our laboratory,[15] contains both loop and helical regions. It therefore serves as a useful model system for developing new structure determination methods.

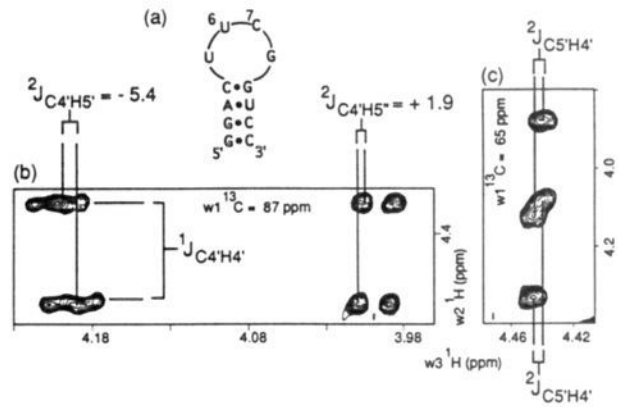


Figure 1. (a) **UUCG** RNA hairpin. (b) Expanded region of 3D HMQC-TOCSY[16] spectrum of **UUCG** at U6 C4′ carbon plane with H5′H4′ and H5″H4′ cross-peaks. (c) Expanded region at U6 C5′ carbon plane with H4′H5′ and H4′H5″ cross-peaks. The spectrum was acquired on a Bruker AMX-600 spectrometer at 26 °C with a 68-ms "clean" TOCSY[21] mixing sequence. ^1H width $\omega3 = \omega2 = 2000$ Hz and ^{13}C width $\omega1 = 6{,}250$ Hz. Obtained as $\omega3$ 512 (complex) × $\omega2$ 160 (TPPI) × $\omega1$ 40 (TPPI) with a relaxation delay of 2 s and a total acquisition time of 3 days. Data processed with FELIX (Biosym). Window functions applied: $\omega3 = \omega2$ = sine bell phase shifted by 30° and skewed by a factor of 0.7; $\omega1$ = sine bell phase shifted by 90° and skewed by 0.7. The spectrum was zero-filled to a final size of 1024 × 256 × 128 real points.

A 3D HMQC-TOCSY[16] was acquired without carbon decoupling. At the appropriate carbon plane (e.g., $\omega1$ = C4′ for U6, Figure 1), the $\omega3\omega2$ ^1H–^1H TOCSY cross-peak (e.g., $\omega3$ = H5′, $\omega2$ = H4′) is split along $\omega2$ by the large $^1J_{CH}$ scalar coupling constant (in this case, $^1J_{C4'H4'}$) and offset along $\omega3$ by either the $^2J_{CH}$ or $^3J_{CH}$ long-range scalar coupling constant (in this case $^2J_{C4'H5'}$). The E.COSY-type cross-peak arrangement, when split by the large $^1J_{CH}$, considerably reduces the interference of line width in the determination of small scalar coupling constants.[17] A similar E.COSY-type technique has been used to directly measure $^nJ_{HP}$ and $^nJ_{CP}$ couplings and indirectly measure (with curve-fitting) some $^nJ_{CH}$ in a DNA duplex at natural abundance.[4]

The 3D HMQC-TOCSY cross-peaks were readily assigned from the known 2D proton and carbon spectra of the **UUCG** hairpin.[15] The sign of the long-range carbon–proton scalar coupling can be determined by the direction of offset of the peak

† Present address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England.

(1) Varani, G.; Tinoco, I., Jr. *Q. Rev. Biophys.* **1991**, *24*, 479–532.

(2) Altona, C. *Recl. Trav. Chim. Pays-Bays* **1982**, *101*, 413–433.

(3) Blommers, M. J. J.; van de Ven, F. J. M.; van der Marel, G. A.; van Boom, J. H.; Hilbers, C. W. *Eur. J. Biochem.* **1991**, *201*, 33–51.

(4) Although it is not discussed by the authors, it appears that $^3J_{HP}$ and $^3J_{CP}$ were used to stereospecifically assign H5′/H5″ in Schmieder, P.; Ippel, J. H.; van den Elst, H.; van der Marel, G. A.; van Boom, J. H.; Altona, C.; Kessler, H. *Nucleic Acids Res.* **1992**, *20*, 4747–4751.

(5) Kline, P. C.; Serianni, A. S. *Magn. Reson. Chem.* **1988**, *26*, 120–123.

(6) The H5′/H5″ labels follow the terminology in ref 2. The prochiral descriptions of *pro-S* and *pro-R* have been added to further reduce ambiguity.

(7) (a) Bax, A.; Grzesiek, S. *Acc. Chem. Res.* **1993**, *26*, 131–138. (b) McIntosh, L. P.; Dahlquist, F. W. *Q. Rev. Biophys.* **1990**, *23*, 1–38. (c) Fesik, S. W.; Zuiderweg, E. R. *Q. Rev. Biophys.* **1990**, *23*, 97–131.

(8) (a) Batey, R. T.; Inada, M.; Kujawinski, E.; Puglisi, J. D.; Williamson, J. R. *Nucleic Acids Res.* **1992**, *20*, 4515–4523. (b) Nikonowicz, E. P.; Sirr, Q.; Legault, P.; Jucker, F. M.; Baer, L. M.; Pardi, A. *Nucleic Acids Res.* **1992**, *20*, 4507–4513. (c) Michnicka, M. J.; Harper, J. W.; King, G. C. *Biochemistry* **1993**, *32*, 395–400. (d) Polson, A. G.; Crain, P. F.; Pomerantz, S. C.; McCloskey, J. A.; Bass, B. L. *Biochemistry* **1991**, *30*, 11507–11514.

(9) (a) Nikonowicz, E. P.; Pardi, A. *Nature* **1992**, *355*, 184–186. (b) Nikonowicz, E. P.; Pardi, A. *J. Am. Chem. Soc.* **1992**, *114*, 1082–1083.

(10) Uniformly ^{13}C-labeled **UUCG** was prepared by isolating cellular RNA from *Methylobacterium extorquens* grown on 30% ^{13}C-labeled methanol as the sole carbon source. The cellular RNA was isolated by a hot phenol/sodium dodecyl sulfate extraction at pH 5.5[12] and then treated with RNase-free DNase I followed by nuclease P1. The nucleoside monophosphates were converted into the triphosphates by enzymatic phosphorylation,[13] and the RNA was prepared by T7 RNA transcription.[14]

(11) A similar approach has been used to study $^nJ_{CH}$ in proteins. (a) Edison, A. S.; Westler, W. M.; Markley, J. L. *J. Magn. Reson.* **1991**, *92*, 434–438. (b) Stockman, B. J.; Markley, J. L. In *Advances in Biophysical Chemistry*; Bush, C. A., Ed.; JAI Press: Greenwich, CT, 1990; Vol. 1 p 1.

(12) Oelmuller, U.; Kruger, N.; Steinbuchel, A.; Friedrich, C. G. *J. Microbiol. Methods* **1990**, *11*, 73–81.

(13) Simon, E. S.; Grabowski, S.; Whitesides, G. M. *J. Org. Chem.* **1990**, *55*, 1834–1841.

(14) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1987**, *15*, 8783–8789.

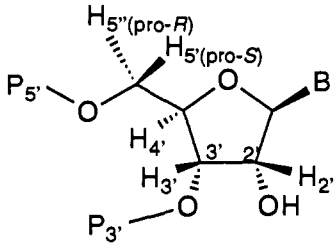
(15) (a) Varani, G.; Cheong, C.; Tinoco, I., Jr. *Biochemistry* **1991**, *30*, 3280–3289. (b) Varani, G.; Tinoco, I., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 9349–9354. (c) Cheong, C.; Varani, G.; Tinoco, I., Jr. *Nature* **1990**, *346*, 680–682.

(16) Wijmenga, S. S.; Hallenga, K.; Hilbers, C. W. *J. Magn. Reson.* **1989**, *84*, 634–642.

(17) Griesinger, C.; Sorensen, O. W.; Ernst, R. R. *J. Magn. Reson.* **1987**, *75*, 474–492.

0002-7863/93/1515-11002$04.00/0 © 1993 American Chemical Society

Table I. Numbering for a Nucleotide Unit, Predicted Signs (see text), and Some Observed Signs and Magnitudes (Hz) of Two-Bond Carbon–Proton Scalar Couplings of UUCG (Each cross-peak offset was measured three times and the values averaged. All values have an error within ±0.4 Hz.)



	predicted[a] sign for γ			observed coupling		
^2J	g^+	t	g^-	U6[b,c]	C7[b,d]	G8[b,e]
C4′H5′	–	+	–	–5.4	–5.5	+1.2
C4′H5″	+	–	–	+1.9	+2.0	–4.7
C5′H4′	+	–	–	+3.7	+4.0	–3.6

	predicted[a] sign for P		observed coupling			
^2J	2′-endo	3′-endo	U6[b]	C7[b]	U5[b]	C11[b]
C2′H3′	–	+	–2.0	–2.3	+2.6	+2.6
C3′H2′	+	–	+2.0	+2.3	–1.4	–1.7

[a] See text. [b] For UUCG[13c]:γ(O5′–C5′–C4′–C3′) for U6 = 63, C7 = 60, G8 = 177; sugar pucker P for U6,C7 = 2′-endo and U5,C11 = 3′-endo. [c] ^{1}H δ (ppm) H5′ = 4.19, H5″ = 4.01. [d] ^{1}H δ (ppm) H5′ = 2.67, H5″ = 3.57. [e] ^{1}H δ (ppm) H5′ = 4.15, H5″ = 4.38.

components (Figure 1).[17] The direction of offset corresponding to a positive coupling was established by examining ω3 = H3′, ω2 = H1′ cross-peaks at the C1′ carbon plane of U6 and C7 (data not shown), since $^3J_{C1'H3'}$ couplings are known to be positive.[18]

The sign of $^2J_{CH}$ is exceptionally useful for RNA structure determination. Previous work has established that $^2J_{CH}$ in the system H–C–^{13}C–X, where X is an electronegative atom, depends on the torsion angle θ between H and X.[19] For ethanol, $^2J_{CH}$ is negative when $0° \le \theta < 90°$ and positive when $90° < \theta \le 180°$, with maximum absolute values at 0° and 180°.[19d] At θ = 90°, $^2J_{CH} = 0$. All of the $^2J_{CH}$ couplings which could be observed for the **UUCG** hairpin fit this trend. Some representative $^2J_{CH}$ are shown in Table I. It should be noted that the sign alone of $^2J_{C3'H2'}$ and $^2J_{C2'H3'}$ can be used to distinguish 2′-endo and 3′-endo sugar puckers. The signs of $^2J_{C4'H5'}$, $^2J_{C4'H5''}$, and $^2J_{C5'H4'}$ provide an independent means for the stereospecific assignment of H5′ and H5″ and determination of torsion angle γ. This qualitative information will prove useful in the structure determination of large RNA molecules where line width and overlap may interfere with the quantitative determination of $^nJ_{HH}$, $^nJ_{CH}$, $^nJ_{PH}$, and $^nJ_{CP}$ couplings.

For U6 and C7, the angle γ was constrained to g^+ on the basis of small $^3J_{H4'H5'/H5''}$ scalar couplings without a stereospecific assignment.[15] The positive $^2J_{C5'H4'}$ confirms this arrangement for γ. The stereospecific assignment of H5′ and H5″ is then made by correlating the observed signs of $^2J_{C4'H5'}$ and $^2J_{C4'H5''}$ with the expected signs (derived by examining a Newman projection about the C5′–C4′ bond and using the relationship between θ and the sign of $^2J_{CH}$ as discussed above) for $\gamma = g^+$. For U6, the assignment of the more downfield proton to H5′ agrees with the expected chemical shift for a nucleotide in a double helix. For C7, however, the more downfield proton is H5″. Examination of the **UUCG** structure shows that C7 H5′ is closer and more over the ring of the *syn*-**G8** than H5″ and therefore would experience a larger upfield shift due to ring current shielding.

Originally, γ of G8 could not be determined with $^3J_{H4'H5'/H5''}$ scalar couplings due to overlap of H4′ with the more downfield H5′/H5″ proton and the need to assign H5′ and H5″ stereospecifically. However, all calculated structures showed $\gamma = t$,[15] presumably as a result of other less direct constraints. The signs of the $^2J_{CH}$ couplings confirm this result (Table I). A negative $^2J_{C5'H4'}$ could correspond to either $\gamma = g^-$ or t, but having one positive $^2J_{C4'H5'/H5''}$ and one negative confines γ to trans and, at the same time, stereospecifically assigns H5′ and H5″. Again, the more downfield proton is H5″.

The coherence transfer from H4′ to H5′/H5″ is weak in a TOCSY due to small $^nJ_{HH}$ couplings in A-form helices. This, in addition to low isotopic enrichment and the cross-peaks being split by the $^1J_{CH}$, made it difficult to observe H4′–H5′/H5″ cross-peaks in the more rigid stem region. However, in the loop region, precisely where a stereospecific assignment would be most useful, the cross-peaks were readily observable. We have observed the same correlation of the sign of the $^2J_{CH}$ with the stereospecific assignment of H5′/H5″ in a 3D HMQC-NOESY[20] with no carbon decoupling (data not shown). E.COSY-type cross-peaks between H4′ and H5′/H5″ are readily observable because the magnetization transfer is very efficient due to the short H4′ to H5′/H5″ distance.

In summary, determination of the sign of the two-bond carbon–proton couplings provides a useful new method for the structure determination of RNA molecules. The ribose sugar conformation and torsion angle γ (O5′–C5′–C4′–C3′) can be specified. Furthermore, H5′ and H5″ protons can be stereospecifically assigned, allowing their use in NOE distance constraints and in determining β (P5′–O5′–C5′–C4′). The correlation we have reported between the sign of $^2J_{CH}$ and RNA structure will be particularly useful in the study of large RNAs and RNA–protein complexes.

Acknowledgment. Supported by NIH GM 10840, DOE DE-FG03-86ER60406, and instrumentation Grants DOE DE-FG05-86ER75281, NSF DMB 86-09305, and NSF BBS 86-20134. J.V.H. is supported in part by NIEHS postdoctoral training Grant 2T32 ES 07075. We would like to thank the Stable Isotope Resource at Los Alamos National Laboratory for supplying us with the ^{13}C-labeled bacteria.

(18) Kline, P. C.; Serianni, A. S. *J. Am. Chem. Soc.* **1990**, *112*, 7373–7381.

(19) For review, see: (a) Marshall, J. L. *Carbon–Carbon and Carbon–Proton NMR Couplings*; Verlag Chemie International, Inc.: Deerfield Beach, FL, 1983. (b) Hansen, P. E. *Prog. Nucl. Magn. Reson. Spectrosc.* **1981**, *14*, 175–296. For furanosides, see: (c) Cyr, N.; Perlin, A. S. *Can. J. Chem.* **1979**, *57*, 2504–2511. (d) Bock, K.; Pedersen, C. *Acta Chem. Scand.* **1977**, *B31*, 354–358. For calculations, see: (e) Schwarcz, J. A.; Cyr, N.; Perlin, A. S. *Can. J. Chem.* **1975**, *53*, 1872–1875.

(20) Fesik, S. W.; Zuiderweg, E. R. P. *J. Magn. Reson.* **1988**, *78*, 588–593.

(21) Griesinger, C.; Otting, G.; Wüthrich, K.; Ernst, R. R. *J. Am. Chem. Soc.* **1988**, *110*, 7870–7872.