Supporting Information

Resolution-optimized NMR measurement of $^{1}D_{CH}$, $^{1}D_{CC}$ and $^{2}D_{CH}$ residual dipolar couplings in nucleic acid bases

Jérôme Boisbouvier[a, b], David L. Bryce[a], Erin O'Neil-Cabello[a], Edward P. Nikonowicz,[c] and Ad Bax[a]

[a] Laboratory of Chemical Physics
National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health
Bethesda, Maryland 20892-0520

[b] present address: Laboratoire de RMN
Institut de Biologie Structurale – Jean-Pierre Ebel
UMR 5075 CNRS-CEA-UJF
41 rue Jules Horowitz
38027 GRENOBLE Cedex 1 – FRANCE

[c] Department of Biochemistry and Cell Biology,
Rice University,
POB 1892,
Houston, TX 77251

Supporting Information Table S1. Experimental $^{1}J_{CH}$, $^{1}J_{CC}$ and $^{2}J_{CH}$ spin-spin coupling constants in the 24-nucleotide RNA sample.[a]

Uridine (5 bases)

#	C_5C_6	C_5C_4	C_4H_5	C_6H_5	C_5H_6	C_5H_5	C_6H_6
41	66.1±0.1	64.7±0.1	1.4±0.1	4.0±0.1	1.9±0.1	176.4±0.1	180.2±0.2
44	66.2±0.1	64.5±0.1	1.4±0.1	4.4±0.1	2.4±0.2	177.1±0.1	179.8±0.3
47	66.5±0.1	64.5±0.1	1.4±0.1	4.1±0.1	2.3±0.1	176.1±0.1	180.7±0.1
54	65.9±0.1	64.1±0.1	1.3±0.1	4.3±0.1	2.5±0.1	176.3±0.1	180.3±0.2
55	65.8±0.1	64.1±0.1	1.3±0.1	4.4±0.1	2.4±0.1	176.4±0.1	179.9±0.2

PseudoUridine (1 base)

#	C_5C_6	C_5C_4	C_4H_5	C_6H_5	C_5H_6	C_5H_5	C_6H_6
46					0.9±0.1		178.6±0.3

Cytosine (4 bases)

#	C_5C_6	C_5C_4	C_4H_5	C_6H_5	C_5H_6	C_5H_5	C_6H_6
40	67.2±0.1	54.7±0.1	1.8±0.1	4.3±0.1	2.8±0.1	174.0±0.1	179.2±0.2
58	67.3±0.1	54.5±0.1	1.4±0.1	4.3±0.1	3.0±0.1	173.6±0.1	178.9±0.2
59	67.3±0.1	54.5±0.1	1.7±0.1	4.2±0.1	3.2±0.1	173.6±0.1	179.4±0.2
60	67.6±0.1	54.8±0.1	1.7±0.1	4.4±0.1	2.9±0.1	173.5±0.1	179.6±0.2

Adenine (7 bases)[c]

	C_5C_4	C_5C_6	C_8H_8 [b]	C_2H_2 [b]
42	65.2±0.1	74.6±0.1	214.0±0.1	200.0±0.2
43	64.4±0.1	74.4±0.1	214.1±0.2	200.2±0.2
49	64.5±0.1	74.6±0.1	215.1±0.2	200.2±0.1
50	64.7±0.1	75.4±0.1	215.8±0.2	200.9±0.1
51	65.3±0.1	75.0±0.2	214.0±0.2	200.7±0.1
53	65.4±0.1	75.0±0.1	214.9±0.3	199.9±0.2
56	65.8±0.1	75.0±0.1	214.0±0.1	199.5±0.2

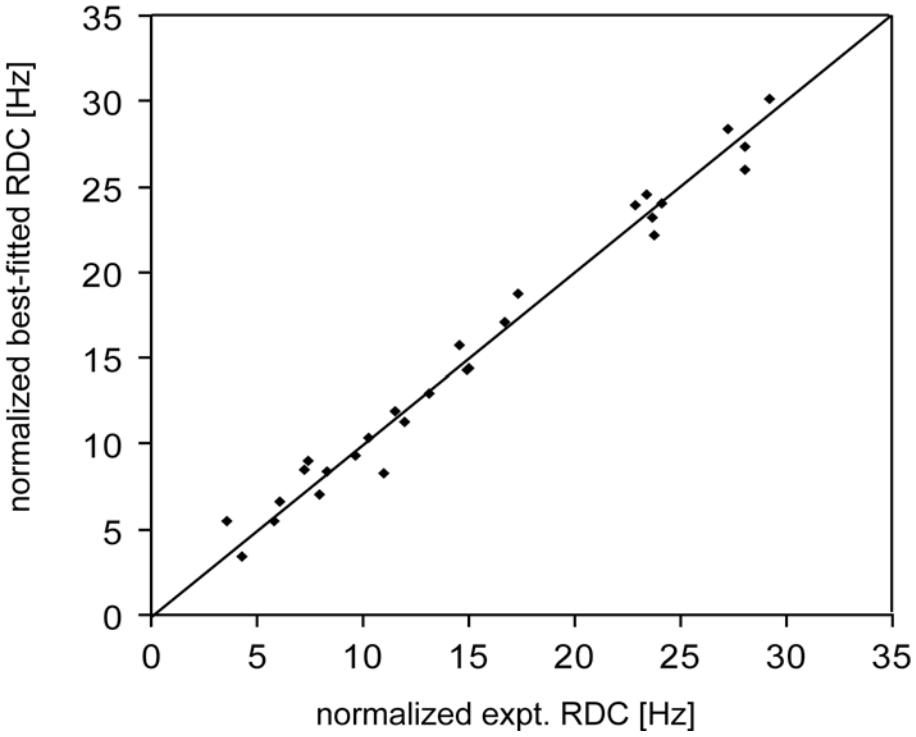
Guanine (3 bases)[d]

	C_5C_4	C_5C_6	C_8H_8 [b]
39	64.0±0.2	85.8±0.3	214.7±0.3
48	63.4±0.1	NA	215.8±0.3
57	63.4±0.1	85.8±0.1	214.3±0.2

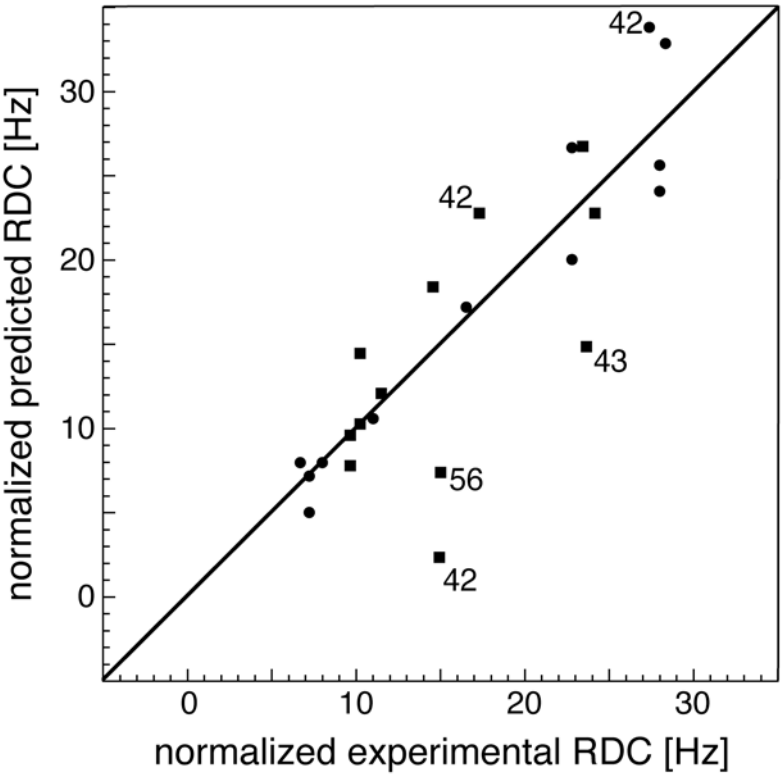
[a] For each base, the experimental value and the standard deviation of the measured J splitting, the experiment used for the measurement, and the corresponding experimental precision are given. Pyrimidine couplings have been measured at $B_0 = 17.6$ T, and purine data at $B_0 = 14.1$ T; no correction for magnetic field induced alignment is included.

[b] measured with an IPAP-HSQC experiment (Ottiger et al., 1998).

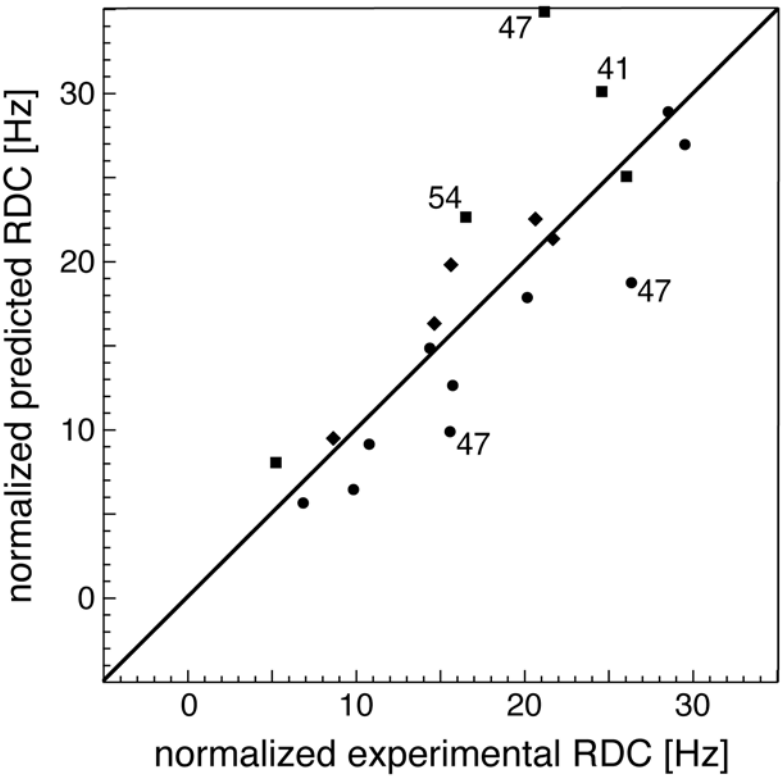
[c] No precise measurements could be made for A52 due to conformational exchange broadening and partial overlap.
[d] Outside the stem region, G-H_8 resonances had largely exchanged with solvent deuterons and were vanishingly weak.



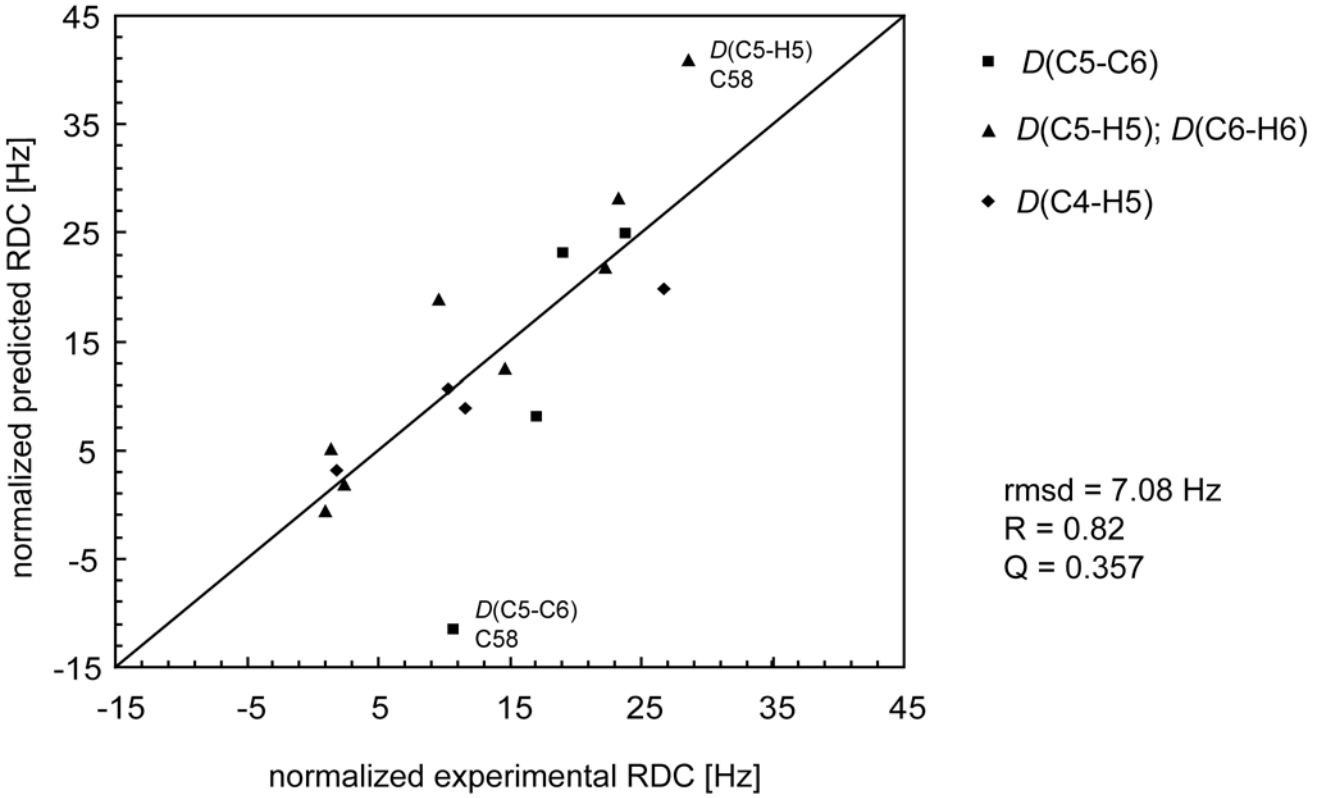
Supporting Information Figure S1. Correlation between experimental and predicted RDC data for adenines 42, 43, 49, 50, 51, 53, and 56. The fit of four dipolar couplings to each adenine involves three adjustable parameters, causing the Pearson's correlation coefficient, R_P = 0.990, to be artificially inflated. RDCs are normalized to the one-bond C8-H8 interaction.



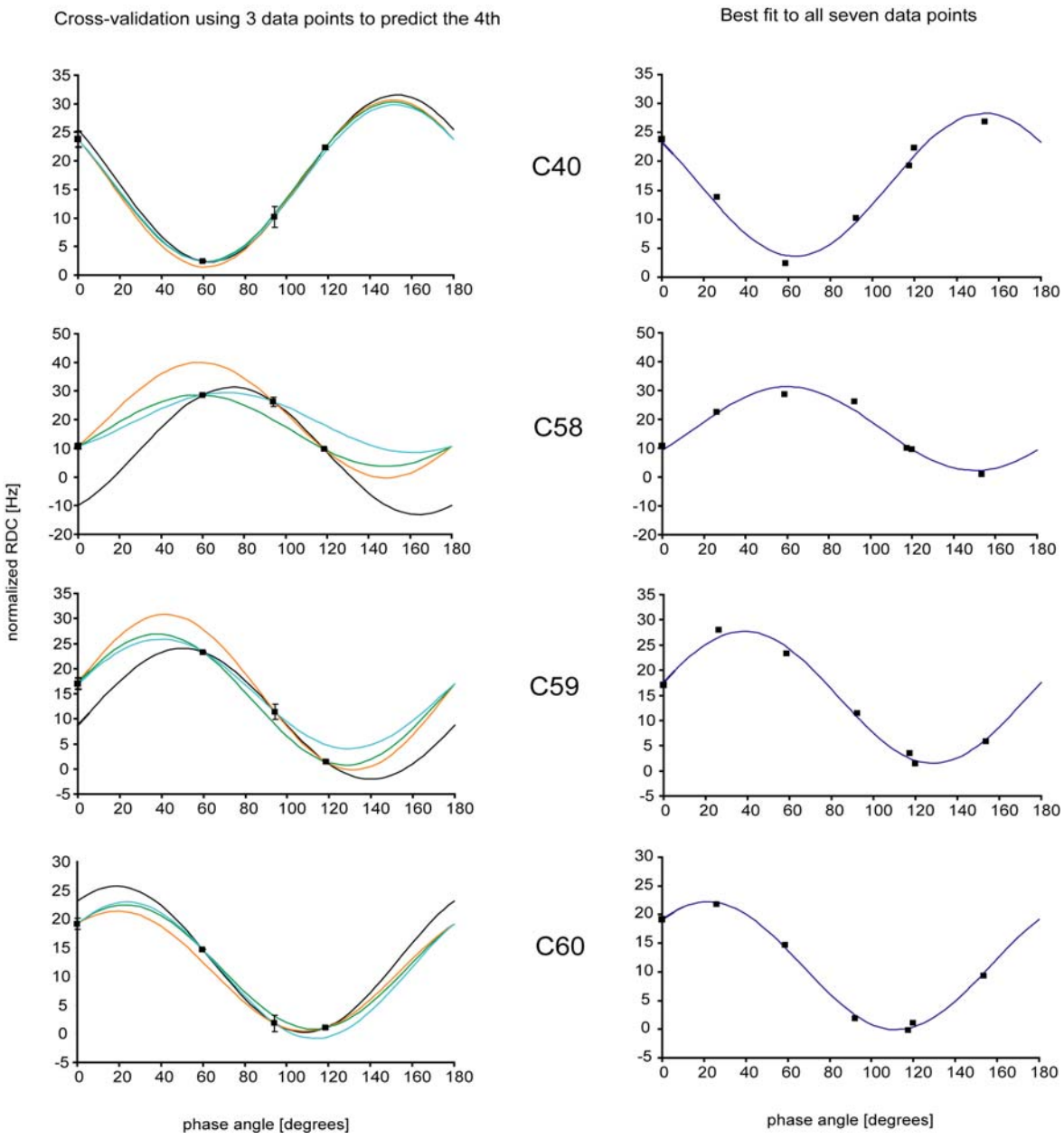
Supporting Information Figure S2. Cross-validation correlation plot for adenines 42, 43, 49, 50, 53, and 56, using ${}^1D_{C2H2}$, ${}^1D_{C8H8}$, ${}^1D_{C5C6}$ and ${}^1D_{C4C5}$ RDCs. For each RDC, the figure shows the experimental value versus the value predicted on the basis of the three other RDCs, when using eq.4 of Bryce and Bax (2004) (analogous to results shown in the left panels of Supporting Information Figure 5). (■) and (●) correspond to the ${}^1D_{CC}$ and ${}^1D_{CH}$ couplings, respectively. The geometry of Figure 5 is used to define the relative orientation of bond vectors. RDCs are normalized to the one-bond C8-H8 interaction. Pearson's correlation coefficient, $R_P = 0.86$; the rmsd between experimental and predicted RDCs equals 4.50 Hz, corresponding to Q = 0.227. Outliers are marked by nucleotide number.



Supporting Information Figure S3. Cross-validation correlation plot for uridine, using only four RDCs, for comparison with adenine (Figure S2). Only $^{1}D_{C6H6}$, $^{1}D_{C5H5}$, $^{1}D_{C5C6}$ and $^{2}D_{C4H5}$ are used. These four RDCs were selected because their orientational distribution is most similar to that for the four available adenine RDCs. For each RDC, the figure shows the experimental value versus the value predicted using the other three RDCs, when using eq. 4 of Bryce and Bax (2004) (analogous to results shown in the left panels of Supporting Information Figure 5). (■), (●) and (◆) correspond to $^{1}D_{CC}$, $^{1}D_{CH}$ and $^{2}D_{CH}$ couplings, respectively. The geometry of Figure 5 is used to define the relative orientation of bond vectors. RDCs are normalized to the aromatic one-bond C-H interaction. Pearson's correlation coefficient, $R_P = 0.84$; the rmsd between experimental and predicted RDCs equals 4.56 Hz, corresponding to Q = 0.23. Outliers are marked by nucleotide number.



Supporting Information Figure S4. Cross-validation correlation plot for cytidines, using only four RDCs. Only $^{1}D_{C6H6}$, $^{1}D_{C5H5}$, $^{1}D_{C5C6}$ and $^{2}D_{C4H5}$ are used. These four RDCs were selected because their orientational distribution is most similar to that for the four available adenine RDCs. For each RDC, the figure shows the experimental value versus the value predicted using the other three RDCs, when using eq. 4 of Bryce and Bax (2004) (analogous to results shown in the left panels of Supporting Information Figure 5). The geometry of Figure 5 is used to define the relative orientation of bond vectors. RDCs are normalized to the aromatic one-bond C-H interaction. Pearson's correlation coefficient, R_P = 0.82; the rmsd between experimental and predicted RDCs equals 7.1 Hz, and is dominated by the large error in the prediction of $^{1}D_{C5C6}$ of C58. Note, however, that this coupling fits well when all 7 couplings are included in the fit (Supporting Information Figure 5).



Supporting Information Figure 5. Fit of base dipolar couplings for the four cytidine bases to $D^{AB}(\varsigma) = D^{AB}_{max} A_{ZZ}\{g_0 + g_2 \cos 2(\varsigma + \psi_2)\}$, where ζ is the known phase angle (0° corresponds to the C_5-C_6 bond orientation), ψ_2 is the fitted phase offset, and $D^{AB}_{max} A_{ZZ} g_0$ and $D^{AB}_{max} A_{ZZ} g_2$ are best-fitted constants, with these three fitted parameters optimized separately for each base. The left panels correspond to the case where three couplings are included in the fit (yielding exact fits as a result of the three adjustable parameters); Black: D_{C5C6} not included in fit; Orange: D_{C5H5} not included; Green: D_{C4H5} not included; Turquoise: D_{C6H6} not included. The right panels correspond to the case where all seven dipolar couplings are included in the fit. In the case of C40, for example, excellent cross-validation is obtained for all 4 data points. For D_{C5C6} in C58, the 3-RDC cross-validation is poor; however, $^1D_{C5C6}$ fits as well as the 6 other RDCs and cross-validates as well as any other data point when 6 RDCs are used in the cross-validation procedure (see Main Text Figure 7).