NMR Determination of Sugar Puckers in Nucleic Acids from CSA−Dipolar Cross-Correlated Relaxation

Jérôme Boisbouvier,[‡] Bernhard Brutscher,[‡] Arthur Pardi,[§] Dominique Marion,[‡] and Jean-Pierre Simorre*,[‡]

Institut de Biologie Structurale, Jean-Pierre Ebel C.N.R.S.-C.E.A. 41, rue Jules Horowitz, 38027 Grenoble Cedex, France
Department of Chemistry and Biochemistry University of Colorado, Boulder, Colorado 80309-0215
Received March 20, 2000
Revised Manuscript Received May 19, 2000

In structural NMR studies of nucleic acids, the local conformations of the phosphodiester backbone and the sugars are generally poorly defined by the available ^{1}H−^{1}H nOe distance constraints. Scalar ^3J coupling constants, which provide a source of torsion angle restraints[1] are difficult to measure accurately for larger oligonucleotides due to increased line widths and relaxation-induced line splittings.[2] It is therefore important to investigate new types of NMR information as additional structural constraints. For example, dipole−dipole cross-correlated relaxation has been shown to give information on the angle between two internuclear vectors,[3] an effect recently exploited for determining sugar puckers in RNA.[4] There has also been a growing interest in chemical shielding anisotropies (CSA) which are sensitive to local molecular geometry.[5] In liquid-state NMR, CSA information can be obtained from NMR relaxation data.[6] In this contribution we present the measurement of cross-correlated relaxation rates involving ^{13}C CSA and ^{13}C−^{1}H dipolar interactions in the sugar moieties of nucleic acids. These rates are sensitive to variation in the amplitude and orientation of the ^{13}C CSA tensor, which depend on the sugar pucker as shown in density functional theory (DFT) calculations on model systems.[7] Experimentally we find a strong correlation between the relaxation rate constants measured for $C_{1'}$ and $C_{3'}$ carbons and the sugar puckers. The good qualitative agreement with DFT calculations of ^{13}C CSA indicates that the measurement of CSA−dipolar cross-correlated relaxation rates offers an attractive new way of determining sugar puckers for larger nucleic acids.

Cross correlation between the ^{13}C CSA and ^{13}C−^{1}H dipolar interactions in scalar coupled ^{13}C−^{1}H spin pairs results in differential transverse relaxation of the two ^{13}C doublet lines.[8] The CSA−dipolar cross-correlated relaxation rate constant $\Gamma^{CSA,DD}_{C,C-H}$ can be accurately measured from a constant time (CT) TROSY experiment[9] as shown in Figure S1 of the Supporting Information. Two ^{1}H−^{13}C correlation spectra are recorded with different phase settings to probe the relaxation of the C^+H^α and C^+H^β lines, respectively. The rate constant $\Gamma^{CSA,DD}_{C,C-H}$ is obtained from the peak intensities measured in the two spectra as $\Gamma^{CSA,DD}_{C,C-H} = \ln(I(C^+H^\alpha)/I(C^+H^\beta))/2T$, with T the length of the CT evolution period. In the simple case of isotropic molecular tumbling and isotropic fast time-scale intramolecular motion, $\Gamma^{CSA,DD}_{C,C-H}$ provides a measure of the CSA-related parameter $\Delta\sigma^*$:

$$\Gamma^{CSA,DD}_{C,C-H} \propto B_0 S^2 \tau_c \Delta\sigma^* \quad (1)$$

with $\Delta\sigma^* = \sum_{j=1,2}(\sigma_{jj} - \sigma_{33})[(3\cos^2(\theta^{jd}) - 1)/2]$.[10] The additional dependence on the molecular tumbling correlation time, τ_c the generalized order parameter, S^2 and the magnetic field strength B_0, can be removed by calculating the ratio $\Gamma^{CSA,DD}_{C_k,C_k-H_k}/\Gamma^{CSA,DD}_{C_l,C_l-H_l}$ measured for two ^{13}C−^{1}H spin pairs within the same sugar, provided that the order parameters S^2_k and S^2_l are approximately the same:

$$\lambda = \frac{\Gamma^{CSA,DD}_{C_k,C_k-H_k}}{\Gamma^{CSA,DD}_{C_l,C_l-H_l}} \cong \frac{\Delta\sigma^*(C_k)}{\Delta\sigma^*(C_l)} \quad (2)$$

NMR experiments using the pulse sequence of Figure S1 have been performed on three uniformly ^{13}C/^{15}N-labeled RNA samples: a 12-mer hairpin containing a CUUG tetraloop,[11] a 30-mer leadzyme,[12] and a 33-mer theophylline-binding RNA aptamer.[13] A table of measured relaxation rate constants $\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ and $\Gamma^{CSA,DD}_{C_3',C_3'-H_3'}$ is provided as Supporting Information. Both of these relaxation rates were measured for 32 nucleotides and a histogram of the computed $\Gamma^{CSA,DD}_{C_3',C_3'-H_3'}/\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ ratios is shown in Figure 1B. To help interpret the conformation of the individual sugars, the measured $\Gamma^{CSA,DD}_{C_3',C_3'-H_3'}/\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ ratios were compared to the $\Delta\sigma^*(C_{3'})/\Delta\sigma^*(C_{1'})$ values calculated for Me-ribofuranose[7] shown in Figure 1A. From eq 2, large values (∼2.5) are predicted for sugars in $C_{3'}$-*endo* conformation, whereas smaller ones (∼0.5) are predicted for sugars in $C_{2'}$-*endo* conformation.

In the presence of anisotropic molecular tumbling, neglected in eqs 1 and 2, the cross-correlated relaxation rates exhibit a complex dependence on the relative orientation of the CSA and dipolar tensors with respect to the molecular diffusion axes.[14] The asymmetry of the diffusion tensor can be estimated from the mass distribution in the RNA structure, which for an axially symmetric model yields $D_\parallel/D_\perp \approx 1.3$, 2.0, and 2.5 for the three RNA molecules. For a complete sampling of sugar orientations with respect to the diffusion tensor ($D_\parallel/D_\perp = 2.5$), the calculated $\Gamma^{CSA,DD}_{C_3',C_3'-H_3'}/\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ ratios vary between 2.0 and 3.3 for $C_{3'}$-*endo*, and between 0.1 and 0.7 for $C_{2'}$-*endo* conformations as shown in Figure S3 of the Supporting Information. While the anisotropy of the molecular tumbling induces some scattering in the measured $\Gamma^{CSA,DD}_{C_3',C_3'-H_3'}/\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ ratios (Figure 1B), the effect remains much smaller than the variations induced by changing the sugar conformation. On the basis of the DFT results, and taking into

* Author for correspondence. E-mail: jps@rmn.ibs.fr.
‡ Institut de Biologie Structurale.
§ University of Colorado.

(1) Pardi, A. *Methods in Enzymology*, 261*;* Academic Press: San-Diego, 1995; pp 350−382.
(2) (a) Harbison, G. S. *J. Am. Chem. Soc.* **1993**, *115*, 3026−3027. (b) Brüschweiler, R. *Chem. Phys. Lett.* **1996**, *257*, 119−122.
(3) Reif, B.; Hennig, M.; Griesinger, C. *Science* **1997**, *276*, 1230−1233.
(4) (a) Felli, I.; Richter, C.; Griesinger, C.; Schwalbe, H. *J. Am. Chem. Soc.* **1999**, *121*, 1956−1957. (b) Richter, C.; Griesinger, C.; Felli, I.; Cole, P. T.; Varani, G.; Schwalbe, H. *J. Biomol. NMR* **1999**, *15*, 241−250.
(5) (a) Halvin, R. H.; Le, H.; Laws, D. D.; deDios, A. C.; Olfield, E. *J. Am. Chem. Soc.* **1997**, *119*, 11951−11958. (b) deDios, A. C. *Prog. Nucl. Magn. Reson. Spectrosc.* **1996**, *29*, 229−278. (c) Sitkoff, D.; Case, D. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 165−190.
(6) (a) Tjandra, N., Szabo, A., Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 6986−6991. (b) Tjandra, N.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 9576−9577. (c) Tessari, M.; Vis, H.; Boelens, R.; Kaptein, R.; Vuister, G. *J. Am. Chem. Soc.* **1997**, *119*, 8985−8990. (d) Fushman, D.; Cowburn, D. *J. Am. Chem. Soc.* **1998**, *120*, 7109−7110. (e) Kroenke, C. D.; Rance, M.; Palmer A., III. *J. Am. Chem. Soc.* **1999**, *121*, 10119−10125. (f) Damberg, P.; Jarvet, J.; Gräslund, A. *J. Biomol. NMR* **1999**, *15*, 27−37.
(7) Dejaegere, A. P.; Case, D. A. *J. Phys. Chem. A* **1998**, *102*, 5280−5289.
(8) (a) Guéron, M.; Leroy, J. L.; Griffey, R. H. *J. Am. Chem. Soc.* **1983**, *105*, 7262−7266. (b) Goldman, M. *J. Magn. Reson.* **1984**, *60*, 437−452.
(9) (a) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *J. Am. Chem. Soc.* **1998**, *120*, 6394−6400. (b) Brutscher B.; Boisbouvier, J.; Pardi, A.; Marion, D.; Simorre, J.-P. *J. Am. Chem. Soc.* **1998**, *120*, 11845−11851.
(10) σ_{jj} (j = 1, 2, 3) are the principal axes values of the ^{13}C CSA.; θ^{jd} is the angle between the principal axis j and the C−H bond vector.
(11) Jucker, F. M.; Pardi, A. *Biochemistry* **1995**, *34*, 14416−14427.
(12) Hoogstraten, C. G.; Legault, P.; Pardi, A. *J. Mol. Biol.* **1998**, *284*, 337−350.
(13) (a) Zimmermann, G. R.; Jenison, R. D.; Wick, C. L., Simorre, J.-P.; Pardi, A.; *Nat. Struct. Biol.* **1997**, *4*, 644−649. (b) Zimmermann, G. R.; Shields, T. P.; Jenison, R. D.; Wick, C. L., Pardi, A. *Biochemistry* **1998**, *37*, 9186−9192. (c) Boisbouvier, J.; Brutscher, B.; Simorre, J.-P.; Marion, D. *J. Biomol. NMR* **1999**, *14*, 241−252.
(14) Cuperlovic, M., Palke, W. E., Gerig, J. T., Gray, G. A. *J. Magn. Reson., Ser. B* **1996**, *110*, 26−38.

10.1021/ja000976b CCC: $19.00
Published on Web 06/30/2000

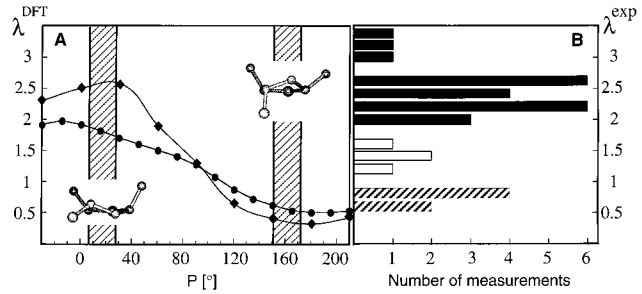

Figure 1. (A) DFT calculations [7] of $\lambda^{DFT} = \Delta\sigma^*(C_{3'})/\Delta\sigma^*(C_{1'})$ on two different model compounds: Me-ribofuranose (filled diamonds) and deoxythymidine (filled circles) as a function of the sugar pucker angle P. The hashed regions correspond to the allowed pseudorotation phase angles P for the two main sugar conformations $C_{2'}$-*endo* (P = 162° ± 10°) and $C_{3'}$-*endo* (P = 18° ± 10°).[16] (B) Histogram of measured $\lambda^{exp} = \Gamma^{CSA,DD}_{C_3',C_3'-H_3}/\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ ratios for three different RNA molecules. Sugars assigned to $C_{3'}$-*endo* ($\Gamma^{CSA,DD}_{C_3',C_3'-H_3}/\Gamma^{CSA,DD}_{C_1',C_1'-H_1'} > 2.0$) and $C_{2'}$-*endo* ($\Gamma^{CSA,DD}_{C_3',C_3'-H_3}/\Gamma^{CSA,DD}_{C_1',C_1'-H_1'} < 1.0$) are indicated by black and hashed bars, respectively. Open bars correspond to sugars (G1 and A8 of the theophylline-binding RNA; C27 and C33 of the leadzyme) undergoing conformational averaging.

account the scattering induced by anisotropic molecular diffusion, the majority of sugars in Figure 1B were assigned to $C_{3'}$-*endo* conformations (black bars), whereas six sugars were assigned to a $C_{2'}$-*endo* conformation (hashed bars). The $\Gamma^{CSA,DD}_{C_3',C_3'-H_3}/\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ ratios for the remaining four nucleotides correspond to neither a $C_{3'}$-*endo* nor a $C_{2'}$-*endo* conformation suggesting conformational averaging in these sugars (open bars).

To help validate these results the CSA−dipolar cross-correlated relaxation rates are compared to previous structural NMR studies based on ^{1}H−^{1}H nOe and $^3J_{HH}$ scalar coupling constants. The sugar puckers determined from CSA−dipolar cross-correlated relaxation experiments here agree with sugar puckers previously reported for these three RNA structures, except for nucleotides G1 and G9 in the leadzyme, and C33 in the theophylline-binding RNA. A further independent source of sugar pucker information is obtained from ^{1}H−^{13}C dipole−dipole cross-correlated relaxation rates. We therefore measured $\Gamma^{DD,DD}_{C_1'-H_1',C_2'-H_2'}$ rate constants in the leadzyme and in the RNA−theophylline complex using a HCCH-type transfer experiment as proposed recently by Schwalbe and co-workers.[4] A correlation plot of the measured $\Gamma^{DD,DD}_{C_1'-H_1',C_2'-H_2'}$ and $\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ rate constants for the RNA−theophylline complex is shown in Figure 2. The clustering of the measured rate constants in two distinct regions of the correlation plot confirms that both methods are equally valid for determining sugar puckers in RNA. The slightly reduced $\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ and $\Gamma^{DD,DD}_{C_1'-H_1',C_2'-H_2'}$ rate constants measured for G14 located in the GA_3 loop of the RNA are most likely due to increased fast time-scale internal mobility. This contribution can be partly removed by calculating the ratio $\Gamma^{CSA,DD}_{C_3',C_3'-H_3}/\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ (see eq 2) which shows that G14 is predominantly in a $C_{3'}$-*endo* conformation. For the leadzyme, the measured $\Gamma^{DD,DD}_{C_1'-H_1',C_2'-H_2'}$ rates also confirm the sugar puckers determined from CSA−dipolar relaxation data (Supporting Information). In particular the terminal nucleotide G1 was constrained to $C_{3'}$-*endo* in the NMR structure, but both the measured relaxation rates, $\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ and $\Gamma^{DD,DD}_{C_1'-H_1',C_2'-H_2'}$ indicate conformational averaging in this flexible region of the RNA. For nucleotide G9 located in the active-site internal loop of the leadzyme, the $C_{2'}$-*endo* conformation proposed in the NMR structure is not confirmed by the cross-correlated relaxation data, which instead indicates predominantly a $C_{3'}$-*endo* conformation.

CSA−dipolar cross-correlated relaxation rates increase linearly with the molecular tumbling correlation time τ_c (eq 1), and accurate measurements are still possible for much larger molecules than those studied here. The method will further benefit from higher magnetic field strengths (see eq 1), and we therefore expect

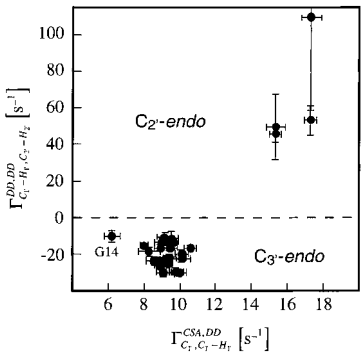

Figure 2. Correlation plot of $\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ and $\Gamma^{DD,DD}_{C_1'-H_1',C_2'-H_2'}$ rate constants measured for the RNA−theophylline complex at a magnetic field strength B_0 = 18.8T. Small $\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ values correspond to small *negative* $\Gamma^{DD,DD}_{C_1'-H_1',C_2'-H_2'}$ values and indicate a $C_{3'}$-*endo* conformation, whereas large $\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ values are accompanied by large *positive* $\Gamma^{DD,DD}_{C_1'-H_1',C_2'-H_2'}$ values and indicate a $C_{2'}$-*endo* conformation. Data for the two terminal nucleotides (G1 and C33) and for C27 are not included because NMR relaxation measurements have shown these residues highly flexible in solution.[13b,c] Error bars were calculated from the signal-to-noise ratio measured in the corresponding spectra.

that it will be advantageous for larger RNAs studied at higher magnetic fields. Use of a *non*-CT TROSY experiment using homonuclear ^{13}C decoupling, combined with a line shape analysis,[15] will further increase the sensitivity of the CSA−dipolar cross-correlated relaxation measurement and thus extend the applicability of this approach to larger molecules.

In conclusion, we have shown that the ^{13}C CSA of $C_{1'}$ and $C_{3'}$ sugar carbons are useful probes of local structure in RNA as was previously predicted by DFT calculations on small model compounds.[7] The variations in the ^{13}C CSA are manifested here in CSA−dipolar cross-correlated relaxation rates $\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ and $\Gamma^{CSA,DD}_{C_3',C_3'-H_3'}$, which can be obtained with high accuracy from the measurement of peak intensities in 2D ^{1}H−^{13}C correlation spectra. For sugars which are known to undergo only restricted fast time-scale internal mobility, measurement of a single rate constant proves to be sufficient (see Figure 2) for distinguishing between the main sugar conformations, whereas for sugars located in flexible loop regions of the RNA the ratio of two rates provides structural information which to a good approximation is independent of the molecular motion (Eq. [2]). This approach should also be applicable to DNA as DFT calculations on deoxythymidine show a similar dependence of the ^{13}C CSAs on the sugar pucker (Figure 1A). The combination of quantum chemical investigations of structure-dependent variations of CSA tensors, and NMR measurements of CSA−dipolar cross-correlated relaxation rates offers new perspectives for structural studies of nucleic acids.

Acknowledgment. We thank G. Zimmermann, A. Brown, P. Legault, and F. Jucker for the preparation of the ^{13}C/^{15}N-labeled RNA samples, M. Rance for recording the CT TROSY experiment on the CUUG hairpin loop, A. Dejaegere for providing DFT data on ^{13}C CSA, and M. Blackledge for stimulating discussions. This work was supported by the C.E.A. and the C.N.R.S. (France), M.S.I. (San Diego, CA), and NIH AI33098 and a NSF International Travel Award (INT-9602955).

Supporting Information Available: Pulse sequences, representative spectra, and experimental details for the measurement of $\Gamma^{CSA,DD}_{C,C-H}$ and $\Gamma^{DD,DD}_{C_1'-H_1',C_2'-H_2'}$, three tables of rate constants measured for the hairpin loop, the leadzyme, and the theophylline-binding RNA aptamer, and a histogram of calculated $\Gamma^{CSA,DD}_{C_3',C_3'-H_3'}/\Gamma^{CSA,DD}_{C_1',C_1'-H_1'}$ ratios, assuming an axially symmetric rotational diffusion tensor (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA000976B

(15) (a) Schwalbe, H.; Samstag, W.; Engels, J. W.; Bermel, W.; Griesinger, C. *J. Biomol. NMR* **1993**, *3*, 479−486. (b) Reif, B., Steinhagen, H.; Junker, B.; Reggelin, M.; Griesinger, C. *Angew. Chem.* **1998**, *37*, 1903−1906.

(16) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1983,