Modulating RNA Alignment Using Directional Dynamic Kinks: Application in Determining an Atomic-Resolution Ensemble for a Hairpin using NMR Residual Dipolar Couplings

Supporting Information

Loïc Salmon[1], George M. Giambaşu[2], Evgenia N. Nikolova[3], Katja Petzold[4], Akash Bhattacharya[5], David A. Case[2], and Hashim M. Al-Hashimi[6]

[1] Department of Molecular, Cellular, and Developmental Biology and Howard Hughes Medical Institute, University of Michigan, Ann Arbor, Michigan, USA

[2] Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey, USA

[3] Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, California 92037, United States

[4] Department of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm, Sweden

[5] The University of Texas Health Science Center, San Antonio Texas, USA

[6] Department of Biochemistry and Chemistry, Duke University School of Medicine, Durham, North Carolina, USA

	S_{xx} (10^{-4})	S_{yy} (10^{-4})	S_{zz} (10^{-4})	η	α (°)	β (°)	γ (°)
C0	-1.90±0.12	-2.08±0.19	3.98±0.10	0.05±0.07	100.5±7.8	-23.1±4.0	-99.0±17.0
C1	-4.02±0.11	-4.15±0.13	8.17±0.13	0.02±0.02	131.7±2.8	-15.0±1.0	169.0±147.2
C2	-6.02±0.35	-6.88±0.40	12.90±0.07	0.07±0.06	141.6±4.2	-20.5±3.5	-124.9±12.3
C3	-4.35±0.59	-6.05±0.30	10.40±0.77	0.16±0.07	-19.8±48.5	-2.3±4.1	-41.0±67.7
C4	-5.21±0.48	-7.89±0.17	13.10±0.53	0.20±0.04	116.5±32.7	-6.1±1.9	-166.6±32.9
C5	-6.71±0.14	-7.14±0.19	13.85±0.18	0.03±0.02	-115.5±14.2	-2.2±0.6	122.1±10.3
C6	-2.25±0.08	-5.92±0.07	8.17±0.10	0.45±0.01	178.5±125.6	-29.1±0.7	178.4±125.8

Table S1: Order tensor parameters obtained for the different constructs (C0-C6, see Figure 1) obtained by fitting RDCs measured in helical residues to an idealized A-form geometry.

	C0	C1	C2	C3	C4	C5	C6
C0	1.00	0.93	0.89	0.79	0.88	0.76	0.54
C1	0.93	1.00	0.98	0.88	0.96	0.89	0.78
C2	0.89	0.98	1.00	0.78	0.89	0.80	0.85
C3	0.79	0.88	0.78	1.00	0.97	0.99	0.62
C4	0.88	0.96	0.89	0.97	1.00	0.97	0.70
C5	0.76	0.89	0.80	0.99	0.97	1.00	0.70
C6	0.54	0.78	0.85	0.62	0.70	0.70	1.00

Table S2: Scalar products between the different order tensors reported in Table S1. Construct numbering is according to Figure 1.

Static	S_{xx} (10^{-4})	S_{yy} (10^{-4})	S_{zz} (10^{-4})	Dynamic	S_{xx} (10^{-4})	S_{yy} (10^{-4})	S_{zz} (10^{-4})
C0	-1.90	-2.08	3.98	C0	-1.67	-2.54	4.21
C1	-4.02	-4.15	8.17	C1	-4.23	-5.17	9.40
C2	-6.02	-6.88	12.90	C2	-6.50	-7.38	13.88
C3	-4.35	-6.05	10.40	C3	-5.29	-5.94	11.28
C4	-5.21	-7.89	13.10	C4	-7.59	-8.50	16.09
C5	-6.71	-7.14	13.85	C5	-6.16	-7.91	14.06
C6	-2.25	-5.92	8.17	C6	-2.78	-6.61	9.39

Table S3: Comparison of the order tensors obtained using the static (as in Table S1) and dynamic (during ensemble selection) approach for the different constructs. Construct numbering is according to Figure 1.

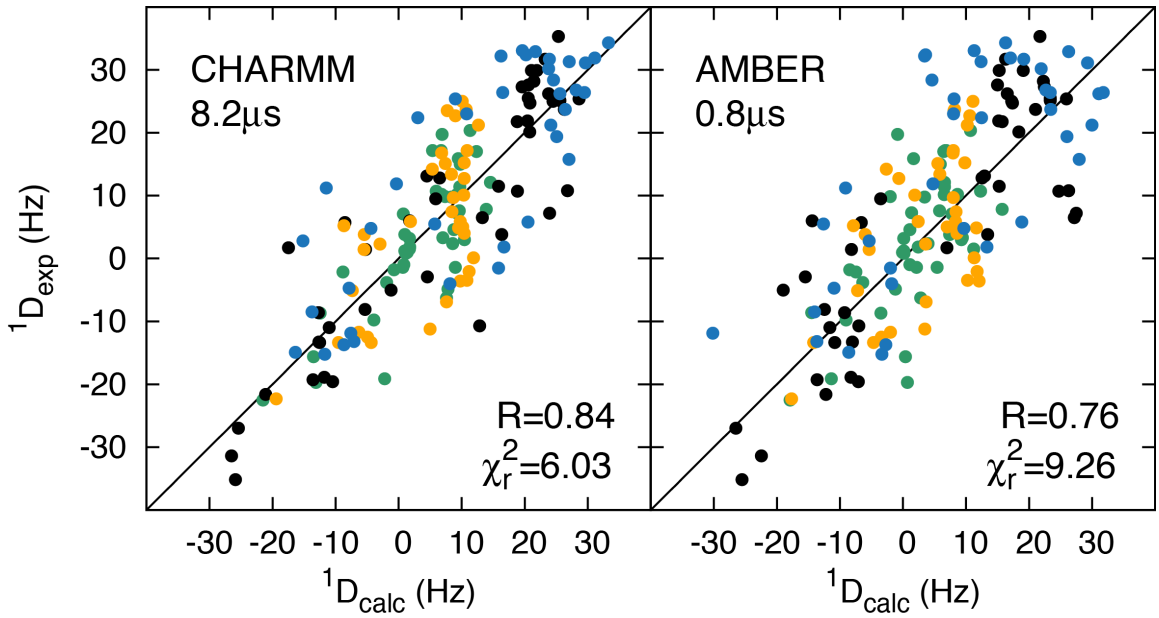

Figure S1: Capability of two different MD trajectories to reproduce experimental RDCs measured in four different HIV-1 TAR constructs reported in the literature[1]. Shown is the data reproduction for (left) an 8.2µs trajectory computed using the CHARMM36 force field and (right) an 0.8µs trajectory computed using the AMBER ff10 force field. The details of the CHARMM trajectory and RDC analysis can be found in ref[1].

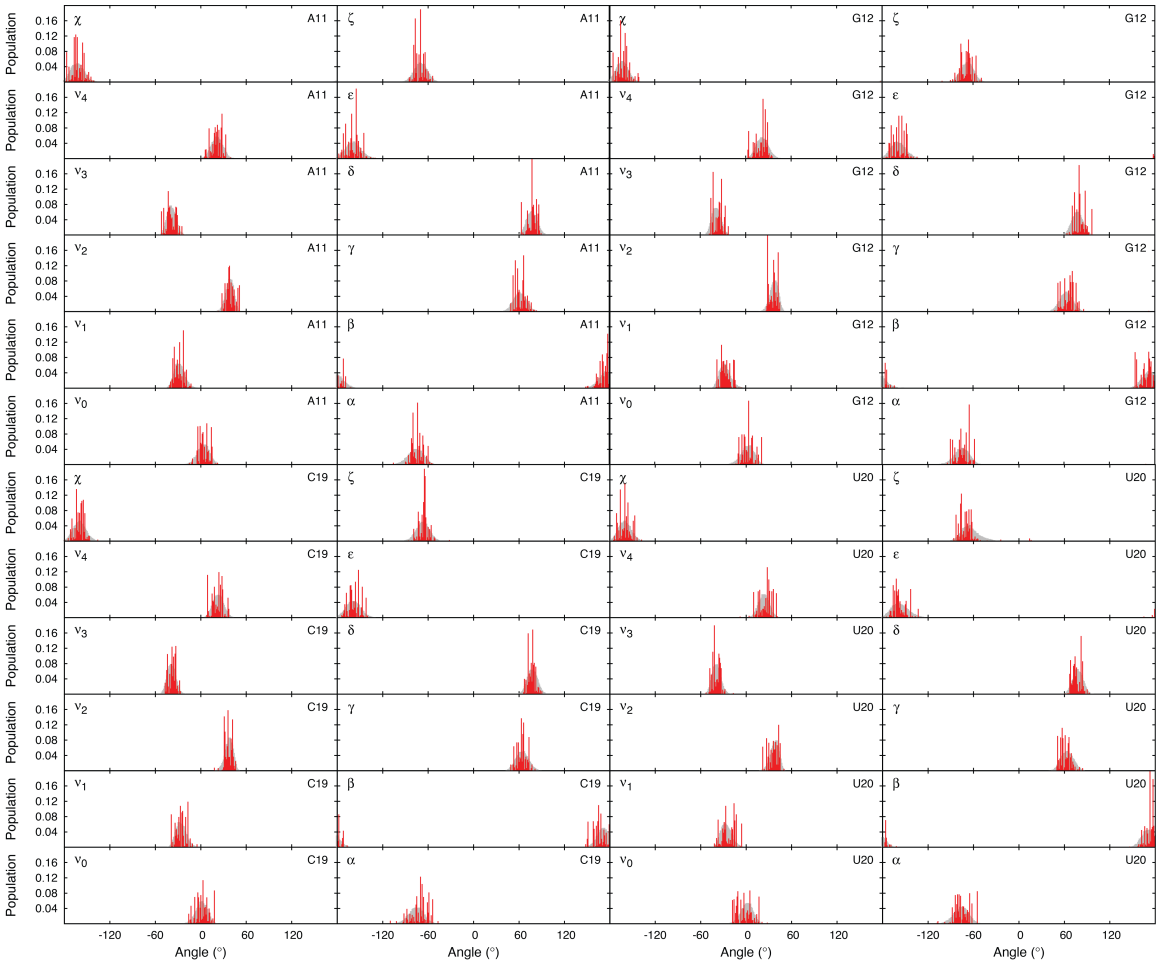

Figure S2: Local dihedral angle distributions for the helical nucleotides. Distribution of local dihedral angle for nucleotides A11, G12, C19 and U20 as observed in the MD pool (grey) and the ensemble selected using the experimental RDCs (red).

	avg.	ν0	ν1	ν2	ν3	ν4	α	β	γ	δ	ε	ζ	χ
A11	ENS	4.41	-28.33	39.78	-38.94	21.86	-74.01	174.87	62.19	77.88	-158.76	-70.71	-162.07
	MD	3.70	-27.41	38.57	-38.14	22.02	-76.38	176.65	61.99	77.43	-157.81	-69.38	-161.93
G12	ENS	3.61	-25.95	36.39	-35.73	20.81	-73.40	170.69	66.23	82.09	-157.59	-68.03	-161.15
	MD	2.72	-26.47	37.90	-37.99	22.61	-74.83	173.96	63.94	78.00	-157.91	-67.33	-162.43
C13	ENS	1.91	-26.14	38.12	-38.50	23.46	-76.37	170.78	65.46	78.42	-152.09	-60.90	-160.53
	MD	1.54	-26.19	38.72	-39.24	24.15	-75.35	172.96	63.79	77.50	-154.36	-61.83	-160.61
U14	ENS	2.49	-26.88	39.06	-38.59	22.98	-76.53	176.01	61.13	74.75	-172.74	-91.08	-155.73
	MD	0.71	-24.58	36.89	-38.10	23.97	-74.97	173.96	63.85	75.45	-173.04	-93.87	-153.47
U15	ENS	-23.59	34.00	-29.79	17.70	3.59	-165.16	153.03	56.48	139.87	-81.34	-72.56	-129.28
	MD	-15.34	27.96	-28.51	20.80	-3.83	-163.70	158.81	56.55	140.85	-85.72	-73.86	-138.58
C16	ENS	-21.21	36.99	-37.29	25.99	-3.34	-72.69	-176.81	59.74	149.27	-91.86	76.09	-132.06
	MD	-21.18	34.87	-34.24	23.16	-1.51	-70.36	-179.03	59.55	146.35	-87.55	77.20	-142.60
G17	ENS	-2.34	-16.81	27.78	-30.18	20.98	68.90	-173.22	173.34	87.70	-149.33	-57.87	57.01
	MD	1.13	-15.71	22.71	-23.14	14.28	67.01	179.62	174.28	95.50	-124.95	-49.11	64.19
G18	ENS	5.40	-27.71	37.69	-35.90	19.45	-173.68	-154.09	163.03	79.92	-153.35	-67.14	-164.78
	MD	6.94	-28.76	37.72	-35.05	18.17	-83.87	114.02	146.96	82.24	-158.49	-64.04	-160.67
C19	ENS	2.36	-25.71	37.08	-37.14	22.40	-70.05	169.33	64.16	76.84	-156.94	-65.23	-159.57
	MD	1.41	-25.45	37.68	-38.17	23.63	-75.72	172.44	65.05	78.03	-158.04	-66.68	-159.39
U20	ENS	-2.24	-22.46	36.54	-38.81	26.42	-74.09	174.68	61.36	76.55	-159.65	-69.57	-159.76
	MD	1.31	-25.60	37.95	-38.60	23.93	-76.64	171.92	64.69	78.21	-156.34	-63.19	-159.57
	var.	ν0	ν1	ν2	ν3	ν4	α	β	γ	δ	ε	ζ	χ
A11	ENS	0.007	0.006	0.005	0.007	0.007	0.009	0.011	0.008	0.006	0.012	0.006	0.010
	MD	0.010	0.007	0.004	0.005	0.008	0.023	0.015	0.017	0.007	0.015	0.011	0.012
G12	ENS	0.011	0.007	0.004	0.006	0.010	0.016	0.015	0.012	0.008	0.012	0.009	0.009
	MD	0.011	0.008	0.004	0.005	0.009	0.025	0.015	0.018	0.007	0.014	0.010	0.009
C13	ENS	0.005	0.003	0.002	0.003	0.004	0.012	0.011	0.005	0.006	0.013	0.008	0.010
	MD	0.008	0.006	0.004	0.004	0.007	0.023	0.012	0.017	0.006	0.013	0.011	0.009
U14	ENS	0.008	0.008	0.005	0.003	0.005	0.014	0.011	0.006	0.005	0.018	0.009	0.005
	MD	0.009	0.008	0.005	0.005	0.007	0.023	0.010	0.017	0.007	0.013	0.013	0.013
U15	ENS	0.012	0.008	0.006	0.008	0.013	0.021	0.015	0.014	0.013	0.015	0.011	0.058
	MD	0.037	0.027	0.012	0.012	0.029	0.028	0.031	0.015	0.013	0.020	0.017	0.090
C16	ENS	0.008	0.005	0.003	0.004	0.008	0.021	0.012	0.016	0.004	0.012	0.010	0.048
	MD	0.013	0.008	0.007	0.012	0.016	0.018	0.012	0.011	0.013	0.018	0.013	0.024
G17	ENS	0.016	0.011	0.011	0.015	0.019	0.019	0.013	0.010	0.016	0.103	0.123	0.038
	MD	0.026	0.012	0.012	0.026	0.036	0.029	0.018	0.021	0.029	0.193	0.248	0.048
G18	ENS	0.017	0.011	0.004	0.006	0.014	0.565	0.514	0.207	0.009	0.012	0.009	0.011
	MD	0.012	0.007	0.004	0.006	0.012	0.326	0.393	0.279	0.008	0.012	0.011	0.012
C19	ENS	0.011	0.008	0.004	0.003	0.009	0.019	0.017	0.006	0.004	0.010	0.006	0.006
	MD	0.009	0.007	0.004	0.004	0.007	0.022	0.012	0.017	0.006	0.014	0.011	0.010
U20	ENS	0.017	0.013	0.007	0.004	0.011	0.017	0.007	0.008	0.006	0.019	0.020	0.009
	MD	0.009	0.007	0.005	0.004	0.007	0.026	0.013	0.019	0.006	0.021	0.032	0.010

Table S4: Average values and corresponding standard deviations of the different backbone dihedral angles sampled in the cUUCGg loop by the RDC-selected (ENS) and MD-generated (MD) trajectory.

	C13		U14		U15		C16		G17		G18	
	ENS	MD	ENS	MD	ENS	MD	ENS	MD	ENS	MD	ENS	MD
O1'exo	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C1'endo	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	2.2	12.4	0.0	0.2
C2'exo	0.9	4.7	1.9	4.2	0.0	0.0	0.0	0.0	10.9	16.0	22.3	25.8
C3'endo	98.6	91.7	90.5	89.6	0.0	0.0	0.0	0.2	64.8	46.8	68.8	71.5
C4'exo	0.5	3.7	7.6	6.1	0.0	0.0	0.0	0.2	22.1	22.1	8.9	2.6
O1'endo	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C1'exo	0.0	0.0	0.0	0.0	16.1	13.4	0.0	9.2	0.0	0.0	0.0	0.0
C2'endo	0.0	0.0	0.0	0.0	81.3	57.7	91.7	76.0	0.0	0.0	0.0	0.0
C3'exo	0.0	0.0	0.0	0.0	2.1	16.7	7.8	14.1	0.0	0.0	0.0	0.0
C4'endo	0.0	0.0	0.0	0.0	0.4	10.2	0.5	0.1	0.0	0.0	0.0	0.0

Table S5: Population of the different sugar puckers sampled in the cUUCGg loop by the RDC-selected (ENS) and MD-generated (MD) trajectory.

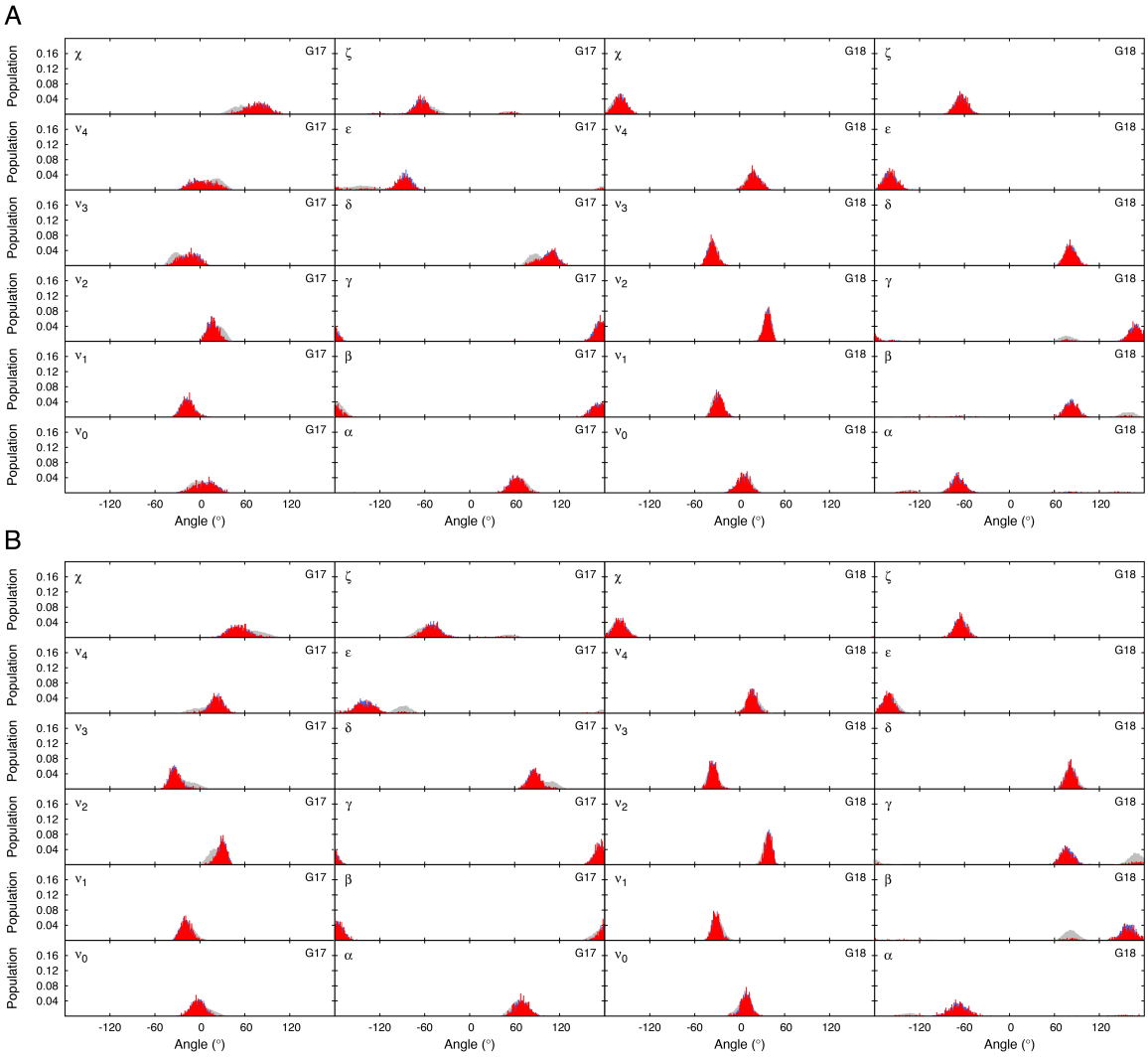

Figure S3: Capturing local dihedral angle dynamics using synthetic RDC data. The synthetic RDCs were used in a selection procedure identical to the one used for the experimental data (see Methods) and the capability of the selection to reproduce the local dihedral angles was investigated (A) Distribution of local dihedral angle for nucleotides 17 and 18 as observed in the MD pool (grey) the synthetic target ensemble 1 (blue) and the ensemble selected using the synthetic data generated from target 1 (red). (B) Distribution of local dihedral angles for nucleotides 17 and 18 as observed in the MD pool (grey) the synthetic target ensemble 2 (blue) and the ensemble selected using the synthetic RDC data generated from target ensemble 2 (red).

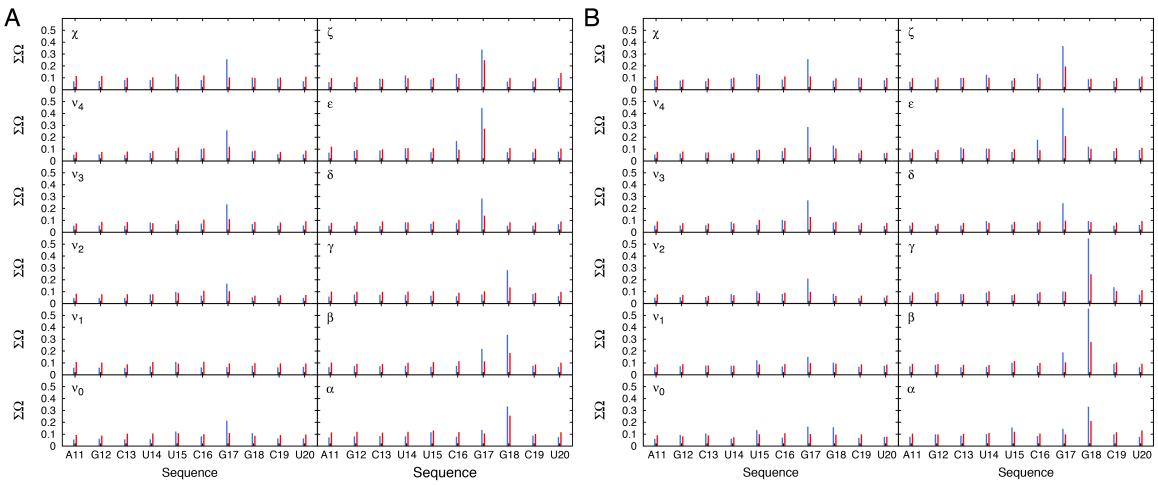

Figure S4: Capturing local dihedral angle dynamics using simulated data as evaluated using REnsemble[2]. Shown as ΣΩ values comparing backbone dihedral angles between the target and the MD-generated ensemble (blue) and the target and RDC-selected ensemble (red). Data are shown for two different target ensembles (A) target ensemble 1 and (B) target ensemble 2.

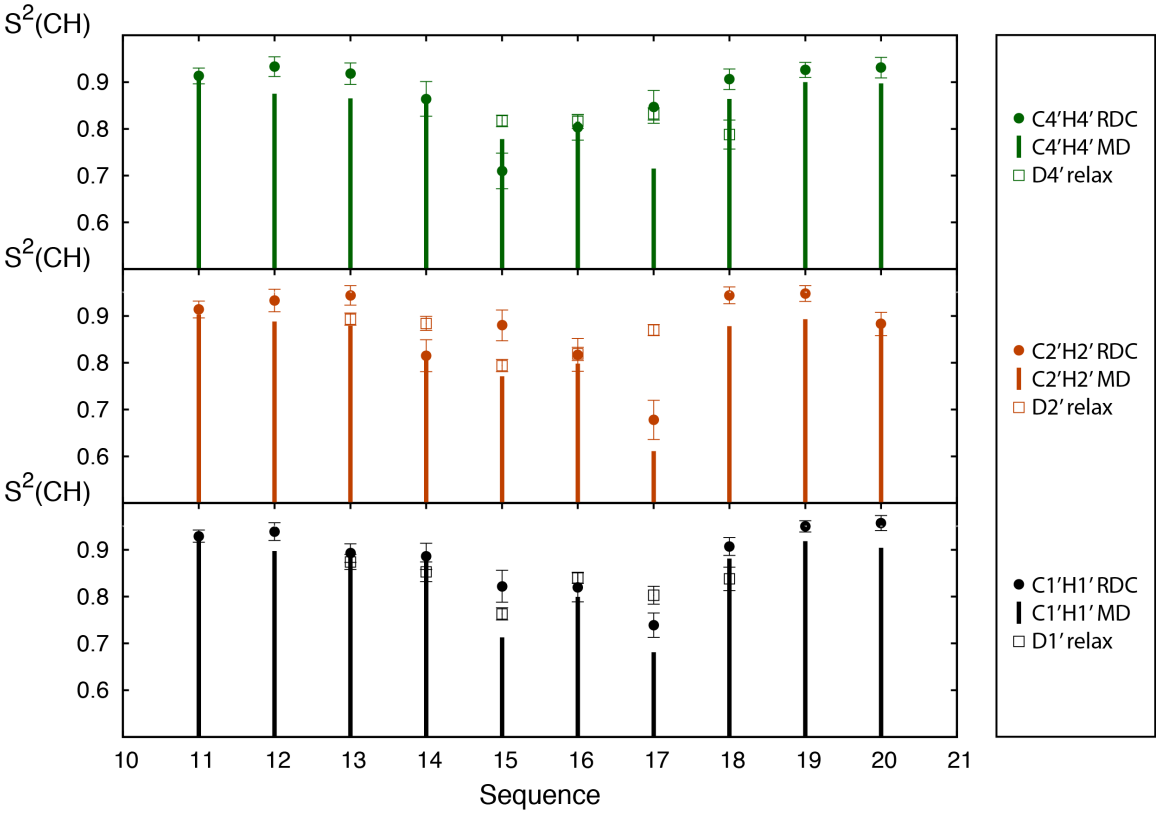

Figure S5: Conformational disorder in the UUCG tetraloop. Comparison of the Lipari-Szabo in relaxation order parameter (S^2) for C4'H4' and D4' (green), C2'H2' and D2' (orange) and C1'H1' and D1' (black) obtained using the RDC-selected ensemble (filled circles), MD-generated ensemble (bars), and ^{2}H spin relaxation data (open squares). measured in a related UUCG tetraloop[3]. Note that differences in the sequence of bps following the closing bp in the uUUCGg loop used in this study and in the ^{2}H relaxation study could contribute to differences for residues 13 and 18. Error bars for the RDC-selected ensemble are obtained using Monte Carlo simulations.

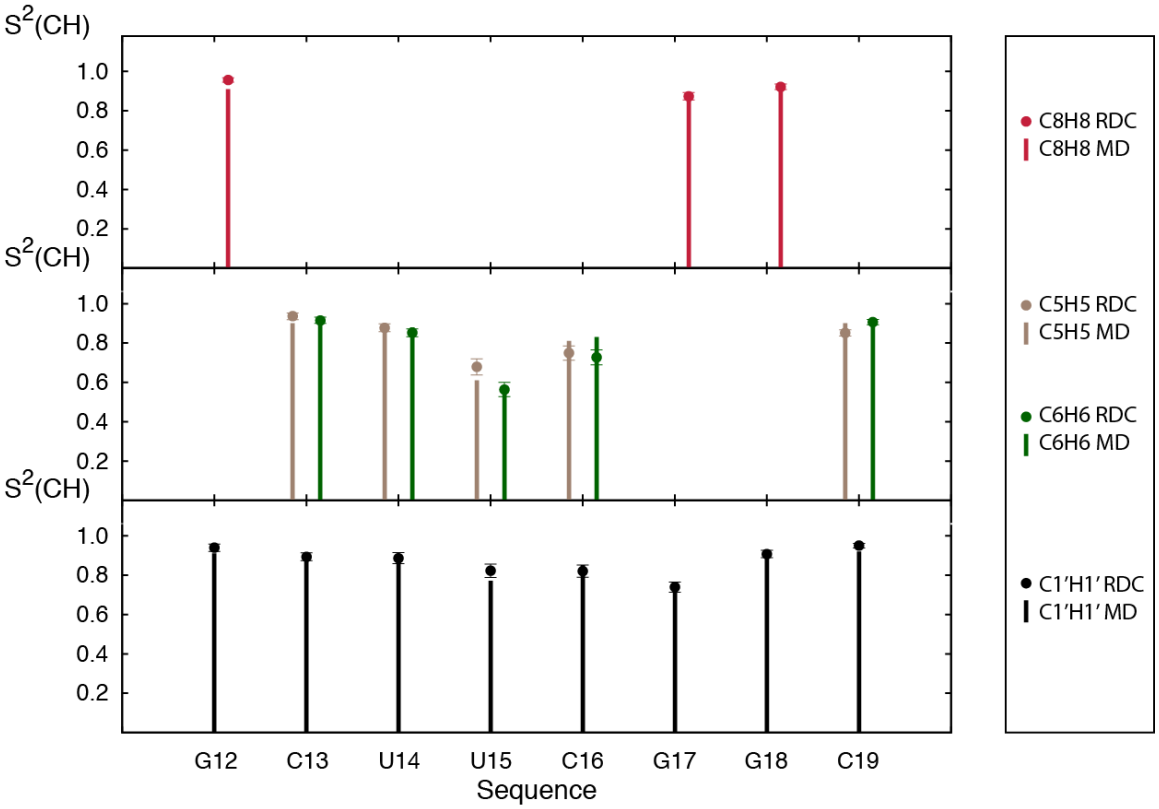

Figure S6: Conformational disorder in the cUUCGg tetraloop. Comparison of the order parameter (S^2) of several bond vectors C8H8 (red), C5H5 (brown), C6H6 (green) and C1'H1' (black). Bars correspond to the order parameters computed for a similar tetraloop using the force-field *ff10*[4] and filled circle to the order parameter computed for the RDC-selected ensemble. Error bars for the RDC-selected ensemble are obtained using Monte Carlo simulations (see Methods).

	PDB ID	Chain	Start	End	Method	Class
1	1AUD	B	28	35	NMR	RNA
2	1BGZ	A	10	15	NMR	RNA
3	1BYJ	A	11	16	NMR	RNA
4	1D6K	B	283	293	NMR	RIBOSOME
5	1DK1	B	8	13	X-Ray	RIBOSOME
6	1DK1	B	32	37	X-Ray	RIBOSOME
7	1EBQ	A	12	17	NMR	RNA
8	1EBR	A	12	17	NMR	RNA
9	1EBS	A	12	17	NMR	RNA
10	1F6Z	A	12	17	NMR	RNA
11	1F79	A	12	17	NMR	RNA
12	1F7G	A	12	17	NMR	RNA
13	1F7I	A	12	17	NMR	RNA
14	1F7Y	B	8	13	X-Ray	RIBOSOME
15	1F7Y	B	32	37	X-Ray	RIBOSOME
16	1FMN	A	16	21	NMR	RNA
17	1FYO	A	11	16	NMR	RNA
18	1FYP	A	11	16	NMR	RNA
19	1G70	A	54	64	NMR	VIRAL
20	1HO6	A	4	9	NMR	DNA
21	1HOQ	A	4	9	NMR	DNA
22	1I6U	C	16	21	X-Ray	RIBOSOME
23	1I6U	D	16	21	X-Ray	RIBOSOME
24	1IKD	A	7	12	NMR	RNA
25	1K2G	A	5	10	NMR	RNA
26	1KP7	A	14	19	NMR	RNA
27	1KUQ	B	8	13	X-Ray	RIBOSOME
28	1KUQ	B	32	37	X-Ray	RIBOSOME
29	1M82	A	10	15	NMR	RNA
30	1MFY	A	14	19	NMR	RNA
31	1PBR	A	11	16	NMR	RNA
32	1RAW	A	21	26	NMR	RNA
33	1TLR	A	10	15	NMR	RNA
34	1ULL	A	14	19	NMR	VIRAL
35	1VX8	A	419	424	X-Ray	RIBOSOME
36	1VXI	A	419	424	X-Ray	RIBOSOME
37	1XSG	A	12	17	NMR	RNA
38	1XSH	A	12	17	NMR	RNA
39	1XST	A	12	17	NMR	RNA
40	1XSU	A	12	17	NMR	RNA

	PDB ID	Chain	Start	End	Method	Class
41	1Z31	A	270	286	NMR	RNA
42	1ZHO	B	15	20	X-Ray	STRUCTURAL
43	1ZHO	D	15	20	X-Ray	STRUCTURAL
44	1ZHO	F	15	20	X-Ray	STRUCTURAL
45	1ZHO	H	15	20	X-Ray	STRUCTURAL
46	2AU4	A	20	25	NMR	RNA
47	2CD3	A	12	17	NMR	NUCLEIC
48	2CD6	A	12	17	NMR	NUCLEIC
49	2GIO	A	13	18	NMR	RNA
50	2GIP	A	12	17	NMR	RNA
51	2HEM	A	10	15	NMR	RNA
52	2HGH	B	32	37	NMR	TRANSCRIPTION/RNA
53	2HUA	A	17	22	NMR	RNA
54	2HW8	B	15	20	X-Ray	STRUCTURAL
55	2J00	A	419	424	X-Ray	RIBOSOME
56	2J02	A	419	424	X-Ray	RIBOSOME
57	2KHY	A	17	22	NMR	RNA
58	2KMJ	A	29	36	NMR	RNA/PEPTIDE
59	2KOC	A	5	10	NMR	RNA
60	2KTZ	A	19	24	NMR	RNA
61	2KU0	A	19	24	NMR	RNA
62	2KZL	A	27	32	NMR	RNA
63	2L2J	A	19	24	NMR	RNA
64	2L2K	A	19	24	NMR	HYDROLASE/RNA
65	2L3E	A	18	23	NMR	RNA
66	2LHP	A	16	21	NMR	RNA
67	2LKR	A	63	68	NMR	RNA
68	2LQZ	A	13	18	NMR	RNA
69	2LU0	A	28	33	NMR	RNA
70	2LUB	A	16	21	NMR	RNA
71	2LUN	A	12	17	NMR	RNA
72	2LWK	A	15	20	NMR	RNA
73	2M4Q	1	11	16	NMR	RNA/ANTIBIOTIC
74	2M58	A	12	17	NMR	RNA
75	2UUA	A	419	424	X-Ray	RIBOSOME
76	2UUB	A	419	424	X-Ray	RIBOSOME
77	2UXC	A	419	424	X-Ray	RIBOSOME
78	2VQE	A	419	424	X-Ray	RIBOSOME
79	2VQF	A	419	424	X-Ray	RIBOSOME
80	3AM1	B	32	37	X-Ray	TRANSFERASE/RNA

	PDB ID	Chain	Start	End	Method	Class
81	3T1Y	A	414	419	X-Ray	RIBOSOME/ANTIBIOTIC
82	3V2C	A	419	424	X-Ray	RIBOSOME/INHIBITOR
83	3V2E	A	419	424	X-Ray	RIBOSOME/INHIBITOR
84	3V7E	C	65	70	X-Ray	RIBOSOMAL
85	3V7E	D	265	270	X-Ray	RIBOSOMAL
86	4AY2	C	8	13	X-Ray	HYDROLASE/RNA
87	4B3M	A	419	424	X-Ray	RIBOSOME
88	4JF2	A	60	65	X-Ray	RNA
89	4JUW	A	419	424	X-Ray	RIBOSOME
90	4KIY	A	419	424	X-Ray	RIBOSOME/ANTIBIOTIC
91	4KIY	A	1449	1454	X-Ray	RIBOSOME/ANTIBIOTIC
92	4KJ0	A	419	424	X-Ray	RIBOSOME/ANTIBIOTIC
93	4KJ0	A	1449	1454	X-Ray	RIBOSOME/ANTIBIOTIC
94	4KJ2	A	419	424	X-Ray	RIBOSOME/ANTIBIOTIC
95	4KJ2	A	1028	1033	X-Ray	RIBOSOME/ANTIBIOTIC
96	4KJ2	A	1449	1454	X-Ray	RIBOSOME/ANTIBIOTIC
97	4KJ4	A	419	424	X-Ray	RIBOSOME/ANTIBIOTIC
98	4KJ4	A	1449	1454	X-Ray	RIBOSOME/ANTIBIOTIC
99	4KJ6	A	207	212	X-Ray	RIBOSOME/ANTIBIOTIC
100	4KJ6	A	419	424	X-Ray	RIBOSOME/ANTIBIOTIC
101	4KJ6	A	1449	1454	X-Ray	RIBOSOME/ANTIBIOTIC
102	4KJ8	A	419	424	X-Ray	RIBOSOME/ANTIBIOTIC
103	4KJ8	A	1449	1454	X-Ray	RIBOSOME/ANTIBIOTIC
104	4KJA	A	419	424	X-Ray	RIBOSOME/ANTIBIOTIC
105	4KJA	A	1028	1033	X-Ray	RIBOSOME/ANTIBIOTIC
106	4KJA	A	1449	1454	X-Ray	RIBOSOME/ANTIBIOTIC
107	4KJC	A	207	212	X-Ray	RIBOSOME
108	4KJC	A	419	424	X-Ray	RIBOSOME
109	4KJC	A	1449	1454	X-Ray	RIBOSOME
110	4PDB	I	17	22	X-Ray	RIBOSOMAL

Table S6: NMR and X-ray crystallography structures with resolution <3Å containing the cUUCGg tetraloop motif obtained using FRABASE.

REFERENCES

1. Salmon, L.; Bascom, G.; Andricioaei, I.; Al-Hashimi, H. M. *J. Am. Chem. Soc.,* **2013**, *135*, 5457-5466.
2. Yang, S.; Salmon, L.; Al-Hashimi, H. M. *Nat Methods,* **2014**, *1*, 552-554.
3. Vallurupalli, P.; Kay, L. E. *J. Am. Chem. Soc.* **2005**, *127*, 6893.
4. Giambasu, G. M.; York, D. M.; Case, D. A. *RNA,* **2015,** 21, 963–974.