

Supplementary Data

Automated and assisted RNA resonance assignment using NMR chemical shift statistics

**Thomas Aeschbacher¹, Elena Schmidt², Markus Blatter¹, Christophe Maris¹,
Olivier Duss¹, Frédéric H.-T. Allain¹, Peter Güntert^{2,3} and Mario Schubert¹**

¹Institute of Molecular Biology and Biophysics, ETH Zürich, 8093 Zurich, Switzerland, ²Institute of Biophysical Chemistry, Center for Biomolecular Magnetic Resonance, and Frankfurt Institute of Advanced Studies, 60438 Frankfurt am Main, Germany, ³Graduate School of Science and Engineering, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan

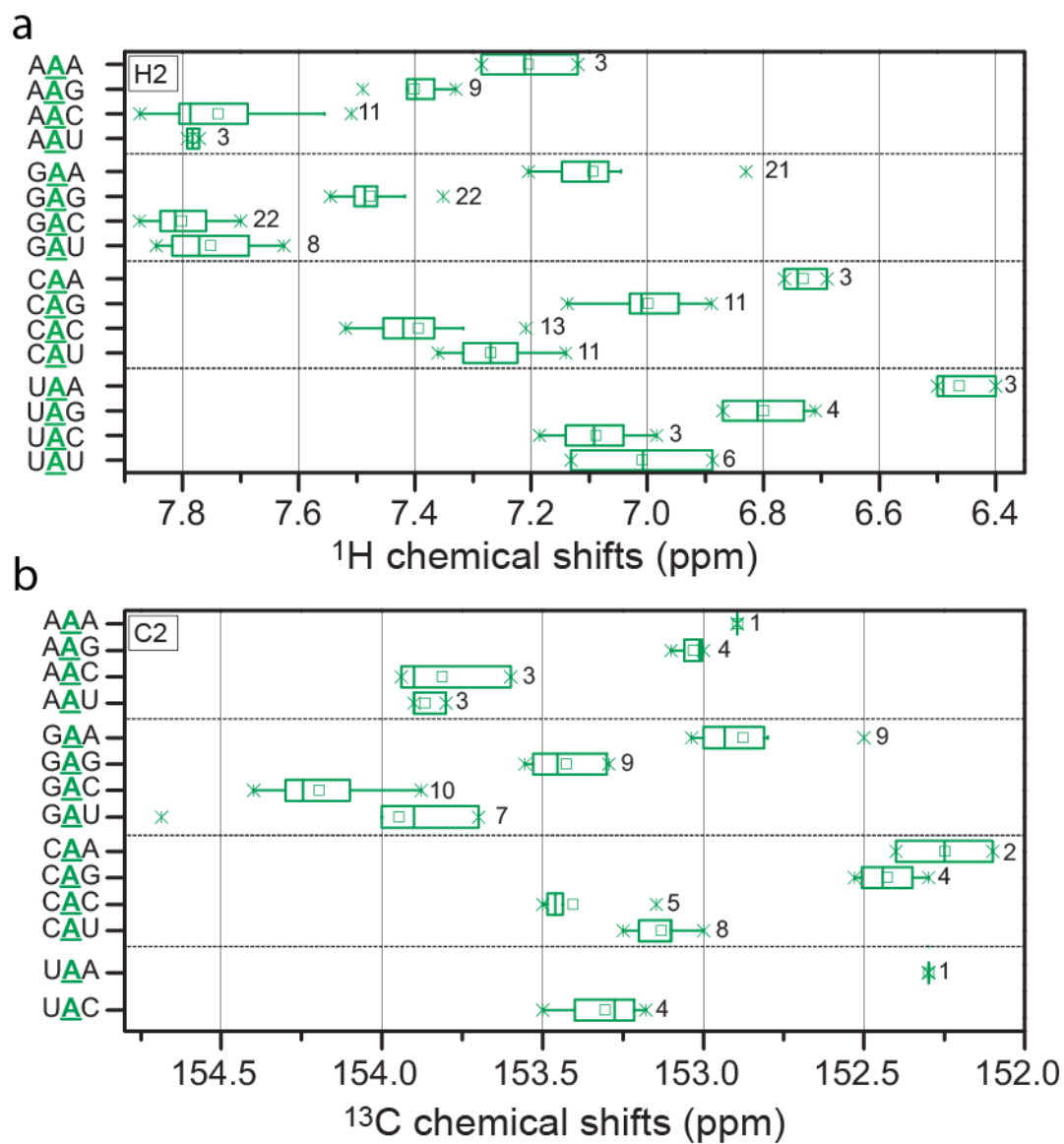


Figure S1. Box plots of adenosine H2 (a) and C2 (b) chemical shifts for Watson-Crick base-paired triplets. See Figure 1 for details.

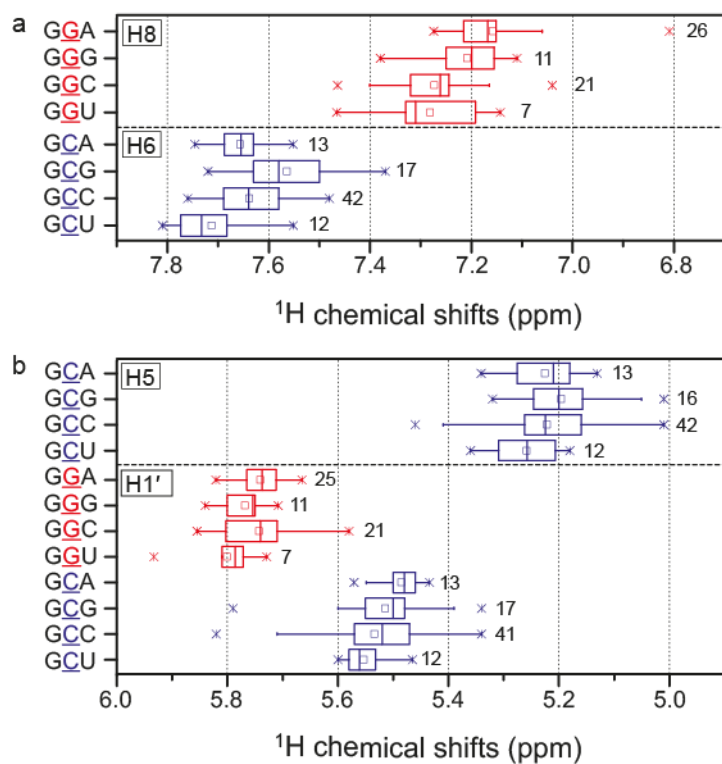


Figure S2. Influence of the 3' neighboring nucleotide on ¹H chemical shifts illustrated with highly abundant Watson-Crick base-paired triplets. (a) ¹H chemical shift statistics of H8 within GGX and H6 within GCX. The number of data points is given next to each box plot. (b) ¹H chemical shift statistics of H5 within GCX triplets and of H1' within GGX and GCX triplets.

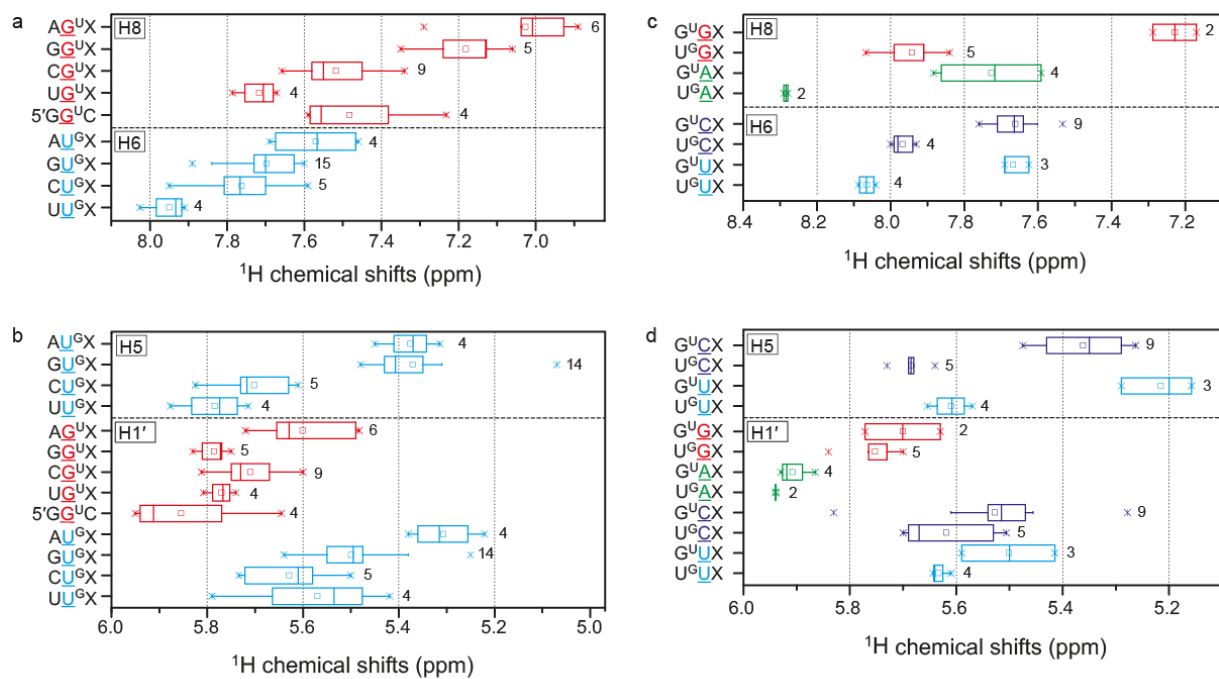


Figure S3. Influence of G•U wobble base-pairs on ^1H chemical shifts. (a) ^1H chemical shift statistics of H8 and H6 of G•U wobble base-pairs flanked by two Watson-Crick base-pairs. The number of data points is given next to each box plot. (b) ^1H chemical shift statistics of H5 and H1' of G•U wobble base-pairs flanked by two Watson-Crick base-pairs. (c) ^1H chemical shift statistics of H8 and H6 of nucleotides that follow a G•U wobble base-pairs. Considered are base-pair triplets containing a wobble base-pair followed by two Watson-Crick base-pairs. (d) ^1H chemical shift statistics of H5 and H1' of nucleotides that follow a G•U wobble base-pairs.

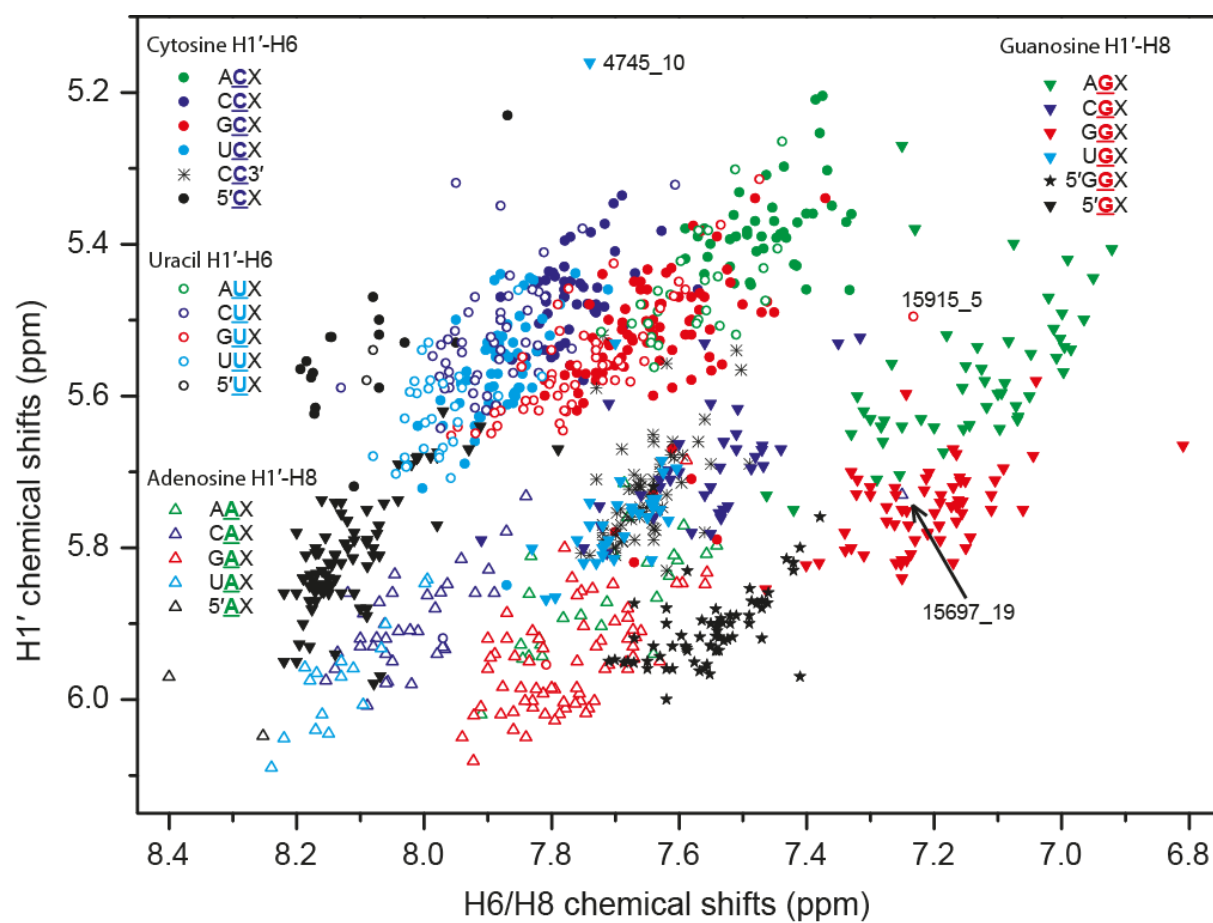


Figure S4. Intranucleotide H1'-H6/H8 chemical shift correlations for Watson-Crick base-paired triplets. Extreme outliers are labeled with the BMRB accession code and the residue number.

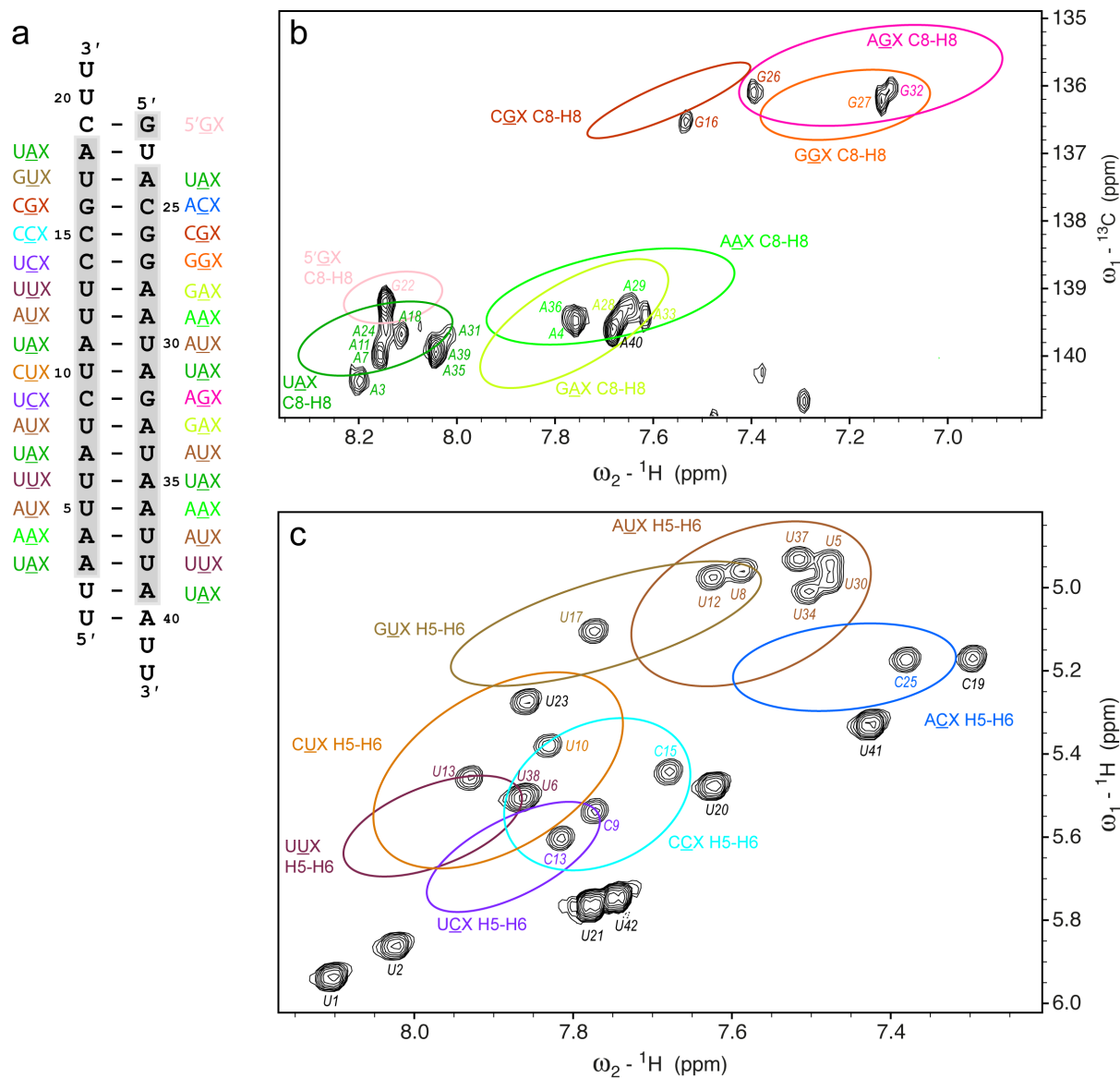


Figure S5. Assisted assignment of the small interfering RNA ELAVL1 (42 nts) using 2D chemical shift statistics. (a) Schematic presentation of the RNA sequence. For the grey highlighted nucleotides 2D chemical shift statistics/ellipses are available. For each of those nucleotides the corresponding categories of the statistics are given with a color code. (b) Natural abundance ^{13}C -HSQC of ELAVL1 recorded with 50 transients at 750 MHz and 313 K displaying purine C8-H8 correlations. 2D statistics are indicated by ellipses (86%) that are color-coded as in panel a. The final assignments for each signal is given in italic using the same color code. G22 can be readily assigned as the solely signal in the ellipse of 5'GX. For other signals the assignment ambiguity can significantly be reduced. The signals of G16 and G26 are examples of the unusual case that the signals are slightly outside the 86% ellipse. (c) 2D TOCSY spectrum of ELAVL1 recorded with 2 transients at 750 MHz and 313 K displaying pyrimidine H5-H6 correlations. Statistics are displayed in form of ellipses (86%) and the final assignments are given in italic. The same color code is used as in panel a. Black labels indicate nucleotides that are in a non-regular environment for which 2D statistics were not derived mainly due to an insufficient amount of data. C25 can be readily assigned. The statistics were also very helpful to assign the signals of C9, C13, U17 to name a few. Together with the 1D statistics and a 2D NOESY the nonexchangeable protons of H8, H6, H5, H2 and H1' could be assigned in a straight-forward way without the requirement of isotope labeling and the recording of 3D spectra.

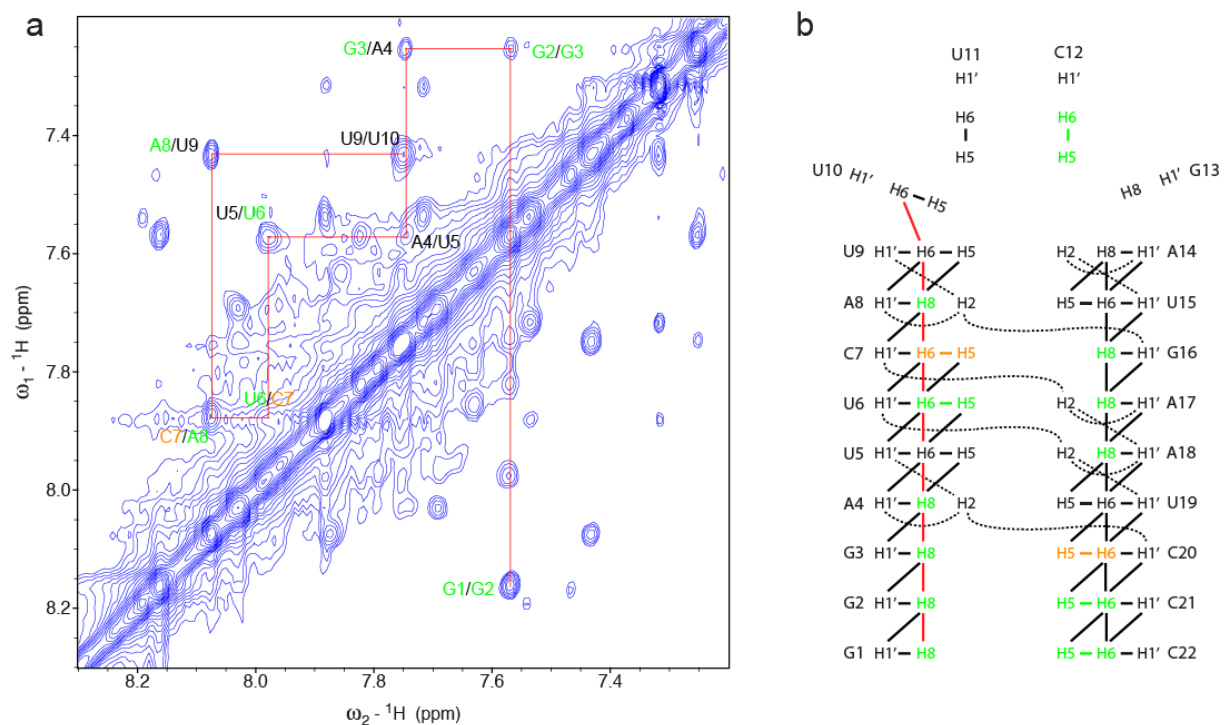


Figure S6. Linking starting points by conventional assignment walks. (a) Region of a 2D NOESY spectrum of TASL1 recorded in D₂O at 700 MHz. The H8/H6–H8/H6 assignment walk between nucleotides G1 to U10 is shown by red lines. Resonances of unambiguous starting points are colored green. Ambiguous starting points are shown in orange. (b) Schematic presentation of TASL1 indicating expected NOE patterns. Starting points are color-coded as in panel a. The same H8/H6–H8/H6 assignment walk as indicated in the NOESY spectrum is shown by red lines.

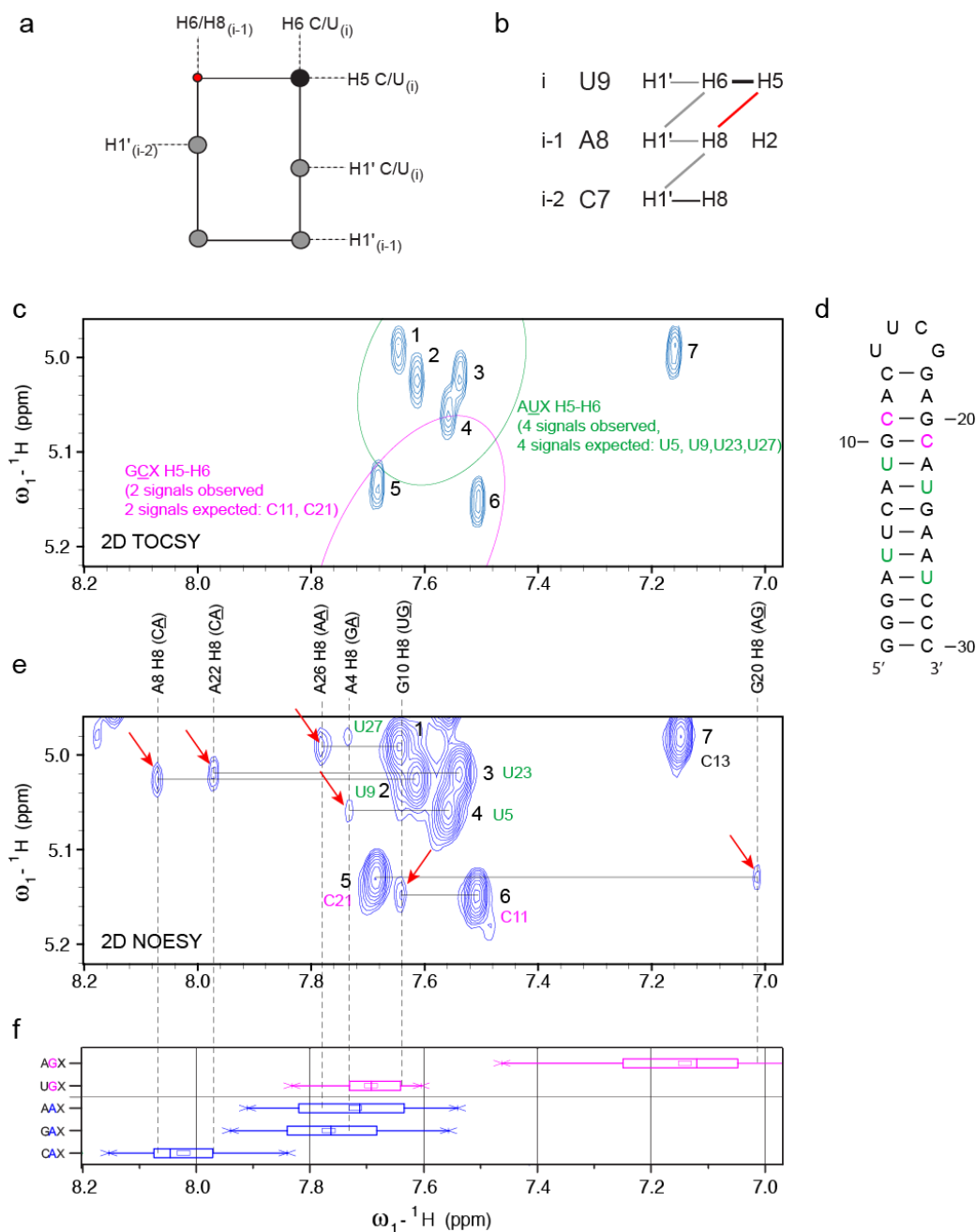


Figure S7. Combined use of 2D and 1D statistics. (a) NOE pattern expected for a pyrimidine (residue i) following a purine (residue $i - 1$). The NOE signals of the H1'-H6/H8 assignment walk are indicated in grey, strong intra-base pyrimidine H5-H6 NOEs are colored black and H5-H8 _{$i - 1$} correlations between pyrimidines and purines are colored red. (b) Example of an RNA sequence in which the same NOE correlations are shown schematically. The same color code is used. (c) Region of a 2D TOCSY spectrum of TASL3 with overlaid covariance ellipses for the categories AUX and GCX. From the C5 chemical shift it is apparent that signals 1-4 are from uracils and signals 5-7 from cytosines. (d) Secondary structure of TASL3 with color-coded pyrimidines that belong to AUX and GCX triplets. For simplicity C11 is included in category GCX although its following nucleotide is not Watson-Crick base paired. (e) Region of a 2D NOESY spectrum corresponding to panel c. H5-H8 _{$i - 1$} correlations between pyrimidines and preceding purines are indicated by red arrows. (f) Statistics of purine H8 chemical shifts of triplets that occur in the sequence. The x-axis matches to the 2D NOESY spectrum shown in panel e.

Table S1. List of all RNA chemical shift entries from the BMRB databank complemented with data of 14 structures found in publications (denoted by their PDB accession codes). Chemical shifts in a Watson-Crick environment were evaluated prior to their use in statistics. Carbon chemical shifts that were evaluated using a set of reference points as described in a previous paper (8) were only used if a set of reference chemical shifts was present and found in the expected ranges. Some datasets could partially be used and some datasets could be used after recalibration. Outliers that are not caused by a systematical mistake were only excluded if they were far off. Proton chemical shifts were excluded, if the chemical shift value differed by more than 0.4 ppm of the median value of the corresponding category. Carbon chemical shifts were excluded if they differed by more than 2 ppm of the median value of the corresponding category.

BMRB number or PDB code ^a	¹³ C chemical shifts ^b	Category of ¹³ C data ^b	Number of nucleotides	Excluded shifts
4120	not used	IV	44	
4125	absent	-	31	
4135	absent	-	9+9	
4175	absent	-	20	
4226	used	I	30	Res 25 C2
4250	absent	-	23	
4253	not used	IV	36	
4346	used	I	33	
4614	absent	-	10	
4745	absent	-	13+12	Strand 2 Res 10 H1'
4750	absent	-	28	
4780	used	Ila	19	Res 1 C8, Res 4 H1'
4816	not used	IV	31	
5007	not used	III	30	
5023	absent	-	13	
5046	absent	-	10	
5170	partially used	Ilb	28	C2, C5 and C1'
5193	absent	-	13	
5256	used	I	17	
5259	used	I	17	
5278	not used	IV	33	
5321	absent	-	29	
5371	used	I	24	
5394	absent	-	9+10	
5395	absent	-	9+10	
5528	absent	-	25	
5530	absent	-	17	
5531	absent	-	17	
5553	not used	IV	31	
5559	not used	IV	20	
5586	absent	-	9	
5587	absent	-	9	
5588	absent	-	9	

BMRB number or PDB code ^a	¹³ C chemical shifts ^b	Category of ¹³ C data ^b	Number of nucleotides	Excluded shifts
5614	absent	-	9	
5632	not used	III	30	
5655	used	I	24	C2', C3' not used, Res22 C1'
5703	not used	IIId	24	
5705	used	I	14	
5773	not used	III	36	
5834 ^c	partially used	I	22	C2, C6 and C8 resonances of Ade and Ura excluded
5852	used	I	23	
5919	partially used	IIb	42	C1' resonances excluded, Res 10 C6
5932	partially used	IIb	15	C1' resonances excluded
5962	used	I	36	Res 1 C1'
5980	absent	-	10	
6042 ^d	absent	-	30	
6062	used	IIa	23	
6076 (1)	used	I	34	
6076 (2)	absent	-	34	
6077 (1)	used	I	34	C5 of Res 11 and 12
6077 (2)	absent	-	34	
6094	not used	IV	101	
6115	absent	-	34	
6239	not used	III	22	
6320	not used	III	32	
6477	not used	IV	47	
6485 ^e	used	I	27	
6509	not used	IV	31	
6543	used	I	45	Res 4 C1'
6562 (1)	not used	IV	18	
6562 (2)	not used	IV	18	
6633	not used	IIId	41	
6652	not used	IV	43	
6756	not used	III	27	
6979	absent	-	10+10	
7090	not used	III	18	
7098	used	I	35	Res 25 C2'
7230	absent	-	9+9	
7403	used (recal.)	IIc	24	
7404	used (recal.)	IIc	23	
7405	used (recal.)	IIc	18	Res 15 C6, Res 15 C1'
10014	absent	-	12	
10018	absent	-	36	
15080	used	I	20	
15081	not used	IV	20	
15157	absent	-	8	
15319	absent	-	17	
15331	absent	-	17	
15342	absent	-	17	
15362	absent	-	17	
15417	used	I	34	

BMRB number or PDB code ^a	¹³ C chemical shifts ^b	Category of ¹³ C data ^b	Number of nucleotides	Excluded shifts
15538	not used	III	29	
15571	used	I	10+10	
15572	used	I	11+10	
15656 (1)	not used	III	29	
15656 (2)	partially used	IIb	29	C5 shifts excluded
15656 (3)	absent	-	29	
15697	absent	-	22	Res 19 H8
15745	not used	IV	22	
15780	used	I	9+8	
15781	used	I	9+8	
15786 (1)	not used	III	37	
15786 (2)	partially used	IIb	37	C2, C5, C6 and C8 resonances excluded
15856	not used	III	29	
15858	not used	IV	7+7	
15859	not used	III	22	
15869	used (recal.)	IIc	30	
15915 (1)	absent	-	12+12	Strand 2 Res 5 H6
15915 (2)	absent	-	12+12	Strand 2 Res 5 H6
17326	used	I	20	
17559	used	I	22	
17560	used	I	21	
17566	used	I	22	
17567	used	I	26	
17568	used	I	30	
17RA	used	I	21	
1AFX	partially used	IIb	12	C2, C5, C6 and C8 resonances excluded
1ATO	absent	-	19	
1C00	absent	-	14	
1EBR	absent	-	30	
1MIS	absent	-	8	
1MWG	absent	-	8	
1PBM	absent	-	6	
1QES	absent	-	30	
1RHT	absent	-	24	
1RNG	not used	IV	12	
1SCL	used (recal.)	IIc	29	
1UUU	used (recal.)	I	19	
1YFV	not used	IIId	8	

^a For entries containing two or more sets of chemical shifts, the individual lists are indicated by numbers in brackets.

^b The ¹³C chemical shifts were evaluated according to Aeschbacher et al. (8). ¹³C data classified in category I and IIa was used, category II b was partially used, category IIc was used after re-referencing and categories IIId, III and IV were excluded.

^c C6 shifts of uridines and C8 shifts of adenosines are shifted by around 2.7 ppm. Since RNA selectively labeled at uridines and adenosines was used, we suspect that the spectrum of the AU labeled RNA was not calibrated properly.

^d Entry contains only coupling constants.

^e The initial entry was corrected 2011, see Aeschbacher et al. (8).

Table S2. 1D statistical values of all different categories of analyzed chemical shifts displayed in Figure 1 and Supplementary Figure 1.

	<i>N</i>	Mean	Std	Sum	Skew	Min	25 th Perc.	Median	75 th Perc.	Max
H8 chemical shifts										
AGX	50	7.142	0.128	357.094	0.422	6.921	7.048	7.123	7.250	7.463
GGX	65	7.214	0.098	468.907	-0.827	6.809	7.162	7.212	7.261	7.464
GGGA	26	7.158	0.092	186.12	-2.295	6.809	7.151	7.168	7.215	7.274
GGG	11	7.209	0.073	79.30	1.179	7.110	7.155	7.199	7.250	7.380
GGC	21	7.274	0.087	152.75	-0.436	7.040	7.244	7.262	7.320	7.464
GGU	7	7.282	0.109	50.97	0.482	7.143	7.192	7.310	7.330	7.466
G^UGX	2	7.229	0.083	14.46	--	7.170	7.170	7.229	7.288	7.288
CGX	43	7.568	0.107	325.436	0.547	7.316	7.510	7.550	7.630	7.910
UGX	33	7.695	0.057	253.926	0.489	7.604	7.643	7.697	7.730	7.831
U^GGX	5	7.943	0.087	39.72	0.510	7.840	7.910	7.910	7.990	8.067
5'GGX	68	7.553	0.076	513.610	-0.036	7.379	7.513	7.546	7.610	7.710
5'GX	87	8.114	0.076	705.960	-1.547	7.790	8.080	8.140	8.169	8.220
A^GUX	6	7.027	0.141	42.16	1.570	6.890	6.927	7.009	7.038	7.290
G^GUX	5	7.18	0.114	35.91	0.809	7.060	7.128	7.130	7.240	7.350
C^GUX	9	7.519	0.105	67.67	-0.682	7.340	7.450	7.550	7.580	7.658
U^GUX	4	7.717	0.050	30.87	1.070	7.671	7.681	7.706	7.754	7.786
5'G^GUX	4	7.484	0.170	29.94	-1.876	7.232	7.383	7.557	7.585	7.590
AAAX	25	7.722	0.107	193.038	-0.026	7.541	7.635	7.713	7.820	7.910
GAX	71	7.768	0.093	551.507	-0.241	7.556	7.683	7.763	7.840	7.940
G^UAX	4	7.727	0.157	30.91	0.068	7.591	7.592	7.717	7.862	7.884
CAX	37	8.023	0.078	296.862	-0.510	7.840	7.971	8.046	8.075	8.154
UAX	16	8.129	0.071	130.064	-0.547	7.995	8.082	8.140	8.175	8.240
U^GAX	2	8.285	0.008	16.57	--	8.279	8.279	8.285	8.290	8.290
5'AX	2	8.327	0.104	16.653	--	8.253	8.253	8.327	8.400	8.400
H6 chemical shifts										
ACX	45	7.462	0.071	335.808	-0.094	7.329	7.414	7.467	7.513	7.590
G^CCX	84	7.638	0.086	641.556	-0.422	7.370	7.580	7.643	7.700	7.810
G^CCA	13	7.657	0.059	99.54	-0.331	7.552	7.630	7.655	7.688	7.746
G^CCG	17	7.565	0.095	128.60	-0.226	7.370	7.500	7.580	7.630	7.720
G^CCC	42	7.639	0.068	320.85	-0.163	7.480	7.580	7.640	7.689	7.760
G^CCU	12	7.714	0.085	92.57	-0.934	7.551	7.683	7.733	7.775	7.810
G^UCX	9	7.662	0.068	68.96	-0.574	7.533	7.640	7.664	7.710	7.760
CCX	56	7.772	0.059	435.240	-0.903	7.570	7.735	7.780	7.810	7.880
UCX	65	7.878	0.054	512.069	-0.401	7.711	7.850	7.882	7.911	8.010
U^GCX	5	7.968	0.031	39.84	-0.437	7.930	7.940	7.980	7.990	8.000
3'CC	66	7.652	0.056	505.052	-0.824	7.490	7.626	7.658	7.692	7.754
5'CX	16	8.101	0.091	129.612	-1.315	7.869	8.070	8.128	8.173	8.194

	<i>N</i>	Mean	Std	Sum	Skew	Min	25 th Perc.	Median	75 th Perc.	Max
AUX	26	7.578	0.075	197.029	-0.270	7.438	7.536	7.571	7.639	7.722
GUX	54	7.761	0.098	419.108	-0.531	7.474	7.713	7.775	7.820	7.957
G ^U UX	3	7.668	0.038	23.01	-1.732	7.625	7.625	7.690	7.690	7.690
CUX	49	7.897	0.079	386.934	-0.700	7.606	7.865	7.900	7.951	8.130
UUX	29	7.979	0.056	231.388	-1.192	7.815	7.954	7.990	8.015	8.080
U ^G UX	4	8.065	0.023	32.26	-0.079	8.040	8.045	8.066	8.085	8.088
5'UX	2	8.085	0.007	16.170	--	8.080	8.080	8.085	8.090	8.090
AU ^G X	4	7.571	0.121	30.28	0.0429	7.460	7.467	7.567	7.675	7.690
GU ^G X	15	7.699	0.084	115.48	0.959	7.600	7.626	7.700	7.730	7.890
CU ^G X	5	7.763	0.133	38.81	0.220	7.590	7.700	7.766	7.808	7.950
UU ^G X	4	7.951	0.052	31.80	1.652	7.910	7.918	7.933	7.984	8.026
H5 chemical shifts										
ACX	45	5.189	0.062	233.482	-1.180	4.973	5.160	5.185	5.222	5.307
G ^C CX	83	5.223	0.080	433.533	0.062	5.010	5.174	5.220	5.276	5.460
G ^C CA	13	5.226	0.064	67.94	0.350	5.131	5.180	5.210	5.276	5.341
G ^C CG	16	5.197	0.087	83.14	-0.525	5.010	5.157	5.200	5.246	5.320
G ^C CC	42	5.223	0.085	219.35	0.474	5.010	5.160	5.225	5.262	5.460
G ^C CU	12	5.259	0.059	63.10	-0.011	5.180	5.207	5.257	5.309	5.360
G ^U CX	9	5.362	0.084	48.26	0.221	5.263	5.290	5.350	5.430	5.475
CCX	56	5.500	0.091	307.975	-0.153	5.240	5.450	5.498	5.553	5.730
UCX	66	5.650	0.065	372.931	0.004	5.478	5.599	5.662	5.694	5.857
U ^G CX	5	5.685	0.032	28.43	0.000	5.640	5.680	5.685	5.690	5.730
3'CC	66	5.476	0.086	361.395	-0.661	5.220	5.437	5.481	5.530	5.710
5'CX	16	5.972	0.116	95.545	-1.974	5.642	5.972	5.990	6.033	6.090
AUX	26	5.042	0.099	131.083	2.586	4.893	4.991	5.033	5.073	5.440
GUX	56	5.088	0.083	284.937	-0.483	4.858	5.033	5.098	5.150	5.272
G ^U UX	3	5.216	0.067	15.65	1.008	5.158	5.158	5.200	5.290	5.290
CUX	49	5.442	0.118	266.667	0.757	5.153	5.401	5.430	5.470	5.824
UUX	30	5.568	0.071	167.033	-1.706	5.345	5.545	5.564	5.620	5.681
U ^G UX	4	5.610	0.035	22.44	0.275	5.570	5.585	5.609	5.636	5.654
5'UX	2	5.825	0.007	11.650	--	5.820	5.820	5.825	5.830	5.830
AU ^G X	4	5.376	0.056	21.50	0.634	5.314	5.342	5.370	5.410	5.450
GU ^G X	14	5.372	0.109	75.20	-1.888	5.070	5.350	5.407	5.430	5.480
CU ^G X	5	5.702	0.086	28.51	0.453	5.610	5.630	5.717	5.730	5.825
UU ^G X	4	5.785	0.069	23.14	0.857	5.714	5.737	5.774	5.833	5.877
H1' chemical shifts										
AGX	50	5.574	0.094	278.678	-0.880	5.270	5.528	5.593	5.638	5.750
G ^G GX	63	5.750	0.054	362.270	-0.632	5.580	5.720	5.750	5.790	5.855
G ^G GA	25	5.741	0.045	143.52	0.306	5.665	5.712	5.737	5.765	5.821
G ^G GG	11	5.769	0.039	63.46	0.542	5.708	5.750	5.754	5.800	5.840

	N	Mean	Std	Sum	Skew	Min	25 th Perc.	Median	75 th Perc.	Max
GGC	21	5.744	0.071	120.62	-0.726	5.580	5.710	5.740	5.803	5.855
GGU	7	5.801	0.064	40.61	1.695	5.729	5.771	5.786	5.810	5.933
G ^U GX	2	5.701	0.010	11.40	--	5.630	5.630	5.701	5.771	5.771
CGX	41	5.700	0.070	233.698	-0.920	5.523	5.665	5.710	5.750	5.807
UGX	32	5.763	0.061	184.408	-1.581	5.530	5.741	5.762	5.801	5.868
U ^G GX	5	5.753	0.054	28.77	1.303	5.700	5.730	5.730	5.765	5.840
5'GGX	67	5.907	0.050	395.763	-1.343	5.720	5.881	5.918	5.944	6.000
5'GX	87	5.808	0.079	505.281	-0.384	5.618	5.77	5.814	5.858	5.979
A ^U UX	6	5.601	0.096	33.61	-0.360	5.483	5.490	5.629	5.655	5.720
GG ^U UX	5	5.786	0.033	28.93	0.484	5.750	5.770	5.772	5.810	5.830
CG ^U UX	9	5.7103	0.073	51.39	-0.309	5.600	5.670	5.730	5.750	5.812
UG ^U UX	4	5.771	0.028	23.08	0.727	5.740	5.753	5.767	5.789	5.808
5'GG ^U UX	4	5.855	0.142	23.42	-1.844	5.645	5.7695	5.912	5.940	5.950
AAX	26	5.865	0.071	152.501	-0.040	5.715	5.810	5.869	5.928	6.020
GAX	72	5.946	0.070	428.090	-0.847	5.685	5.909	5.950	6.002	6.081
G ^U AX	4	5.908	0.029	23.63	-1.711	5.865	5.890	5.919	5.927	5.930
CAX	38	5.896	0.065	224.060	-0.814	5.730	5.860	5.910	5.938	6.008
UAX	17	5.975	0.069	101.575	-0.465	5.842	5.950	5.970	6.022	6.090
U ^G AX	2	5.940	0.001	11.88	--	5.939	5.939	5.940	5.940	5.940
5'AX	2	6.009	0.055	12.018	--	5.970	5.970	6.009	6.048	6.048
ACX	44	5.372	0.068	236.384	-1.578	5.137	5.360	5.386	5.414	5.461
GCX	83	5.526	0.087	458.644	1.013	5.340	5.476	5.520	5.567	5.820
GCA	13	5.486	0.039	71.31	0.993	5.434	5.460	5.480	5.500	5.571
GCG	17	5.515	0.095	93.75	1.178	5.340	5.479	5.500	5.550	5.790
GCC	41	5.535	0.101	226.94	0.879	5.340	5.470	5.520	5.570	5.820
GCU	12	5.554	0.039	66.64	-1.067	5.465	5.532	5.561	5.580	5.600
G ^U CX	9	5.5282	0.146	49.75	0.614	5.278	5.470	5.515	5.540	5.830
CCX	54	5.472	0.056	295.494	-0.273	5.336	5.448	5.470	5.509	5.580
UCX	62	5.565	0.072	345.041	0.843	5.439	5.522	5.570	5.606	5.850
U ^G CX	5	5.619	0.093	28.10	-0.586	5.506	5.530	5.670	5.690	5.700
3'CC	64	5.726	0.068	366.479	-1.179	5.520	5.690	5.743	5.779	5.830
5'CX	16	5.540	0.101	88.643	-1.681	5.231	5.522	5.543	5.583	5.720
AUX	26	5.464	0.071	142.053	-1.346	5.265	5.432	5.484	5.508	5.563
GUX	54	5.561	0.094	300.319	0.619	5.315	5.522	5.567	5.618	5.955
G ^U UX	3	5.502	0.088	16.51	0.086	5.415	5.415	5.500	5.590	5.590
CUX	49	5.534	0.094	271.165	0.759	5.320	5.490	5.538	5.580	5.920
UUX	30	5.617	0.066	168.516	-0.789	5.451	5.581	5.622	5.677	5.709
U ^G UX	4	5.634	0.016	22.54	-1.931	5.610	5.625	5.641	5.643	5.644
5'UX	2	5.560	0.028	11.120	--	5.540	5.540	5.560	5.580	5.580
AU ^G X	4	5.308	0.069	21.23	-0.513	5.221	5.256	5.315	5.360	5.380
GU ^G X	15	5.501	0.094	82.51	-1.234	5.250	5.475	5.495	5.550	5.640

	<i>N</i>	Mean	Std	Sum	Skew	Min	25 th Perc.	Median	75 th Perc.	Max
CU^GX	5	5.629	0.099	28.15	-0.146	5.500	5.580	5.610	5.721	5.734
UU^GX	4	5.570	0.157	22.28	1.243	5.419	5.476	5.535	5.664	5.790
H2 chemical shifts (ppm)										
AAA	3	7.205	0.083	21.616	-0.307	7.120	7.12	7.211	7.285	7.285
AAG	9	7.402	0.056	66.617	0.617	7.330	7.367	7.399	7.414	7.490
AAC	11	7.740	0.113	85.140	-1.275	7.510	7.688	7.787	7.806	7.874
AAU	3	7.782	0.011	23.346	-0.423	7.771	7.771	7.783	7.792	7.792
GAA	21	7.094	0.087	148.975	-1.784	6.830	7.067	7.100	7.147	7.205
GAG	22	7.479	0.046	164.527	-1.215	7.351	7.465	7.486	7.504	7.546
GAC	22	7.802	0.049	171.652	-0.529	7.700	7.760	7.813	7.838	7.875
GAU	8	7.753	0.083	62.021	-0.638	7.626	7.686	7.771	7.818	7.846
CAA	3	6.732	0.038	20.195	-1.042	6.690	6.690	6.741	6.764	6.764
CAG	11	6.999	0.067	76.986	0.419	6.890	6.946	7.010	7.030	7.138
CAC	13	7.395	0.095	96.134	-0.933	7.209	7.367	7.420	7.455	7.520
CAU	11	7.270	0.065	79.970	-0.602	7.140	7.224	7.270	7.317	7.360
UAA	3	6.463	0.055	19.390	-1.668	6.400	6.400	6.490	6.500	6.500
UAG	4	6.800	0.082	27.200	-0.200	6.710	6.730	6.810	6.870	6.870
UAC	6	7.089	0.071	42.534	-0.183	6.984	7.041	7.092	7.140	7.186
UAU	3	7.009	0.122	21.026	0.062	6.888	6.888	7.007	7.131	7.131
C8 chemical shifts										
AGX	22	135.861	0.372	2988.946	1.398	135.40	135.647	135.724	136.00	136.76
GGX	30	136.297	0.262	4088.924	0.670	135.94	136.10	136.256	136.45	136.926
CGX	15	136.227	0.280	2043.412	0.896	135.80	136.012	136.20	136.40	136.853
UGX	19	136.359	0.149	2590.828	0.059	136.00	136.30	136.322	136.40	136.70
5'GGX	31	136.986	0.207	4246.552	1.119	136.64	136.816	136.975	137.084	137.60
5'GX	36	139.151	0.193	5009.440	-0.564	138.70	139.023	139.193	139.30	139.438
AAX	8	139.138	0.357	1113.106	2.575	138.90	138.98	139.00	139.123	140.00
GAX	29	139.534	0.474	4046.493	2.501	138.903	139.33	139.434	139.60	141.40
CAX	17	139.586	0.135	2372.962	0.069	139.40	139.50	139.60	139.70	139.824
UAX	4	139.760	0.273	559.040	0.210	139.44	139.57	139.75	139.95	140.10
C6 chemical shifts (ppm)										
ACX	13	140.646	0.452	1828.401	0.288	139.98	140.20	140.579	140.979	141.524
GCX	34	141.081	0.491	4796.745	0.073	140.10	140.841	141.033	141.346	142.02
CCX	30	141.482	0.363	4244.465	1.024	140.799	141.36	141.50	141.6	142.772
UCX	22	141.872	0.442	3121.175	1.694	141.20	141.704	141.826	142.031	143.379
3'CC	30	141.817	0.266	4254.496	-1.293	140.90	141.70	141.857	141.972	142.3
5'CX	4	143.215	0.535	572.860	1.741	142.80	142.905	143.03	143.525	144.00
AUX	11	141.374	0.4729	1555.112	0.542	140.80	140.93	141.50	141.60	142.323
GUX	20	141.759	0.481	2835.188	0.513	141.10	141.42	141.75	142.00	142.705
CUX	17	142.050	0.320	2414.850	-0.162	141.50	141.994	142.08	142.195	142.612
UUX	13	142.402	0.254	1851.227	-0.298	142.00	142.30	142.40	142.60	142.788

	<i>N</i>	Mean	Std	Sum	Skew	Min	25 th Perc.	Median	75 th Perc.	Max
C5 chemical shifts (ppm)										
ACX	12	97.347	0.337	1168.161	-0.576	96.60	97.169	97.345	97.58	97.9
G<u>C</u>X	32	97.232	0.256	3111.428	0.301	96.80	97.047	97.20	97.434	97.822
C<u>C</u>X	22	97.687	0.361	2149.111	0.946	97.14	97.41	97.67	97.88	98.681
U<u>C</u>X	31	97.641	0.331	3026.875	0.402	97.037	97.39	97.59	97.91	98.4
3'<u>C</u>C	25	98.155	0.287	2453.886	-0.222	97.51	97.953	98.197	98.311	98.701
5'<u>C</u>X	4	98.798	0.134	395.190	-1.765	98.60	98.72	98.845	98.875	98.9
A<u>U</u>X	9	102.885	0.273	925.962	-0.073	102.50	102.622	102.90	103.182	103.2
G<u>U</u>X	20	102.856	0.308	2057.117	0.340	102.30	102.618	102.792	103.10	103.4
C<u>U</u>X	19	103.389	0.214	1964.384	0.087	103.00	103.292	103.381	103.526	103.82
U<u>U</u>X	15	103.274	0.345	1549.105	0.103	102.70	102.948	103.286	103.463	103.9
C1' chemical shifts (ppm)										
A<u>G</u>X	14	92.731	0.302	1298.237	0.699	92.30	92.59	92.60	93.00	93.304
G<u>G</u>X	18	93.077	0.266	1675.378	0.594	92.713	92.87	93.011	93.21	93.60
C<u>G</u>X	11	93.045	0.243	1023.493	0.377	92.71	92.92	93.00	93.23	93.50
U<u>G</u>X	16	92.678	0.327	1482.848	-0.148	91.957	92.470	92.713	92.895	93.35
5'<u>G</u>G<u>X</u>	25	92.925	0.341	2323.137	-2.413	91.599	92.82	92.991	93.12	93.51
5'<u>G</u>X	31	91.718	0.699	2843.245	-0.085	90.224	91.21	91.685	92.31	93.023
A<u>A</u>X	7	92.688	0.206	648.817	-0.816	92.40	92.411	92.81	92.851	92.87
G<u>A</u>X	32	93.075	0.167	2978.412	-0.808	92.70	92.98	93.098	93.207	93.3
C<u>A</u>X	16	93.060	0.128	1488.957	0.126	92.90	92.928	93.10	93.147	93.28
U<u>A</u>X	5	92.832	0.186	464.160	0.414	92.66	92.67	92.77	93.01	93.05
A<u>C</u>X	16	93.895	0.388	1502.321	1.185	93.32	93.613	93.925	94.10	94.959
G<u>C</u>X	28	93.998	0.244	2631.955	-0.158	93.50	93.814	94.00	94.198	94.48
C<u>C</u>X	13	94.213	0.173	1224.763	-0.344	93.847	94.105	94.20	94.342	94.509
U<u>C</u>X	22	94.126	0.210	2070.762	0.580	93.796	94.00	94.085	94.30	94.544
3'<u>C</u>C	23	92.962	0.146	2138.128	0.271	92.713	92.86	92.942	93.10	93.24
A<u>U</u>X	9	93.181	0.214	838.626	1.012	92.99	93.03	93.09	93.34	93.577
G<u>U</u>X	20	93.867	0.278	1877.343	0.696	93.52	93.630	93.775	94.091	94.50
C<u>U</u>X	17	93.828	0.206	1595.078	0.275	93.50	93.70	93.83	93.986	94.239
U<u>U</u>X	13	93.588	0.372	1216.641	-1.990	92.582	93.565	93.739	93.77	93.968
C2 chemical shifts (ppm)										
A<u>A</u>A	1	152.895	--	152.895	--	152.895	152.895	152.895	152.895	152.895
A<u>A</u>G	4	153.032	0.046	612.126	1.833	153.00	153.004	153.013	153.060	153.10
A<u>A</u>C	3	153.813	0.186	461.440	-1.642	153.60	153.60	153.90	153.94	153.94
A<u>A</u>U	3	153.867	0.058	461.600	-1.732	153.80	153.80	153.90	153.90	153.90
G<u>A</u>A	9	152.879	0.169	1375.910	-1.522	152.50	152.812	152.934	153.00	153.037
G<u>A</u>G	9	153.428	0.110	1380.849	-0.154	153.293	153.30	153.454	153.53	153.556
G<u>A</u>C	10	154.196	0.159	1541.959	-1.008	153.877	154.10	154.246	154.30	154.40
G<u>A</u>U	7	153.947	0.345	1077.631	2.063	153.70	153.70	153.90	154.00	154.686

	<i>N</i>	Mean	Std	Sum	Skew	Min	25 th Perc.	Median	75 th Perc.	Max
CAA	2	152.250	0.212	304.5000	--	152.10	152.10	152.25	152.40	152.40
CAG	4	152.428	0.101	609.713	-0.592	152.30	152.35	152.442	152.507	152.53
CAC	5	153.406	0.147	767.029	-2.089	153.147	153.437	153.46	153.485	153.50
CAU	8	153.131	0.080	1225.050	-0.022	153.00	153.10	153.10	153.20	153.25
UAA	1	152.300	--	152.300	--	152.30	152.30	152.30	152.30	152.30
UAG	0	--	--	--	--	--	--	--	--	--
UAC	4	153.308	0.137	613.230	1.250	153.18	153.215	153.275	153.40	153.50
UAU	0	--	--	--	--	--	--	--	--	--

Table S3. Statistics of the intraresidual H8-C8 2D chemical shifts correlations of Figure 2.

H8-C8 chemical shift correlations												
	AAX		CAX		GAX		UAX					
	H8	C8	H8	C8	H8	C8	H8	C8				
Mean	7.6886	139.14	8.0153	139.59	7.7626	139.53	8.167	139.76				
Std	0.1260	0.3565	0.0659	0.1350	0.0957	0.4739	0.0786	0.2728				
Corr	0.49428		0.5142		0.68402		0.46485					
N	8		17		29		4					
	AGX		CGX		GGX		UGX		5'GGX		5'GX	
	H8	C8	H8	C8	H8	C8	H8	C8	H8	C8	H8	C8
Mean	7.1566	135.86	7.5684	136.23	7.2099	136.30	7.6776	136.36	7.5737	136.99	8.1313	139.15
Std	0.1343	0.3720	0.0833	0.2797	0.0863	0.2621	0.0530	0.1487	0.0731	0.2071	0.0450	0.1931
Corr	0.3456		0.88163		0.33701		-0.02944		0.29544		0.29131	
N	22		15		30		19		31		36	

Table S4. Statistics of the intraresidual H5-H6 2D chemical shifts correlations of Figure 2.

H5-H6 chemical shift correlations												
	ACX		CCX		GCX		UCX		3'CC		5'CX	
	H5	H6	H5	H6	H5	H6	H5	H6	H5	H6	H5	H6
Mean	5.1934	7.4603	5.4996	7.7721	5.2233	7.6383	5.6510	7.8778	5.4759	7.6507	5.9716	8.1008
Std	0.0529	0.0701	0.0905	0.0586	0.0799	0.0868	0.0656	0.0545	0.0866	0.0549	0.1162	0.0915
Corr	0.20926		0.25932		0.5595		0.63389		0.58006		0.34914	
N	44		56		83		65		65		16	
	AUX		CUX		GUX		UUX		5'UX			
	H5	H6	H5	H6	H5	H6	H5	H6	H5	H6		
Mean	5.0417	7.5780	5.4422	7.8966	5.0888	7.7613	5.5747	7.9789	5.825	8.085		
Std	0.0991	0.0751	0.1176	0.0786	0.0749	0.0978	0.0606	0.0562	0.0071	0.0071		
Corr	0.36128		0.46968		0.69167		0.53948		NA			
N	26		49		54		29		2			

Table S5. Input peak lists for automatic resonance assignment with FLYA

Spectrum	Expected peaks ^a	Measured peaks ^b	Assigned ^c	Complete ^d
FZL4:				
NOESY	560	270	74.1%	42.7%
TOCSY	18	18	100.0%	100.0%
¹ H- ¹³ C HSQC	56	60	88.3%	98.2%
TASL1:				
NOESY	657	431	61.3%	51.6%
TOCSY	24	24	100.0%	100.0%
¹ H- ¹³ C HSQC	61	60	98.3%	100.0%
SL23:				
NOESY	600	297	69.7%	50.0%
TOCSY	28	28	100.0%	100.0%
¹ H- ¹³ C HSQC	62	59	91.5%	100.0%
FZL2:				
NOESY	500	215	76.7%	47.4%
TOCSY	18	18	100.0%	100.0%
¹ H- ¹³ C HSQC	58	55	96.4%	98.3%
ELAVL1 siRNA:				
NOESY	1419	616	82.1%	57.9%
TOCSY	46	45	100.0%	100.0%
¹ H- ¹³ C HSQC	121	109	93.6%	98.4%

^a Expected peaks: Number of peaks expected by FLYA for H2, H5, H6, H8, H1', and their corresponding carbon resonances based on magnetization transfer rules for TOCSY and ¹H-¹³C HSQC, or, for the NOESY spectrum, consistently short distances in a bundle of randomized conformers (see Methods). Only the uracil H5-H6 correlations were considered for the TOCSY spectrum.

^b Measured peaks: Number of peaks picked in the measured spectrum.

^c Assigned: Percentage of measured peaks that were assigned by FLYA. The theoretical maximum of 100% corresponds to having all measured peaks assigned. Note that several expected peaks can be mapped to the same measured peak, i.e. assignments of measured peaks can be unambiguous or ambiguous. Remaining unassigned measured peaks are likely to be artifacts.

^d Complete: Percentage of expected peaks that FLYA mapped to a measured peak. The theoretical maximum of 100% corresponds to the situation that the measured peak list contains all expected peaks. Each expected peak can be mapped to at most one measured peak. Remaining expected peaks correspond to missing peaks in the measured peak list.

Table S6. FLYA chemical shift assignments for FZL4

For each assigned atom the table contains the following information: ‘Ref’, manually determined reference chemical shift (ppm). ‘Shift’, consolidated chemical shift determined by FLYA (ppm). ‘Dev’, deviation between the consolidated chemical shift determined by FLYA and the reference chemical shift (ppm). ‘Extent’, number of individual runs of the FLYA algorithm in which the atom was assigned. ‘inside’, percentage of the chemical shift values from the individual FLYA runs that deviate by less than the tolerance of 0.02 ppm for ¹H and 0.3 ppm for ¹³C from the consolidated chemical shift value. Assignments for which this percentage is at least 80% are labelled as ‘strong’. ‘inref’, percentage of the chemical shift values from the individual FLYA runs that agree with the reference chemical shift value within the tolerance. Atoms for which the consolidated and reference chemical shifts agree within the tolerance are labelled with ‘=’ (green). Disagreement between the two shifts is indicated by ‘!’ (magenta). For erroneous consolidated shifts the sequentially closest atom with a reference chemical shift within the tolerance is indicated in parenthesis if it is located in the same residue (no residue number given) or in the immediately preceding or following residue (residue number given). Atoms for which our chemical shift statistics could be used are highlighted in cyan.

Atom	Residue	Ref	Shift	Dev	Extent	inside	inref		H6	RCYT	13	7.724	7.724	0.000	50.0	100.0	100.0	strong=
C1'	RGUA	1	91.579	91.579	0.000	50.0	100.0	strong=	C1'	RGUA	14	91.893	92.446	-0.553	50.0	42.0	2.0	!
H1'	RGUA	1	5.791	5.793	-0.002	50.0	100.0	strong=	H1'	RGUA	14	5.392	5.672	-0.280	50.0	73.9	2.0	!
C8	RGUA	1	139.129	139.129	0.000	50.0	100.0	strong=	C8	RGUA	14	138.097	139.709	-1.612	49.0	93.9	0.0	strong! (C8 15)
H8	RGUA	1	8.122	8.120	0.002	50.0	100.0	strong=	H8	RGUA	14	7.850	7.922	-0.072	50.0	76.0	0.0	!
C1'	RGUA	2	92.817	92.817	0.000	50.0	100.0	strong=	C1'	RADE	15	93.197	93.197	0.000	50.0	100.0	100.0	strong=
H1'	RGUA	2	5.916	5.917	-0.001	50.0	100.0	strong=	H1'	RADE	15	5.949	5.950	-0.001	50.0	100.0	100.0	strong=
C8	RGUA	2	136.839	136.839	0.000	50.0	100.0	strong=	C2	RADE	15	153.724	153.724	0.000	50.0	100.0	100.0	strong=
H8	RGUA	2	7.514	7.516	-0.002	50.0	100.0	strong=	H2	RADE	15	7.627	7.628	-0.001	50.0	99.9	100.0	strong=
C1'	RGUA	3	93.290	93.290	0.000	50.0	100.0	strong=	C8	RADE	15	139.895	139.895	0.000	50.0	100.0	100.0	strong=
H1'	RGUA	3	5.769	5.776	-0.007	50.0	97.3	strong=	H8	RADE	15	7.833	7.835	-0.002	50.0	100.0	100.0	strong=
C8	RGUA	3	136.166	136.166	0.000	50.0	100.0	strong=	C1'	URA	16	93.063	93.063	0.000	50.0	100.0	100.0	strong=
H8	RGUA	3	7.190	7.191	-0.001	50.0	100.0	strong=	H1'	URA	16	5.406	5.402	0.004	50.0	99.8	100.0	strong=
C1'	URA	4	93.766	93.766	0.000	50.0	100.0	strong=	C5	URA	16	103.152	103.152	0.000	50.0	100.0	100.0	strong=
H1'	URA	4	5.556	5.558	-0.002	50.0	100.0	strong=	H5	URA	16	4.977	4.979	-0.002	50.0	100.0	100.0	strong=
C5	URA	4	102.610	102.610	0.000	50.0	100.0	strong=	C6	URA	16	140.727	140.711	0.016	50.0	99.8	100.0	strong=
H5	URA	4	5.068	5.070	-0.002	50.0	100.0	strong=	H6	URA	16	7.444	7.445	-0.001	50.0	100.0	100.0	strong=
C6	URA	4	141.696	141.696	0.000	50.0	100.0	strong=	C1'	RGUA	17	92.392	92.392	0.000	50.0	100.0	100.0	strong=
H6	URA	4	7.713	7.712	0.001	50.0	100.0	strong=	H1'	RGUA	17	5.699	5.700	-0.001	50.0	100.0	100.0	strong=
C1'	RCYT	5	93.967	93.967	0.000	50.0	100.0	strong=	C8	RGUA	17	136.482	136.482	0.000	50.0	100.0	100.0	strong=
H1'	RCYT	5	5.519	5.520	-0.001	50.0	99.9	strong=	H8	RGUA	17	7.624	7.625	-0.001	50.0	100.0	100.0	strong=
C5	RCYT	5	98.203	98.203	0.000	50.0	100.0	strong=	C1'	RADE	18	92.765	92.765	0.000	50.0	100.0	100.0	strong=
H5	RCYT	5	5.673	5.674	-0.001	50.0	100.0	strong=	H1'	RADE	18	5.885	5.885	0.000	50.0	100.0	100.0	strong=
C6	RCYT	5	141.386	141.386	0.000	50.0	100.0	strong=	C2	RADE	18	154.103	154.103	0.000	50.0	100.0	100.0	strong=
H6	RCYT	5	7.801	7.799	0.002	50.0	100.0	strong=	H2	RADE	18	7.699	7.710	-0.011	50.0	99.9	100.0	strong=
C1'	RADE	6	93.098	93.098	0.000	50.0	100.0	strong=	C8	RADE	18	139.596	139.596	0.000	50.0	100.0	100.0	strong=
H1'	RADE	6	5.834	5.833	0.001	50.0	100.0	strong=	H8	RADE	18	7.870	7.867	0.003	50.0	100.0	100.0	strong=
C2	RADE	6	153.072	153.072	0.000	50.0	100.0	strong=	C1'	RCYT	19	93.733	93.733	0.000	50.0	100.0	100.0	strong=
H2	RADE	6	7.226	7.225	0.001	50.0	100.0	strong=	H1'	RCYT	19	5.350	5.351	-0.001	50.0	100.0	100.0	strong=
C8	RADE	6	139.749	139.749	0.000	50.0	100.0	strong=	C5	RCYT	19	97.193	97.193	0.000	50.0	100.0	100.0	strong=
H8	RADE	6	8.045	8.046	-0.001	50.0	100.0	strong=	H5	RCYT	19	5.158	5.158	0.000	50.0	100.0	100.0	strong=
C1'	URA	7	93.362	93.362	0.000	50.0	100.0	strong=	C6	RCYT	19	140.684	140.699	-0.015	50.0	99.8	100.0	strong=
H1'	URA	7	5.422	5.419	0.003	50.0	100.0	strong=	H6	RCYT	19	7.453	7.452	0.001	50.0	100.0	100.0	strong=
C5	URA	7	102.493	102.493	0.000	50.0	100.0	strong=	C1'	RCYT	20	94.315	94.315	0.000	50.0	100.0	100.0	strong=
H5	URA	7	4.932	4.933	-0.001	50.0	100.0	strong=	H1'	RCYT	20	5.484	5.485	-0.001	50.0	99.2	100.0	strong=
C6	URA	7	141.565	141.565	0.000	50.0	100.0	strong=	C5	RCYT	20	97.637	97.637	0.000	50.0	100.0	100.0	strong=
H6	URA	7	7.596	7.598	-0.002	50.0	100.0	strong=	H5	RCYT	20	5.405	5.406	-0.001	50.0	99.8	100.0	strong=
C1'	RCYT	8	93.823	93.823	0.000	50.0	100.0	strong=	C6	RCYT	20	141.510	141.510	0.000	50.0	100.0	100.0	strong=
H1'	RCYT	8	5.460	5.461	-0.001	50.0	100.0	strong=	H6	RCYT	20	7.731	7.730	0.001	50.0	100.0	100.0	strong=
C5	RCYT	8	97.632	97.632	0.000	50.0	100.0	strong=	C1'	RCYT	21	92.825	92.825	0.000	50.0	100.0	100.0	strong=
H5	RCYT	8	5.375	5.381	-0.006	50.0	99.9	strong=	H1'	RCYT	21	5.742	5.748	-0.006	50.0	100.0	100.0	strong=
C6	RCYT	8	141.058	141.058	0.000	50.0	100.0	strong=	C5	RCYT	21	98.102	98.102	0.000	50.0	100.0	100.0	strong=
H6	RCYT	8	7.559	7.561	-0.002	50.0	100.0	strong=	H5	RCYT	21	5.465	5.470	-0.005	50.0	99.7	100.0	strong=
C1'	RADE	9	92.484	92.470	0.014	50.0	55.8	=	C6	RCYT	21	141.876	141.836	0.040	50.0	98.0	98.0	strong=
H1'	RADE	9	5.670	5.666	0.004	50.0	57.4	=	H6	RCYT	21	7.649	7.648	0.001	50.0	99.5	100.0	strong=
C2	RADE	9	154.507	154.534	-0.027	50.0	98.0	strong=										
H2	RADE	9	7.254	7.254	0.000	50.0	98.0	strong=										
C8	RADE	9	139.655	141.836	-2.181	49.0	51.0	!										
H8	RADE	9	7.922	8.229	-0.307	50.0	50.0	36.0	!									
C1'	RGUA	10	89.976	89.935	0.041	50.0	58.0	=										
H1'	RGUA	10	5.102	5.105	-0.003	50.0	58.0	58.0	!									
C8	RGUA	10	139.909	138.097	1.812	49.0	67.3	28.6	!									
H8	RGUA	10	7.289	7.852	-0.563	50.0	70.0	28.0	!	(H8 11)								
C1'	RGUA	11	89.284	91.861	-2.577	50.0	54.0	2.0	!									
H1'	RGUA	11	5.660	5.392	0.268	50.0	54.0	22.0	!									
C8	RGUA	11	140.238	138.097	2.141	21.0	61.9	9.5	!									
H8	RGUA	11	7.853	7.257	0.596	50.0	59.2	30.0	!									
C1'	RADE	12	90.619	90.619	0.000	50.0	100.0	strong=										
H1'	RADE	12	5.828	5.828	0.000	50.0	99.9	strong=										
C2	RADE	12	155.084	155.005	0.079	50.0	79.5	76.0	=									
H2	RADE	12	7.923	7.905	0.018	50.0	75.7	76.0	=									
C8	RADE	12	141.834	141.836	-0.002	49.0	95.9	95.9	strong=									
H8	RADE	12	8.228	8.229	-0.001	50.0	96.0	96.0	strong=									
C1'	RCYT	13	91.890	91.918	-0.028	50.0	93.9	94.0	strong=									
H1'	RCYT	13	5.857	5.859	-0.002	50.0	93.9	94.0	strong=									
C5	RCYT	13	98.627	98.627	0.000	50.0	100.0	strong=										
H5	RCYT	13	5.753	5.754	-0.001	50.0	100.0	strong=										
C6	RCYT	13	143.158	143.158	0.000	50.0	100.0	strong=										

Table S9. FLYA chemical shift assignments for FZL2

For each assigned atom the table contains the following information: ‘Ref’, manually determined reference chemical shift (ppm). ‘Shift’, consolidated chemical shift determined by FLYA (ppm). ‘Dev’, deviation between the consolidated chemical shift determined by FLYA and the reference chemical shift (ppm). ‘Extent’, number of individual runs of the FLYA algorithm in which the atom was assigned. ‘inside’, percentage of the chemical shift values from the individual FLYA runs that deviate by less than the tolerance of 0.02 ppm for ¹H and 0.3 ppm for ¹³C from the consolidated chemical shift value. Assignments for which this percentage is at least 80% are labelled as ‘strong’. ‘inref’, percentage of the chemical shift values from the individual FLYA runs that agree with the reference chemical shift value within the tolerance. Atoms for which the consolidated and reference chemical shifts agree within the tolerance are labelled with ‘=’ (green). Disagreement between the two shifts is indicated by ‘!’ (magenta). For erroneous consolidated shifts the sequentially closest atom with a reference chemical shift within the tolerance is indicated in parenthesis if it is located in the same residue (no residue number given) or in the immediately preceding or following residue (residue number given). Atoms for which our chemical shift statistics could be used are highlighted in cyan.

Atom	Residue	Ref	Shift	Dev	Extent	inside	inref		C8	RADE	13	141.486	141.510	-0.024	50.0	88.0	88.0	strong=	
C1	RGIA 1	90.739	90.739	0.000	50.0	98.0	98.0	strong=	H8	RADE 13	8.279	8.281	-0.002	50.0	88.0	88.0	strong=		
H1	RGIA 1	5.786	5.787	-0.001	50.0	99.6	100.0	strong=	C1'	RCYT 14	91.912	91.707	-0.205	48.0	80.8	43.8	strong=		
C8	RGIA 1	139.212	139.188	0.024	50.0	100.0	100.0	strong=	H1'	RCYT 14	5.852	5.867	-0.015	50.0	91.6	92.0	strong=		
H8	RGIA 1	8.155	8.155	0.000	50.0	99.7	100.0	strong=	C5	RCYT 14	98.111	98.059	0.052	50.0	94.8	94.0	strong=		
C1'	RGIA 2	92.357	92.432	-0.075	49.0	64.6	55.1	=	H5	RCYT 14	5.625	5.625	0.000	50.0	93.9	94.0	strong=		
H1'	RGIA 2	5.934	5.936	-0.002	50.0	54.0	54.0	=	C6	RCYT 14	142.743	142.743	0.000	50.0	94.0	94.0	strong=		
C8	RGIA 2	137.127	137.160	-0.033	50.0	98.0	98.0	strong=	H6	RCYT 14	7.734	7.733	0.001	50.0	94.0	94.0	strong=		
H8	RGIA 2	7.657	7.664	-0.007	49.0	92.1	98.0	strong=	C1'	RGIA 15	92.867	92.861	0.006	47.0	39.0	36.2	=		
C1'	RGIA 2	93.419	93.383	0.036	49.0	80.2	81.6	strong=	H1'	RGIA 15	5.667	5.498	-0.169	50.0	38.0	26.0	!		
H1'	RGIA 3	5.808	5.802	0.006	50.0	83.7	84.0	strong=	C8	RGIA 15	138.109	139.706	-1.597	49.0	63.9	14.3	!		
C8	RGIA 3	136.822	136.856	-0.034	50.0	48.0	48.0	=	H8	RGIA 15	7.907	7.853	0.054	50.0	60.2	22.0	!		
H8	RGIA 3	7.289	7.293	-0.004	50.0	85.9	86.0	strong=	C1'	RADE 16	93.085	93.070	0.015	50.0	100.0	100.0	strong=		
C1'	RCYT 4	93.843	93.843	0.000	48.0	89.6	89.6	strong=	H1	RADE 16	5.933	5.935	-0.002	50.0	100.0	100.0	strong=		
H1'	RCYT 4	5.523	5.527	-0.006	50.0	86.0	86.0	strong=	C2	RADE 16	153.698	153.698	0.000	49.0	95.9	95.9	strong=		
C5	RCYT 4	97.400	97.421	-0.021	50.0	92.0	94.0	strong=	H2	RADE 16	7.641	7.635	0.006	50.0	93.6	94.0	strong=		
H5	RCYT 4	5.348	5.348	0.000	50.0	86.0	86.0	strong=	C8	RADE 16	139.743	139.778	-0.035	50.0	100.0	100.0	strong=		
C6	RCYT 4	140.546	140.513	0.033	50.0	86.8	86.0	strong=	H8	RADE 16	7.812	7.817	-0.005	50.0	100.0	100.0	strong=		
H6	RCYT 4	7.672	7.676	-0.004	50.0	86.0	86.0	strong=	C1'	URA 17	93.028	93.044	-0.016	50.0	100.0	100.0	strong=		
C1'	RCYT 5	93.950	93.880	0.070	50.0	100.0	100.0	strong=	H1'	URA 17	5.432	5.430	0.002	50.0	100.0	100.0	strong=		
H1'	RCYT 5	5.458	5.462	-0.004	50.0	100.0	100.0	strong=	C5	URA 17	103.075	103.075	0.000	50.0	100.0	100.0	strong=		
C5	RCYT 5	98.113	98.111	0.002	50.0	100.0	100.0	strong=	H5	URA 17	4.994	4.996	-0.002	50.0	100.0	100.0	strong=		
H5	RCYT 5	5.532	5.531	0.001	50.0	100.0	100.0	strong=	C6	URA 17	140.817	140.817	0.000	50.0	100.0	100.0	strong=		
C6	RCYT 5	140.954	140.954	0.000	50.0	100.0	100.0	strong=	H6	URA 17	7.467	7.470	-0.003	50.0	100.0	100.0	strong=		
H6	RCYT 5	7.728	7.729	-0.001	50.0	100.0	100.0	strong=	C1'	RGIA 18	92.463	92.474	-0.011	50.0	95.5	100.0	strong=		
C1'	RADE 6	93.151	93.132	0.019	50.0	100.0	100.0	strong=	H1'	RGIA 18	5.764	5.765	-0.001	50.0	100.0	100.0	strong=		
H1'	RADE 6	5.901	5.902	-0.001	50.0	100.0	100.0	strong=	C8	RGIA 18	136.284	136.288	-0.004	50.0	100.0	100.0	strong=		
C2	RADE 6	153.190	153.190	0.000	50.0	100.0	100.0	strong=	H8	RGIA 18	7.617	7.622	-0.005	50.0	99.9	100.0	strong=		
H2	RADE 6	7.303	7.303	0.000	50.0	100.0	100.0	strong=	C1'	RGIA 19	92.867	92.867	0.000	50.0	100.0	100.0	strong=		
C8	RADE 6	139.736	139.750	-0.014	50.0	100.0	100.0	strong=	H1'	RGIA 19	5.666	5.666	0.000	50.0	99.9	100.0	strong=		
H8	RADE 6	8.061	8.066	-0.005	50.0	100.0	100.0	strong=	C8	RGIA 19	135.978	135.982	-0.004	50.0	100.0	100.0	strong=		
C1'	URA 7	93.275	93.388	-0.113	50.0	100.0	100.0	strong=	H8	RGIA 19	7.214	7.216	-0.002	50.0	100.0	100.0	strong=		
H1'	URA 7	5.462	5.462	0.000	50.0	100.0	100.0	strong=	C1'	URA 20	93.764	92.327	1.437	50.0	50.6	14.0	!		
C5	URA 7	102.615	102.615	0.000	50.0	100.0	100.0	strong=	H1'	URA 20	5.485	5.935	-0.450	50.0	50.0	46.0	!		
H5	URA 7	4.967	4.969	-0.002	50.0	100.0	100.0	strong=	C5	URA 20	104.078	104.078	0.000	50.0	100.0	100.0	strong=		
C6	URA 7	141.523	141.366	0.157	50.0	96.3	90.0	strong=	H5	URA 20	5.386	5.386	0.000	50.0	100.0	100.0	strong=		
H6	URA 7	7.633	7.635	-0.002	50.0	99.7	100.0	strong=	C6	URA 20	140.628	140.546	0.082	50.0	100.0	100.0	strong=		
C1'	RCYT 8	93.707	93.773	-0.066	47.0	94.8	100.0	strong=	H6	URA 20	7.690	7.688	0.002	50.0	100.0	100.0	strong=		
H1'	RCYT 8	5.495	5.495	0.000	50.0	68.2	64.0	=	C1'	RCYT 21	93.947	93.740	0.207	46.0	47.4	45.7	=		
C5	RCYT 8	97.630	97.604	0.026	50.0	96.8	100.0	strong=	H1'	RCYT 21	5.636	5.490	0.146	50.0	49.9	12.0	!		
H5	RCYT 8	5.419	5.420	-0.001	50.0	86.0	86.0	strong=	C5	RCYT 21	97.593	97.593	0.000	50.0	100.0	100.0	strong=		
C6	RCYT 8	141.180	141.086	0.094	50.0	85.7	86.0	strong=	H5	RCYT 21	5.672	5.670	0.002	50.0	99.7	100.0	strong=		
H6	RCYT 8	7.624	7.634	-0.010	50.0	85.9	86.0	strong=	C6	RCYT 21	142.541	142.541	0.000	50.0	100.0	100.0	strong=		
C1'	RADE 9	92.727	91.528	1.199	50.0	44.0	28.0	!	H6	RCYT 21	7.944	7.943	0.001	50.0	99.9	100.0	strong=		
H1'	RADE 9	5.641	5.497	1.144	50.0	50.0	36.0	!	C1'	RCYT 22	92.789	92.748	0.041	50.0	99.9	100.0	strong=		
C2	RADE 9	153.570	153.582	-0.012	49.0	87.1	87.8	strong=	H1'	RCYT 22	5.752	5.752	0.000	50.0	82.0	82.0	strong=		
H2	RADE 9	6.955	6.955	0.000	50.0	78.0	78.0	=	C5	RCYT 22	98.176	98.166	0.010	50.0	100.0	100.0	strong=		
C8	RADE 9	139.728	139.921	-0.193	50.0	82.4	82.0	strong=	H5	RCYT 22	5.560	5.560	0.000	50.0	100.0	100.0	strong=		
H8	RADE 9	7.855	7.792	0.063	50.0	62.1	20.0	!	C6	RCYT 22	141.808	141.779	0.029	50.0	100.0	100.0	strong=		
C1'	RADE 10	91.522	91.678	-0.156	50.0	80.5	56.0	strong=	H6	RCYT 22	7.656	7.656	0.000	50.0	99.6	100.0	strong=		
H1'	RADE 10	5.498	5.866	-0.368	50.0	75.8	22.0	!											
C2	RADE 10	155.299	155.214	0.085	50.0	73.0	76.0	=											
H2	RADE 10	7.905	7.908	-0.003	50.0	74.9	76.0	=											
C8	RADE 10	139.954	140.095	-0.141	50.0	83.8	88.0	strong=											
H8	RADE 10	7.794	7.912	-0.118	50.0	67.6	22.0	!											
C1'	RGIA 11	89.516	89.516	0.000	50.0	76.0	76.0	=											
H1'	RGIA 11	5.141	5.139	0.002	50.0	76.0	76.0	=											
C8	RGIA 11	140.716	140.666	0.050	50.0	50.5	50.0	=											
H8	RGIA 11	7.289	7.288	0.001	50.0	80.0	80.0	=											
C1'	RGIA 12	89.330	92.882	-3.552	49.0	55.1	24.5	!											
H1'	RGIA 12	5.783	5.665	0.118	50.0	56.2	0.0	!											
C8	RGIA 12	140.144	138.116	2.028	50.0	58.0	20.0	!											
H8	RGIA 12	7.931	7.906	0.025	50.0	61.6	2.0	!											
C1'	RADE 13	91.532	92.723	-1.191	49.0	55.1	6.1	!											
H1'	RADE 13	5.871	5.629	0.242	50.0	83.3	4.0	strong!											
C2	RADE 13	155.022	155.109	-0.087	49.0	91.5	95.9	strong=											
H2	RADE 13	7.918	7.907	0.011	50.0	90.6	94.0	strong=											

The .wv file was edited to exclude the G-U wobble:
 watsoncrick strand1=1-2 strand2=21-22
 watsoncrick strand1=4-8 strand2=15-19

The program CHES2FLYA uses as input a file with secondary structure information in the .ct file format. This can be created manually or e.g. on "www.rnasoft.ca". ches2flya creates then a .prot file containing chemical shift prediction of all nucleotides of RNAs. Additionally an angle restraint file (.aco file), a sequence file (.seq file) are prepared. These files are then used as an input for FLYA automated assignment algorithm within CYANA (P.Guntert). ches2flya requires the statistics file Statfile.tab.

Usage: ches2flya [options]

```
#####
-h help
-c prediction interval defined by the percentage of resonances in it [1-11]
  (default 2) 1: 50%, 2: 60%, 3: 70%, 4: 80, 5: 90%, 6: 95%
  7: 98%, 8: 99%, 9: 99.5%, 10: 99.8%, 11: 99.9%
-f file with statistics table (default Statfile.tab)
-s file with RNA sequence and secondary structure (default rnass.ct)
  The file has to be in the .ct format and can be created e.g. on http://www.rnasoft.ca/ .
#####
```

```
./ches2flya -h #--> Help appears
./ches2flya #--> Creates 60% prediction intervals for each shift of the secondary structure file 'rnass.ct' (Standard input file)
./ches2flya -s 'input.ct' #--> Creates 60% prediction intervals for each shift of the secondary structure file 'input.ct'
./ches2flya -s 'input.ct' -c 3 #--> Creates 70% prediction intervals for each shift of the secondary structure file 'input.ct'
```

```
#####
# Given example:
# RNA stemloop FZL4
#####
cd CHES2FLYA/demo
./ches2flya -s FZL4.ct -f ../Statfile.tab #--> generates: cyana.aco, cyana.prot, cyana.seq, cyana.wc
```

```
#####
#-----FLYA----(Peter Guntert)-----#
#####
# requires Cyana3.9
```

```
# Unpack the FLYA.tar file by
tar -xvf FLYA.tar
```

```
# FLYA uses two directories, one with the original input data ("orig"), one for the
# actual calculation with the name of the project ("FZL4")
# "orig" should be accessible from "FZL4" via "../orig"
mkdir FZL4 #--> already created from the tar file
mkdir orig #--> already created from the tar file
```

```
#required files in these directories:
# peak lists of the 1H-13C HSQC, 2D TOCSY and 2D NOESY spectra and for comparison also
# the assigned shift lists (if available) in xeasy format
```

```
./orig/FZL4_HSQC_PeakList
./orig/FZL4_NOESY_PeakList
./orig/FZL4_TOCSY_PeakList
./orig/FZL4_HSQC_ShiftList
./orig/FZL4_NOESY_ShiftList
./orig/FZL4_TOCSY_ShiftList
./orig/FZL4.aco #--> from cp ./CHES2FLYA/cyana.aco ./orig/FZL4.aco
./orig/FZL4.prot #--> from cp ./CHES2FLYA/cyana.prot ./orig/FZL4.prot
./orig/FZL4.wc #--> from cp ./CHES2FLYA/cyana.wc ./orig/FZL4.wc
```

```
# CYANA run macros
./FZL4/REP.cya #--> please see the content of REP.cya at the end
./FZL4/init.cya #--> please see the content of init.cya at the end
./FZL4/FLYA.cya #--> please see the content of FLYA.cya at the end
./FZL4/FZL4.seq #--> from cp ./CHES2FLYA/cyana.seq ./orig/FZL4.seq
```

```
# in FZL4 directory:
# source CYANA 3.9 (or later) executables and run REP.cya followed by FLYA.cya
```

```
cyana PREP.cya --> this will generate the following files:
stat.prot
ref.prot
NOESY.peaks
TOCSY.peaks
C13H1.peaks
all.aco
```

```
cyana FLYA.cya --> this will generate the following files:
flya.tab
flya.prot
TOCSY_exp.peaks
TOCSY_asn.peaks
NOESY_exp.peaks
NOESY_asn.peaks
C13H1_exp.peaks
C13H1_asn.peaks
flya.txt
```

```
#####--PREP.cya#####
dir := ../orig
input := FZL4_NOESY_PeakList,FZL4_TOCSY_PeakList,FZL4_HSQC_PeakList
spectra := NOESY,TOCSY,C13H1
format := NOESY H1 H2, TOCSY H1 H2, C13H1 C H
```

```
read prot $dir/FZL4.prot unknown=warn
write prot stat.prot
```

```
read prot $dir/FZL4_NOESY_ShiftList| unknown=warn
read prot $dir/FZL4_TOCSY_ShiftList| unknown=warn add
read prot $dir/FZL4_HSQC_ShiftList| unknown=warn add
write prot ref.prot
```

```
do i 1 length('spectra')
  read prot ref.prot
  read peaks $dir/$input(i) format="$format(i)"
  write peaks $spectra(i) names
end do
```

```
read $dir/FZL4.aco unknown=warn
write all.aco
```

```
system "cp -a $dir/FZL4.wc wc.cya"
```

```
#####--FLYA.cya#####
#noesy:=NOESY
```

```

#scalar:=TOCSY,C13H1
calibration := # NOE calibration parameters
structures := 500,20 # number of initial, final structures
steps := 10000 # number of torsion angle dynamics steps
upl_values := 2,4,5,2
randomseed := 57923 # random number generator seed
run_assign_reference:=ref.prot
run_assign_iterations:=15000
run_assign_population:=100
run_assign_statistics:=stat.prot
analyze_assign_group:="Sugar: C1\ H1\ / CONSOLIDATED, Bases: C2 H2 C5 H5 C6 H6 C8 H8 / CONSOLIDATED, Stem: 1..8 14..21 C2 H2 C5 H5 C6 H6 C8 H8 C1\ H1\ / CONSOLIDATED, ALL: CONSOLIDATED, Sugar: C1\ H1\, Bases: C2 H2 C5 H5 C6 H6 C8 H8, Stem: 1..8 14..21 C2 H2 C5 H5 C6 H6 C8 H8 C1\ H1\, ALL: *"

if (.not.existfile('start.pdb')) then
./init
read aco all.aco
wc
sugarbond
calc_all 100 #vtfmin
overview start.oww structures=20 pdb
end if

flya refprot=ref.prot noesy=$noesy scalar=$scalar runs=50 plot=
#####

#####--init.cya#####
path:=./cyanadir/macro
name:=FZL4
cyanalib
read seq $name.seq
noesy:=NOESY
scalar:=TOCSY,C13H1

tolerance:=0.02,0.02,0.3
assigncs_accH:=tolerance(1)
assigncs_accC:=tolerance(3)
assigncs_accN:=tolerance(3)

command NOESY_expect peaks
atoms select "C1\ H1\ C2 H2 C5 H5 C6 H6 C8 H8 - C*"
spectrum $peaks append distance=4.0 structures=20 probability=0.9
spectrum $peaks append distance=6.0 structures=20 probability=0.8
spectrum $peaks append distance=9.0 structures=20 probability=0.5
spectrum $peaks append distance=14.0 structures=20 probability=0.3 write
# atom select "*"
# spectrum $peaks append distance=6.0 structures=20 probability=0.01
end

command TOCSY_expect peaks
atoms select "H5 H6 URA RCYT"
spectrum $peaks skipdiagonal append write
end

command C13H1_select
atoms select "C1\ H1\ C2 H2 C5 H5 C6 H6 C8 H8"
end
#####

```