Supplementary Material

Combining asymmetric ^{13}C-labeling and isotopic filter/edit NOESY: a novel strategy for rapid and logical RNA resonance assignment

Regan M. LeBlanc[1,2], Andrew P. Longhini[1], Stuart F.J. Le Grice[2], Bruce A. Johnson[3,4], Theodore K. Dayie[1,*]

[1] Center for Biomolecular Structure and Organization, Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742, United States

[2] Basic Research Laboratory, National Cancer Institute, Frederick, Maryland 21702, United States

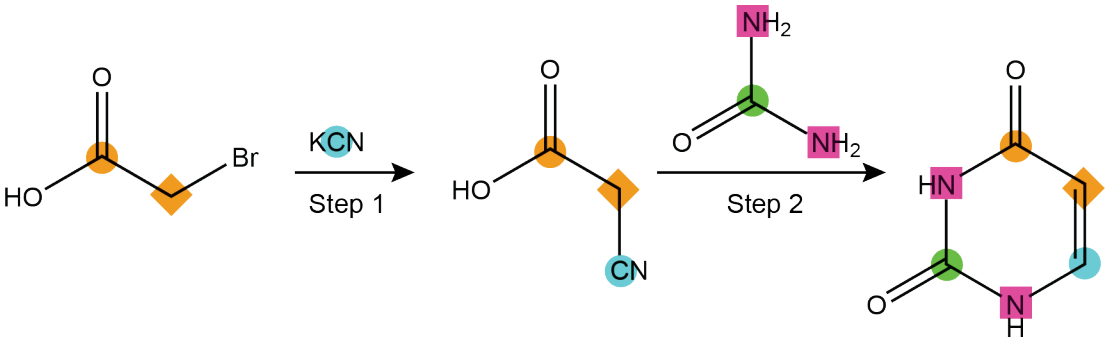
[3] One Moon Scientific, Inc., Westfield NJ 07090, USA

[4] Structural Biology Initiative, Advanced Science Research Center at the Graduate Center of the City University of New York, New York, NY 10031, USA

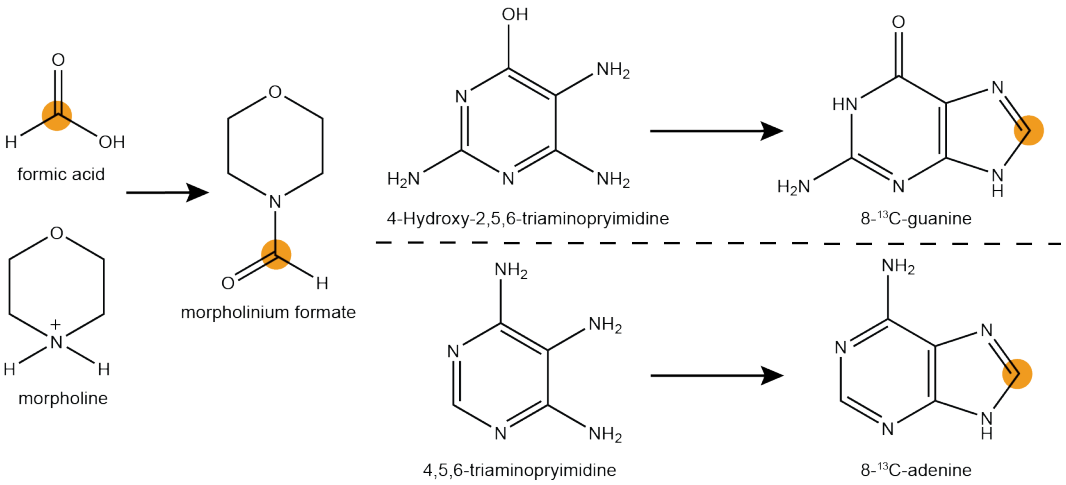
[*] To whom correspondence should be addressed. Tel: +1 301 405 3165; email: dayie@umd.edu

Supplementary Methods

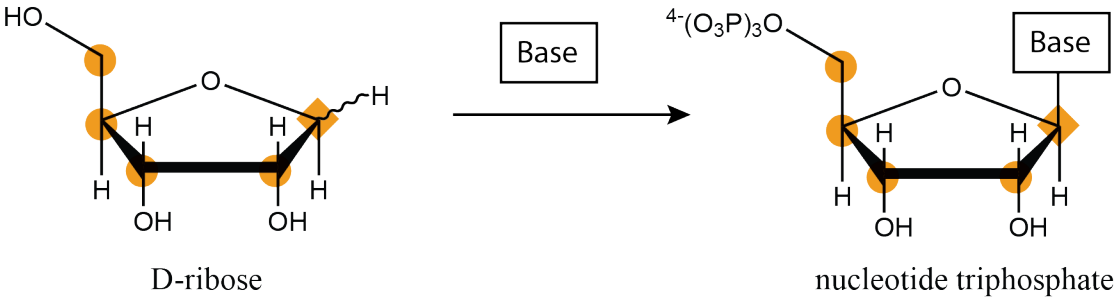
Nucleotide Synthesis: All nucleotides were synthesized based on previously published protocols (1-3). Briefly, chemically synthesized site-specifically labeled base (**Supplementary Figure 1 and 2**) and ribose precursors were enzymatically coupled in one pot reactions (**Supplementary Figure 3**). All reactions were purified over a boronate resin and 4 °C. All reactions had a dATP regeneration system with 0.5 mM dATP, 100 mM Creatine Phosphate, 5 µg/mL Creatine Kinase, 10 mU/µL Myokinase, and 4 mU/µL TIPP. For a 10 mL UTP reactions, 8 mM uracil and 8 mM ribose were added to Buffer A containing 50 mM Na_3PO_4 pH 7.5, 10 mM $MgCl_2$, 2 mM Ampicillin, 10 mM DTT, and 0.1 mg/mL BSA. UMP was synthesized by the combined action of 5 mU/µL ribokinase, 0.3 mU/µL PRPPS, and 5 mU/µL UPRT. After 5 hours at 37 °C, 10 mM KCl, 0.1 mM dATP, 10 mM Creatine Phosphate and 50 µg/mL NMPK were added. For CTP, 2 mM purified UTP with 2 mM $MgCl_2$, 0.5 mM $MgCl_2$, 100 mM Creatine Phosphate, 5 µg/mL Creatine Kinase, 10 mU/µL Myokinase, 0.10 mg/mL CTPS, and 20 mM ^{15}N ammonium chloride pH 8.0 were combined and incubated for 5 hours at 37 °C. ATP and GTP were both formed in Buffer A with 100 mM KCl and the same dATP regeneration system described above. For ATP 8 mM adenine and 8 mM ribose were converted substituting 5 mU/µL APRT for UPRT. For GTP, 5 mM GTP and 6 mM ribose were combined with 5 mU/µL XGPRT and GK for UPRT and NMPK.



Supplementary Figure 1. The chemical synthesis of uracil. The potential isotope labeling sites of uracil are color coded to match the source of enrichment from the precursor compounds. The two carbon position of 2-bromoacetic acid are labeled as either a circle or diamond to reflect the corresponding position in uracil. Step 1) 2-bromoacetic acid, Na_2CO_3 and KCN are refluxed in H_2O at 80 °C for 3 h while stirring and then allowed to stir at room temperature for another 20 h. Step 2) Isolated 3-cyanoacetic acid is combined with urea in diethyl ether and heated to 90 °C for 30 min then 5% $Pd/BaSO_4$ is added under a H_2 atmosphere until uracil is obtained. More detailed synthesis conditions are provided elsewhere (1).



Supplementary Figure 2. The chemical synthesis of isotopically enriched purines. The C8 atom of adenine and guanine are ^{13}C enriched from formic acid precursor. Typical yields are 60% for adenine and 90% for guanine. Full experimental details can be found in the supplementary material of our previously published work (2).



Supplementary Figure 3. The enzymatic coupling of isotopically enriched ribose to chemically synthesized pyrimidine or purine base. Either adenine, guanine, or uracil can be coupled to D-ribose using ribokinase, phosphoribosyl pyrophosphate, a base-specific phosphoribosyl transferase, nucleoside monophosphate kinase, and a dATP regeneration system. Carbon labeling positions on the final NTP ribose match the input ribose.

Supplementary Table 1. All NMR experiments used in this study are given in the table along with relevant parameters. The Bruker pulse sequence used or modified for RNA is given under pulse program. The italicized experiments were run in 90/10% H_2O/D_2O. All other experiments were run in 99.98% D_2O. Two samples were used in this study: a 2',8-^{13}C-A; 1',6-^{13}C-1,3-^{15}N-C; 1',8-^{13}C-G; 2',6-^{13}C-1,3-^{15}N-U labeled bacterial Asite (sel-Asite) and a uniform ^{13}C/^{15}N-AUGC bacterial A-site (u-Asite) sample. The carrier, sweep width (SW), and complex points (TD) is given for each dimension of the 2D experiments. All 3D experiments were run as 2D by setting the unobserved dimension to 1 TD. The total number of scans (NS) and recovery delay (d1) is provided for each experiment. The total experimental time (Expt Time) is reported in days (d), hours (h), and minutes (m). All J-coupling transfers were optimized with 1D or short 2D experiments to account for relaxation. The mixing time for the non-exchangeable proton NOESYs was set to 300 ms and 150 ms for the imino NOESY.

			F2 Dimension			**F1 Dimension**					
Sample	**Experiment**	**Bruker Pulse Program**	**Carrier (ppm)**	**SW (ppm)**	**TD**	**Carrier (ppm)**	**SW (ppm)**	**TD**	**NS**	**d1**	**Expt Time**
sel-Asite	F1eF2e NOESY	noesyhsqcetgpsi3d	4.7	8	1024	4.7	8	256	32	5 s	12h22m
sel-Asite	F1eF2f NOESY	hsqcgpnowgx33d	4.7	8	1024	4.7	8	256	32	1.5 s	4h32m
sel-Asite	F1fF2e NOESY	noesyhsqcgpwgx13d	6.5	7	1024	6.5	8	256	32	1.5 s	4h34m
sel-Asite	^{13}C HSQC	hsqcetgpsi2	4.7	8	1024	115	90	1024	4	1.5 s	1h52m
u-Asite	*1D imino*	*kgzgwg*	*4.7*	*22*	*2048*	*117*	*NA*	*NA*	*16*	*1.5 s*	*2.5m*
u-Asite	*Imino NOESY*	*noesyesfbgpph*	*4.72*	*22*	*2048*	*7*	*15*	*512*	*128*	*2 s*	*1d17h20m*
u-Asite	*^{15}N HSQC*	*hsqcetf3gpsidec*	*4.7*	*22*	*1024*	*153*	*28*	*256*	*32*	*1.2 s*	*3h41m*
u-Asite	*2 bond ^{15}N HSQC*	*hsqcetf3gpsidec*	*4.69*	*9*	*1024*	*200*	*84*	*256*	*120*	*1.5 s*	*13h37m*
u-Asite	TROSY HCN	na_trhcnetsi3d	4.7	10	2048	148	12	128	16	1.5 s	56m
u-Asite	^{13}C Ribose HSQC	hsqcetgpdec	4.7	10	1024	79	54	256	64	1.5 s	7h20n
u-Asite	^{13}C Base HSQC	hsqcetgpdec	4.7	13.3	1024	143	24	256	32	1.5 s	3h37m
u-Asite	HCCH COSY	hcchcogp3d	6.0	5.68	2048	6.0	3	128	8	1 s	38m
u-Asite	HNN-COSY	na_hnncosygpphwg	4.69	22	2048	188	100	256	120	1.5 s	14h20m

Supplementary Table 2. All assigned non-exchangeable resonances from the alternatively labeled bacterial A-site RNA. The box for each assignment in the table is color coded to correspond with the experiment necessary for initial assignment: F1fF2e (blue), F1eF2f (salmon), F1eF2e and HSQC (green). The A20-H2 resonance could not be assigned with the selectively labeled sample only and is highlighted in purple. The dash corresponds to resonance that do not exist. The blank boxes are for ^{12}C-labelled positions in the alternatively labeled sample. The NMR chemical shifts were referenced against 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) following previously published protocol (4).

	H1'	H2'	H6/8	H2/5	C1'	C2'	C6/C8
G1	5.69	4.81	8.01	-	91.30		138.93
G2	5.79	4.41	7.49	-	92.88		136.30
C3	5.46	4.58	7.52	5.08	93.41		140.43
G4	5.60	4.45	7.34	-	93.05		136.03
U5	5.31	4.02	7.44	5.06		75.04	141.31
C6	5.55	4.43	7.85	5.59	92.88		141.92
A7	5.87	4.45	8.00	7.73		75.57	139.46
C8	5.12	4.14	7.40	5.23	93.58		140.51
A9	5.71	4.27	8.02	7.34		75.39	139.64
C10	5.31	4.10	7.46	5.06	93.23		140.69
C11	5.36	4.34	7.50	5.28	93.67		140.78
U12	5.48	3.64	7.64	5.63		75.74	140.51
U13	5.99	4.55	7.91	5.74		74.42	144.56
C14	5.84	3.98	7.57	6.01	88.92		142.71
G15	5.85	4.74	7.74	-	94.37		142.80
G16	4.36	4.47	8.19	-	92.79		138.76
G17	5.69	4.39	7.12	-	93.05		136.21
U18	5.42	4.39	7.51	5.02		75.21	140.77
G19	5.57	4.15	7.55	-	91.47		136.82
A20	5.80	4.52	7.99	7.87		75.92	140.60
A21	5.89	4.61	8.17	7.99		76.18	141.13
G22	5.41	4.63	7.26	-	93.23		137.35
U23	5.34	4.07	7.45	5.06		75.13	141.74
C24	5.59	4.45	7.89	5.67	92.44		141.89
G25	5.60	4.45	7.49	-	93.05		136.74
C26	5.44	4.15	7.61	5.08	94.02		141.22
C27	5.64	3.88	7.56	5.39	92.79		141.57

Supplementary Table 3. Filter/Edit NOESY Contacts in HCV Domain IIa IRES. Contacts less than 5 Å are based on measurements made between listed protons in lowest energy structure from PDB 2HUA (5). Intra residue contacts were not counted.

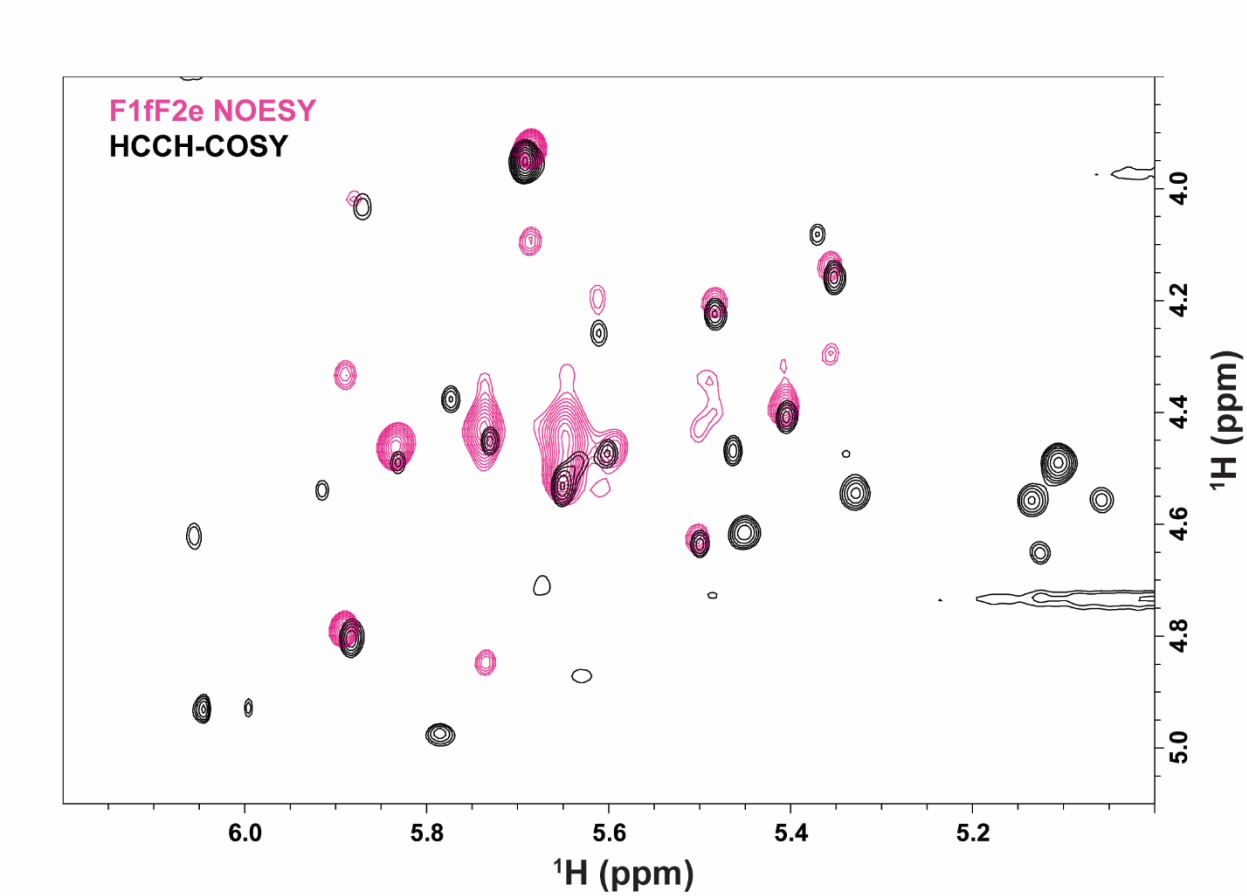
	H2-H1'	H2-H6/H8	H1'-H1'	H1'-H8
A8	1	2	1	0
A30	3	0	0	0
A31	0	1	1	1
A32	2	1	1	1

Supplementary Table 4. Filter/Edit NOESY Contacts in A-Site. Contacts less than 5 Å are based on measurements made between listed protons in lowest energy structure from PDB 1A3M (6). Intra residue contacts were not counted.

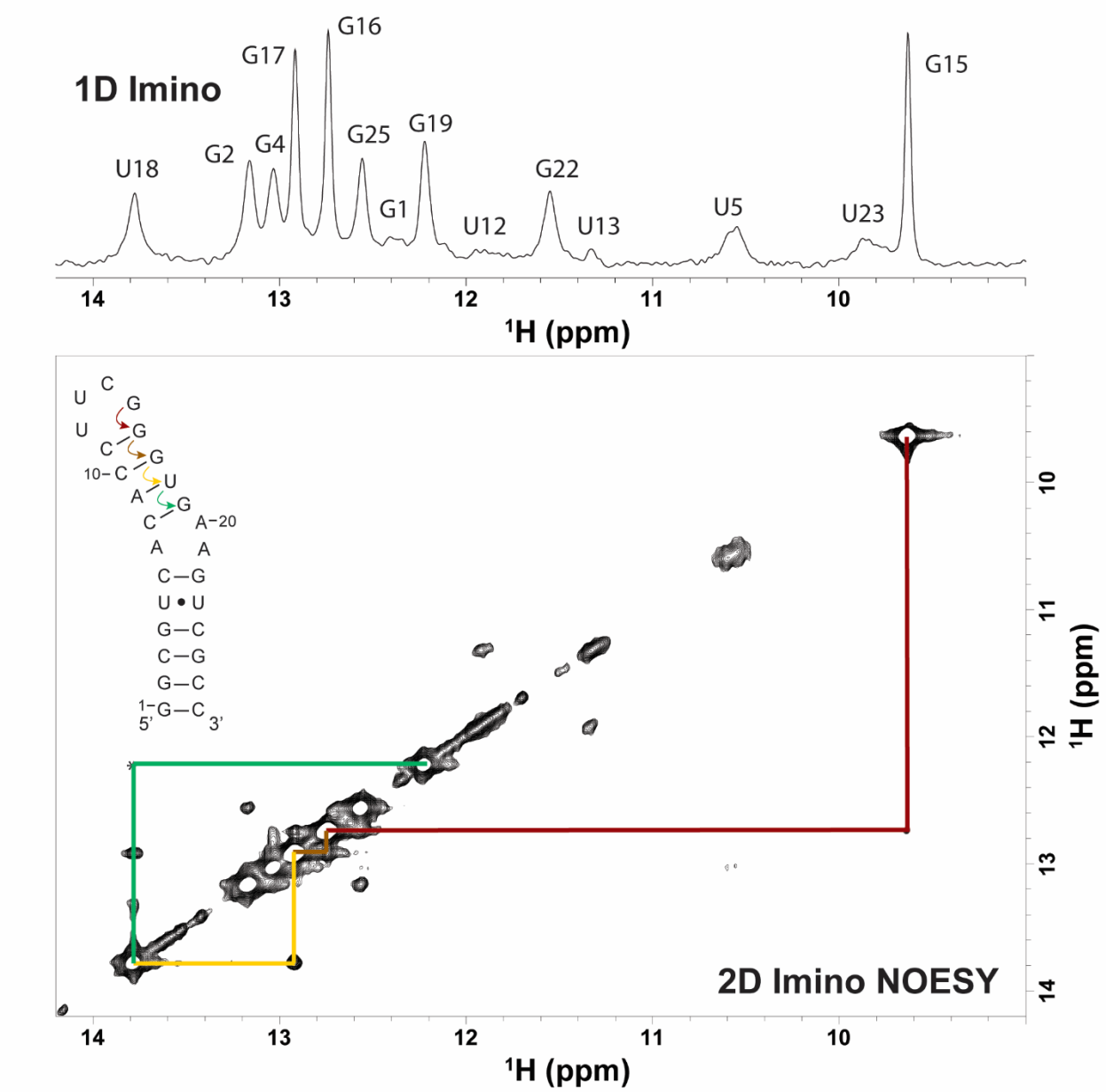
	H2-H1'	H2-H6/H8	H1'-H1'	H1'-H8
A7	2	0	0	1
A20	1	0	0	0
A21	1	1	0	1

Supplementary Table 5. Filter/Edit NOESY Contacts in HIV-Tar. Contacts less than 5 Å are based on measurements made between listed protons in lowest energy structure from PDB 1ANR (7). Intra residue contacts were not counted.

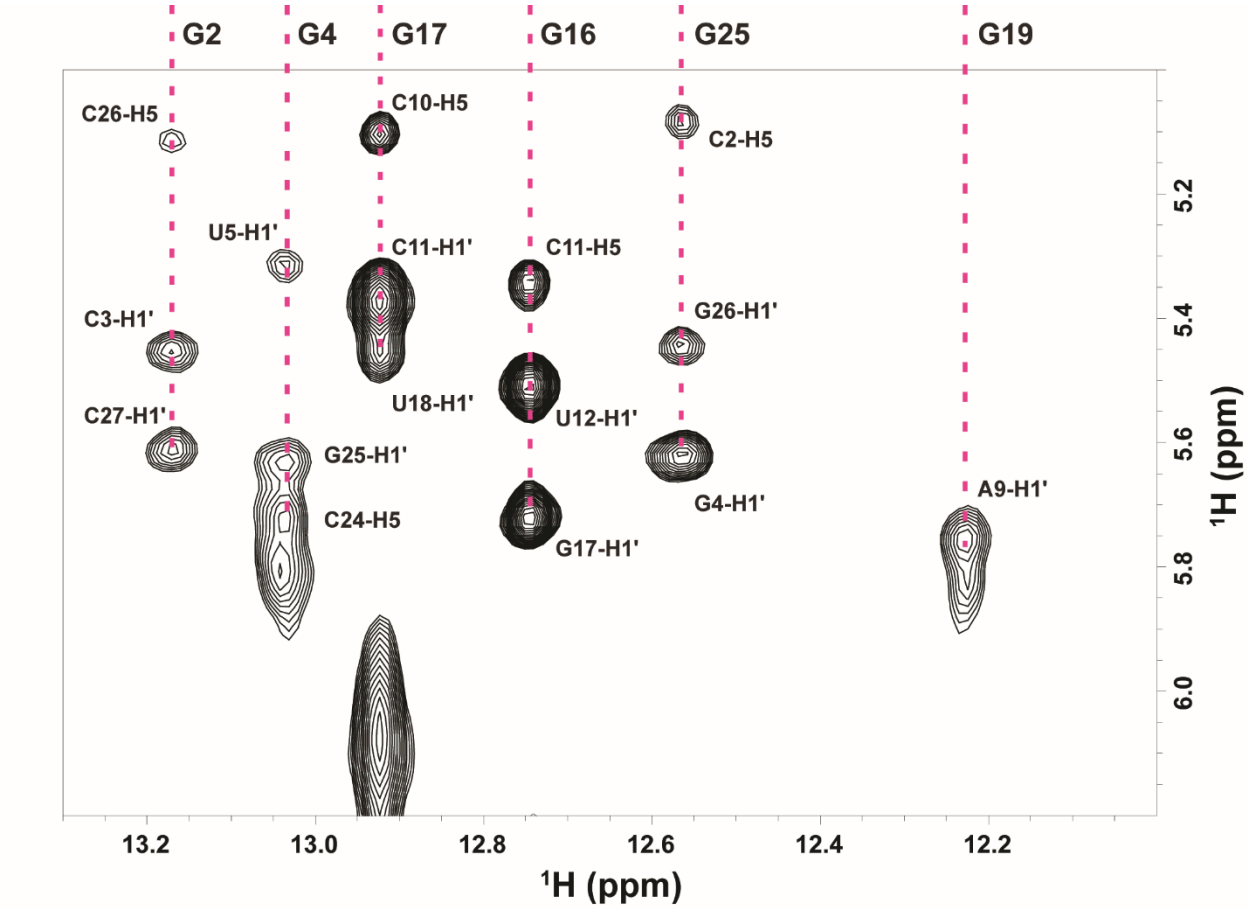
	H2-H1'	H2-H6/H8	H1'-H1'	H1'-H8
A22	1	0	0	1



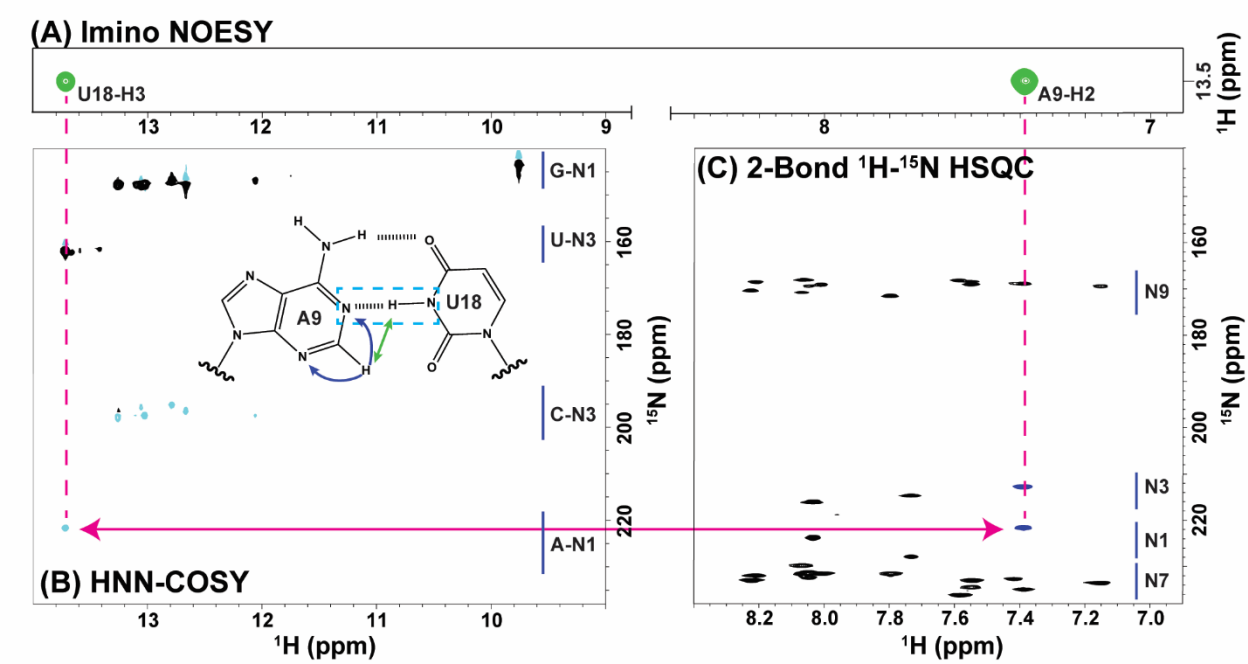
Supplementary Figure 4. The yellow region of the F1eF2f NOESY (magenta peaks) is overlaid on the HCCH-COSY (black peaks) of the uniform bacterial A-site RNA sample. The strong cross-peaks of the F1eF2f NOESY correspond to the through bond correlations in the HCCH-COSY indicating intra-nucleotide H1' and H2' cross-peaks.



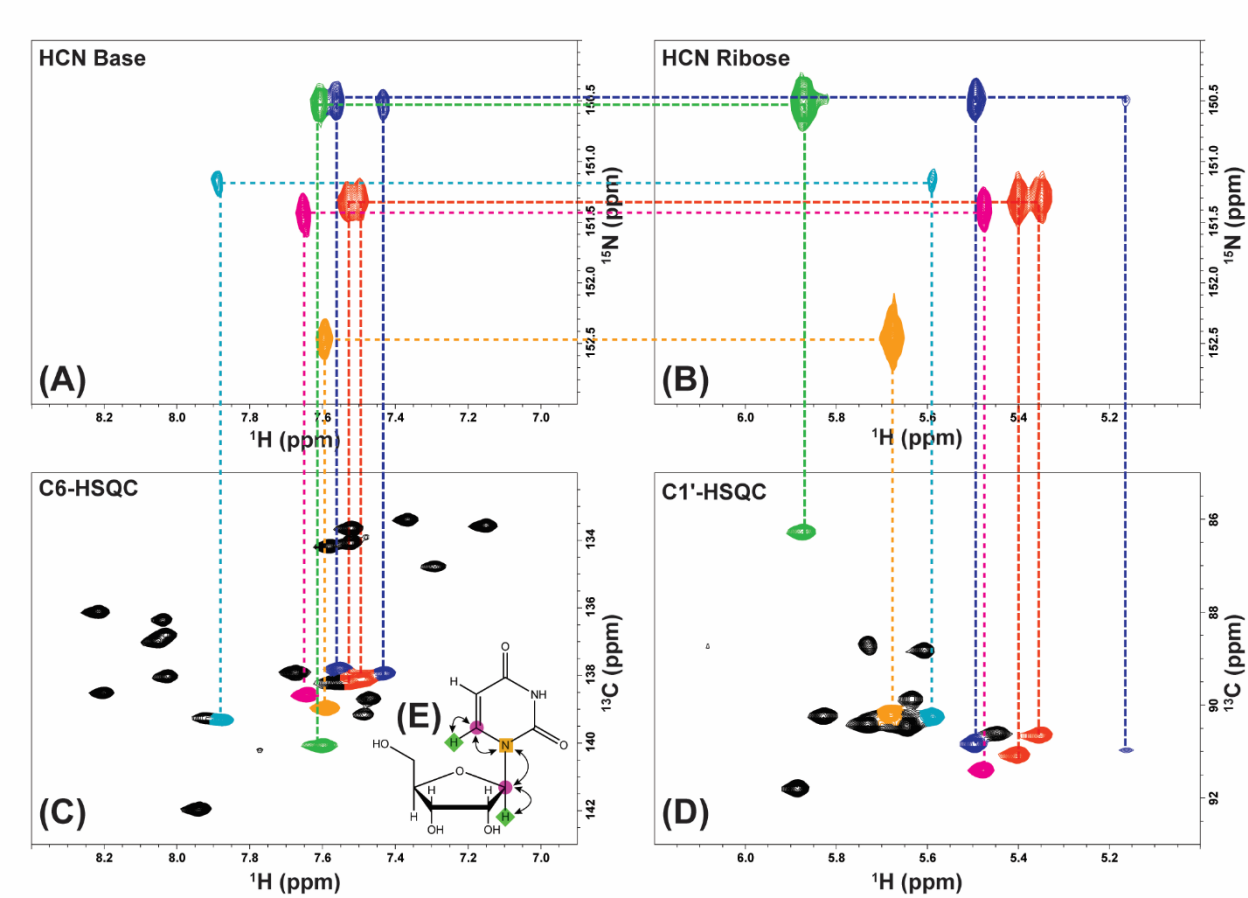
Supplementary Figure 5. A 1D imino spectra is shown projected above a 2D ^{1}H-^{1}H imino NOESY spectra with the sequential walk for the upper stem color coded for the secondary structure of bacterial A-site and the cross-peaks of the imino NOESY spectra. The asterisk represents a peak visible below the contour level of the figure.



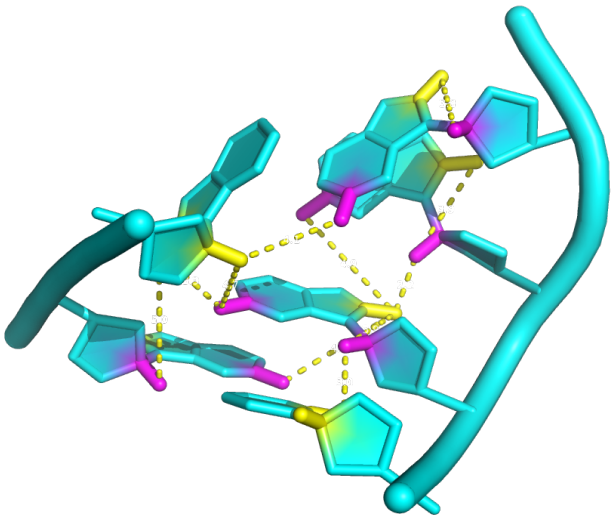
Supplementary Figure 6. The zoomed in G-H1 2D ^{1}H-^{1}H NOESY cross-peaks with H1' and H5 through spin-diffusion are labeled within the spectrum. The G-residue corresponding to the H1 (imino) resonance is labeled above the spectra for each set of cross-peaks next to the dashed magenta lines.



Supplementary Figure 7. A) Imino NOESY spectra showing the through-space crosspeak between A9-H2 and U18-H3 (green peaks) as shown in the A-U base-pair cartoon (green arrow). B) The HNN-COSY hydrogen bond experiment correlates U18-N3 to A9-N1 through U18-H3 and is highlighted in the A-U base-pair by the blue dashed box. C) The 2-bond ^{1}H-^{15}H-HSQC correlates A9-H2 to A9-N3 and A9-N1 (purple arrows in A-U base-pair) and correlates with the A9-N1 resonance from the HNN-COSY experiment (magenta arrow).



Supplementary Figure 8. The H6 and H1' intra-nucleotide pairs for cytidine were confirmed with a base HCN (A) and ribose HCN (B) along a HSQC of the base (C) and ribose (D). The transfer pathway was limited to the desired HCN transfer (E) by eliminating C-C coupling at C1' and C6 greatly enhancing the sensitivity of the experiment. The "connected" pairs were color coded for ease of viewing with similar colors indicating ambiguity due to overlap in the ^{15}N dimension as seen for purple and red peaks. This ambiguity was previously solved with the F1eF2e NOESY walk.



Supplementary Figure 9. Representative contacts seen for the asymmetric bulge of the HCV Domain IIA IRES (PDB:1HUA) (5). Protons marked in yellow (H2 and H1' of adenine) are attached to ^{12}C labeled nuclei. Protons marked in magenta are attached to ^{13}C labeled nuclei. All dotted lines represent measurements that are less than 5 Å and would theoretically have contacts in a F1eF1f NOESY experiment.

1. Alvarado, L.J., LeBlanc, R.M., Longhini, A.P., Keane, S.C., Jain, N., Yildiz, Z.F., Tolbert, B.S., D'Souza, V.M., Summers, M.F., Kreutz, C. *et al.* (2014) Regio-Selective Chemical-Enzymatic Synthesis of Pyrimidine Nucleotides Facilitates RNA Structure and Dynamics Studies. *ChemBioChem*, **15**, 1573-1577.
2. Alvarado, L.J., Longhini, A.P., LeBlanc, R.M., Chen, B., Kreutz, C. and Dayie, T.K. (2014) In Donald, H. B.-A. (ed.), *Methods in Enzymology*. Academic Press, Vol. Volume 549, pp. 133-162.
3. Longhini, A.P., LeBlanc, R.M., Becette, O., Salguero, C., Wunderlich, C.H., Johnson, B.A., D'Souza, V.M., Kreutz, C. and Dayie, T.K. (2015) Chemo-enzymatic synthesis of site-specific isotopically labeled nucleotides for use in NMR resonance assignment, dynamics and structural characterizations. *Nucleic Acids Research*.
4. Aeschbacher,T., Schubert,M. and Allain,F.H.-T. (2012) A procedure to validate and correct the 13C chemical shift calibration of RNA datasets. *J. Biomol. NMR*, **52**, 179–90.
5. Locker, N., Easton, L.E. and Lukavsky, P.J. (2007) HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. *The EMBO Journal*, **26**, 795.
6. Fourmy, D., Yoshizawa, S. and Puglisi, J.D. (1998) Paromomycin binding induces a local conformational change in the A-site of 16 s rRNA1. *Journal of Molecular Biology*, **277**, 333-345.
7. Aboul-ela, F., Karn, J. and Varani, G. (1996) Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. *Nucleic Acids Research*, **24**, 3974-3981.