

Structural analysis of the 5' terminal of human immunodeficiency type 1 genome RNAs transcribed from proviruses with mutations in the transcription start site

Yuka Kashiya¹, Gota Kawai¹, Takeshi Yoshida² and Youichi Suzuki^{3,*}

¹Department of Life Science, Graduate School of Advanced Engineering, Chiba Institute of Technology, Chiba, Japan

²Management Department of Biosafety, Laboratory Animal, and Pathogen Bank, National Institute of Infectious Diseases, Tokyo, Japan

³Laboratory of Biosafety Research, Faculty of Medicine, Osaka Medical and Pharmaceutical University, Takatsuki, Japan

*To whom correspondence should be addressed.

Supplemental Fig. S1. Possible secondary structures of model RNAs used in this study

Supplemental Fig. S2. First derivatives of melting curves of model RNAs

Supplemental Fig. S3. Imino proton spectra of model RNAs and their fragments

Supplemental Fig. S4. Imino proton spectra of model RNAs with C mismatch

Supplemental Fig. S5. Imino proton spectra of model RNAs with U mismatch

Supplemental Fig. S6. Hydrogen bond distances during MD simulations of TP-G1G-46

Supplemental Fig. S7. Hydrogen bond distances during MD simulations of TP-G2GC-47

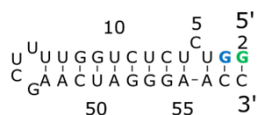
Supplemental Fig. S8. Hydrogen bond distances during MD simulations of TP-G2GU-47

Supplemental Table S1. NMR constraints and structural statistics

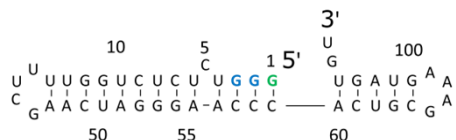
TP-G1G-46



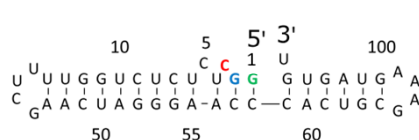
TAR-G1G-29



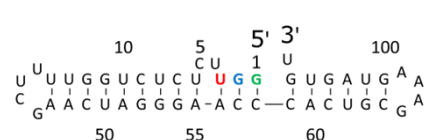
TP-G2GG-47



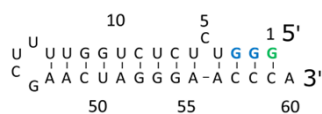
TP-G2GC-47



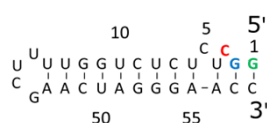
TP-G2GU-47



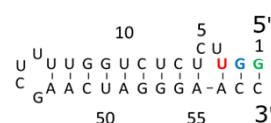
TAR-G2GG-32



TAR-G2GC-30



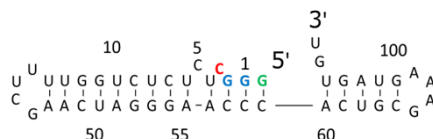
TAR-G2GU-30



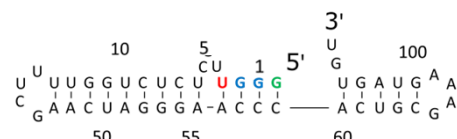
TP-G3GGG-48



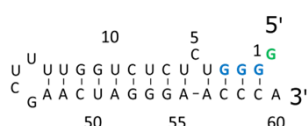
TP-G3GGC-48



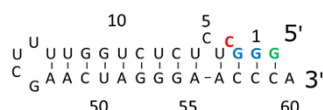
TP-G3GGU-48



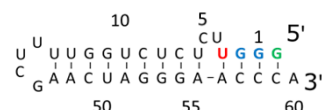
TAR-G3GGG-33



TAR-G3GGC-33

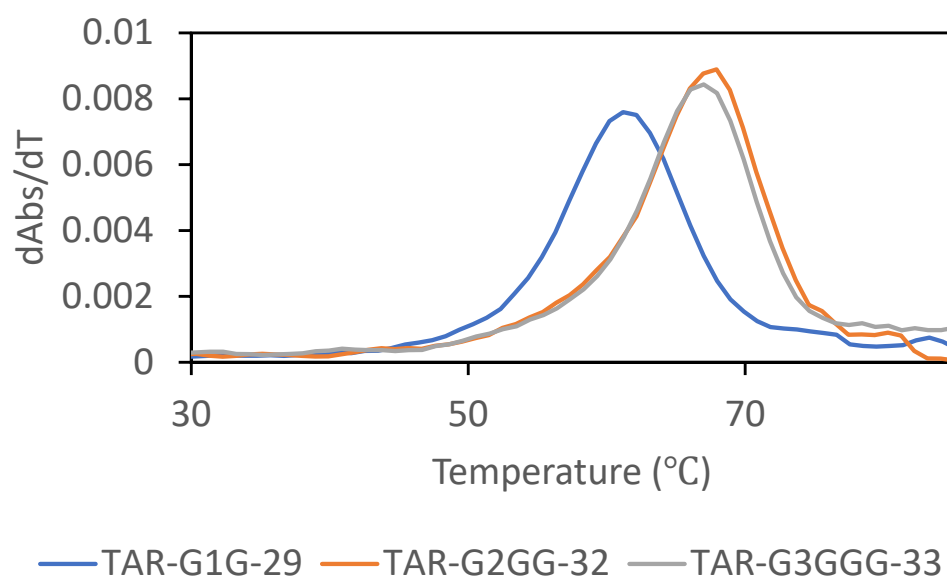
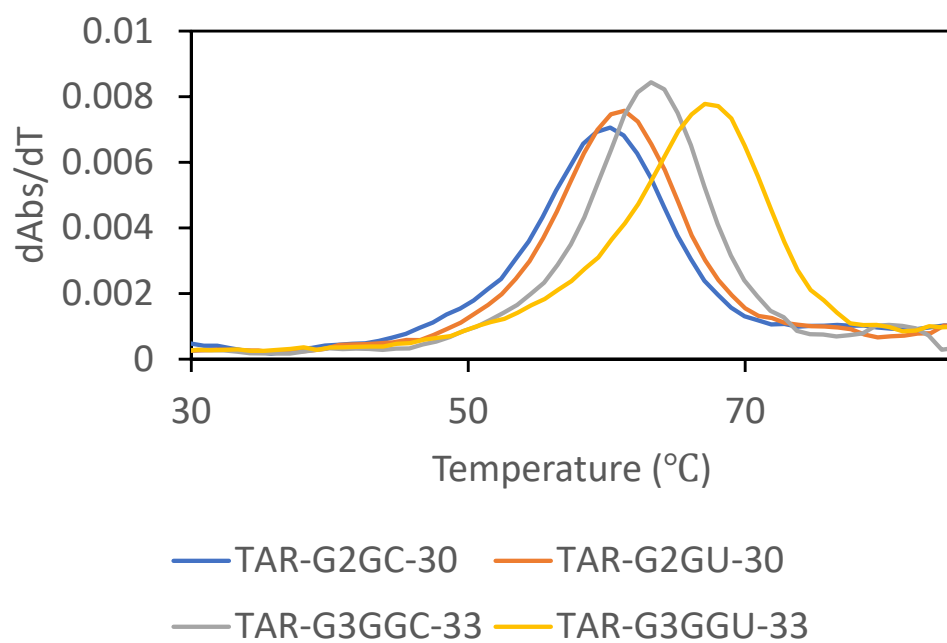


TAR-G3GGU-33

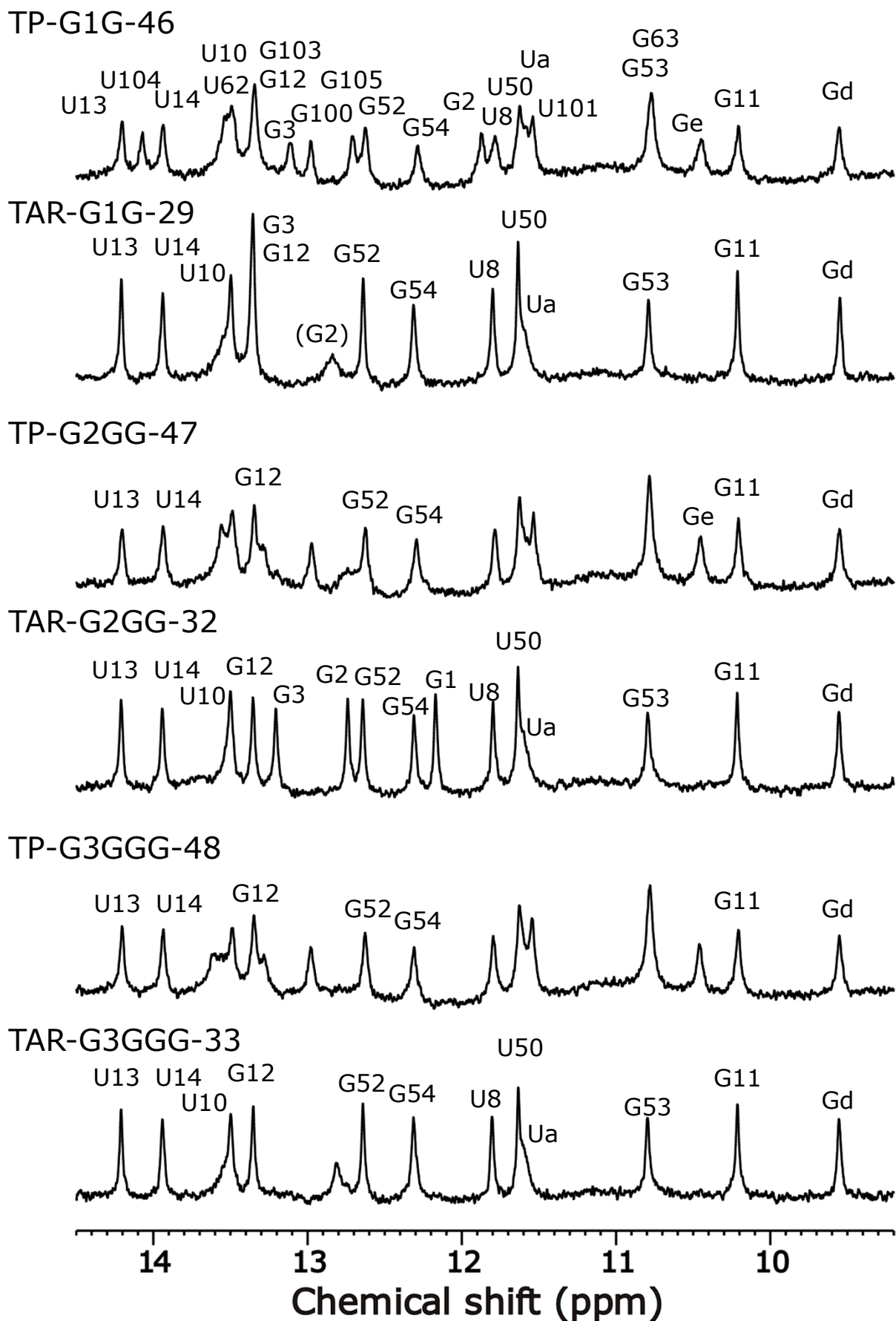


Supplemental Fig. S1. Possible secondary structures of model RNAs used in this study

For each model, the first G residue (green) corresponds to the m⁷G moiety of the Cap structure. G residues transcribed from the GGG regions in the provirus are indicated by blue. The C or U residues transcribed from the mutated residues are indicated by red.

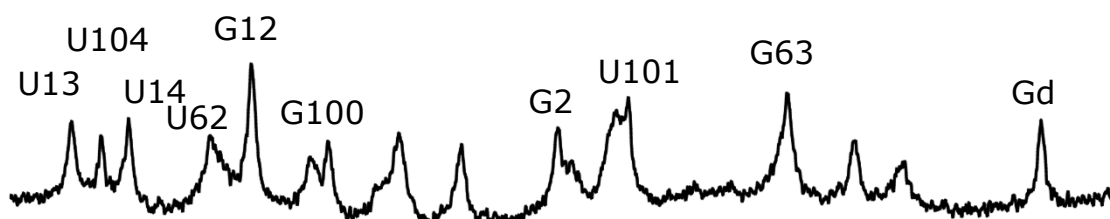


Supplemental Fig. S2. First derivatives of melting curves of model RNAs

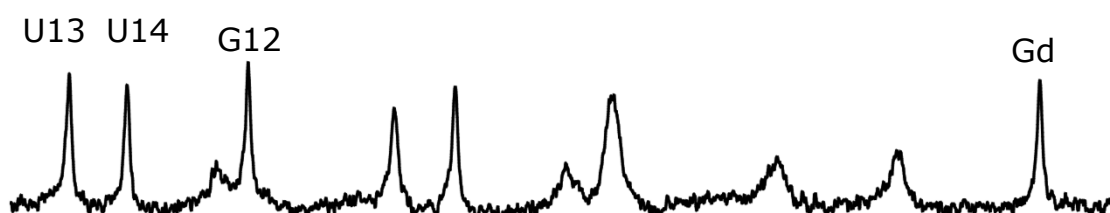


Supplemental Fig. S3. Imino proton spectra of model RNAs and their fragments
Assigned signals were indicated by labels. Residues for UUCG and GAAA loops are indicated by Ua-Gd and Ge-Ah. Parentheses indicate the tentative assignments.

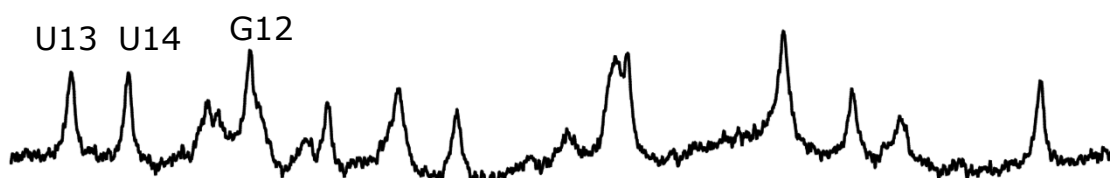
TP-G2GC-47



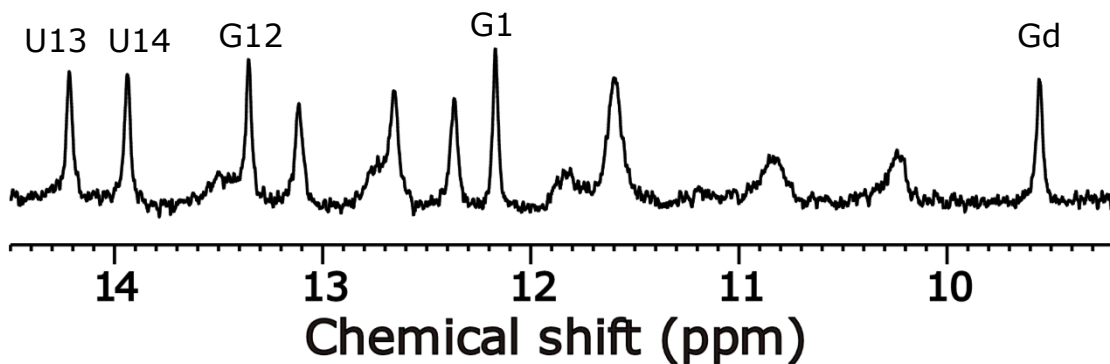
TAR-G2GC-30



TP-G3GGC-48

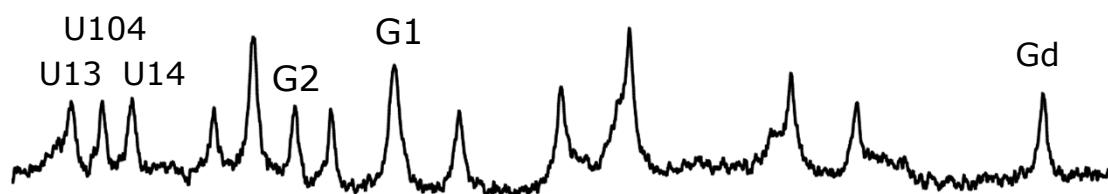


TAR-G3GGC-33

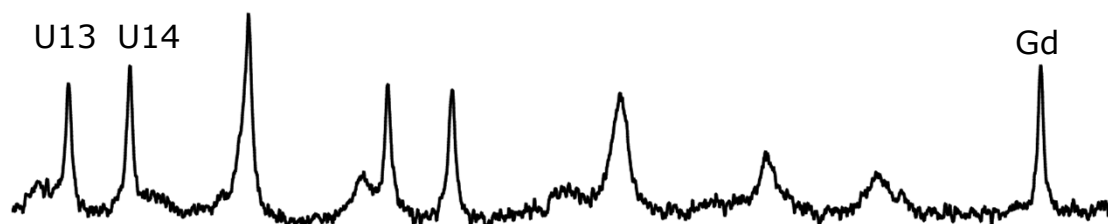


Supplemental Fig. S4. Imino proton spectra of model RNAs with C mismatch
Assigned signals were indicated by labels. Residues for UUCG and GAAA loops are indicated by Ua-Gd and Ge-Ah.

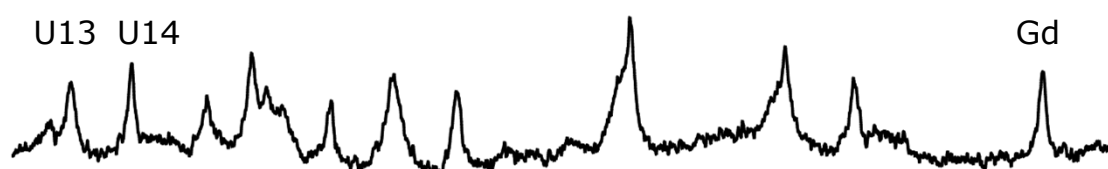
TP-G2GU-47



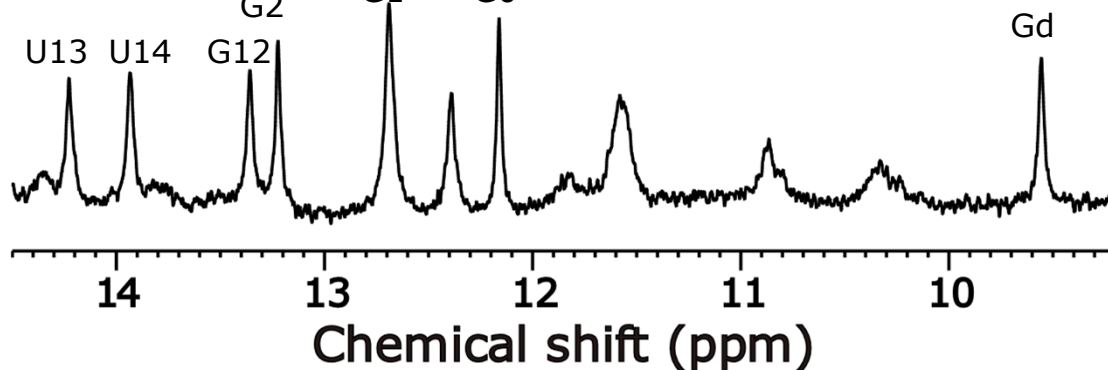
TAR-G2GU-32



TP-G3GGU-48

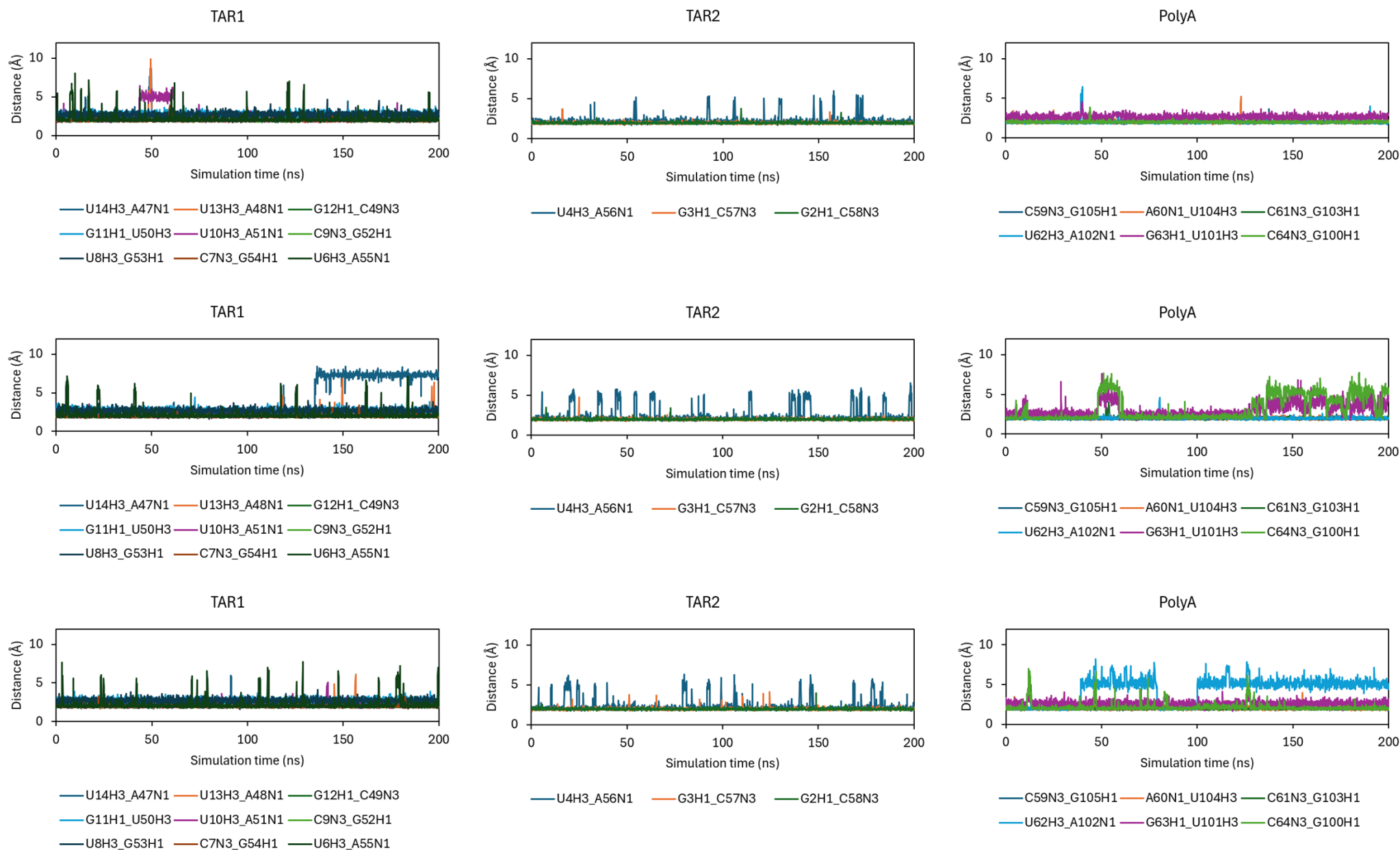


TAR-G3GGU-33

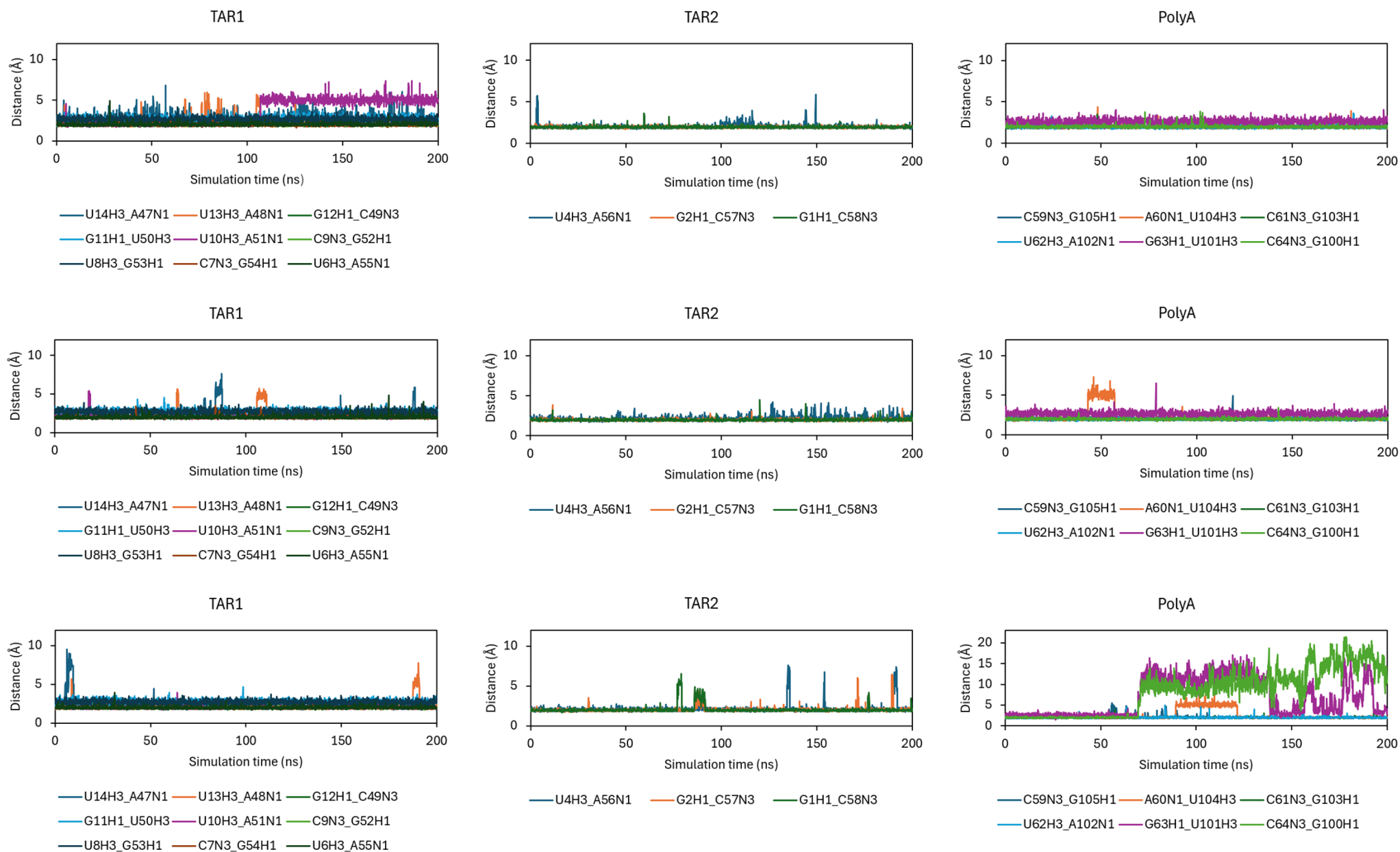


Supplemental Fig. S5. Imino proton spectra of model RNAs with U mismatch

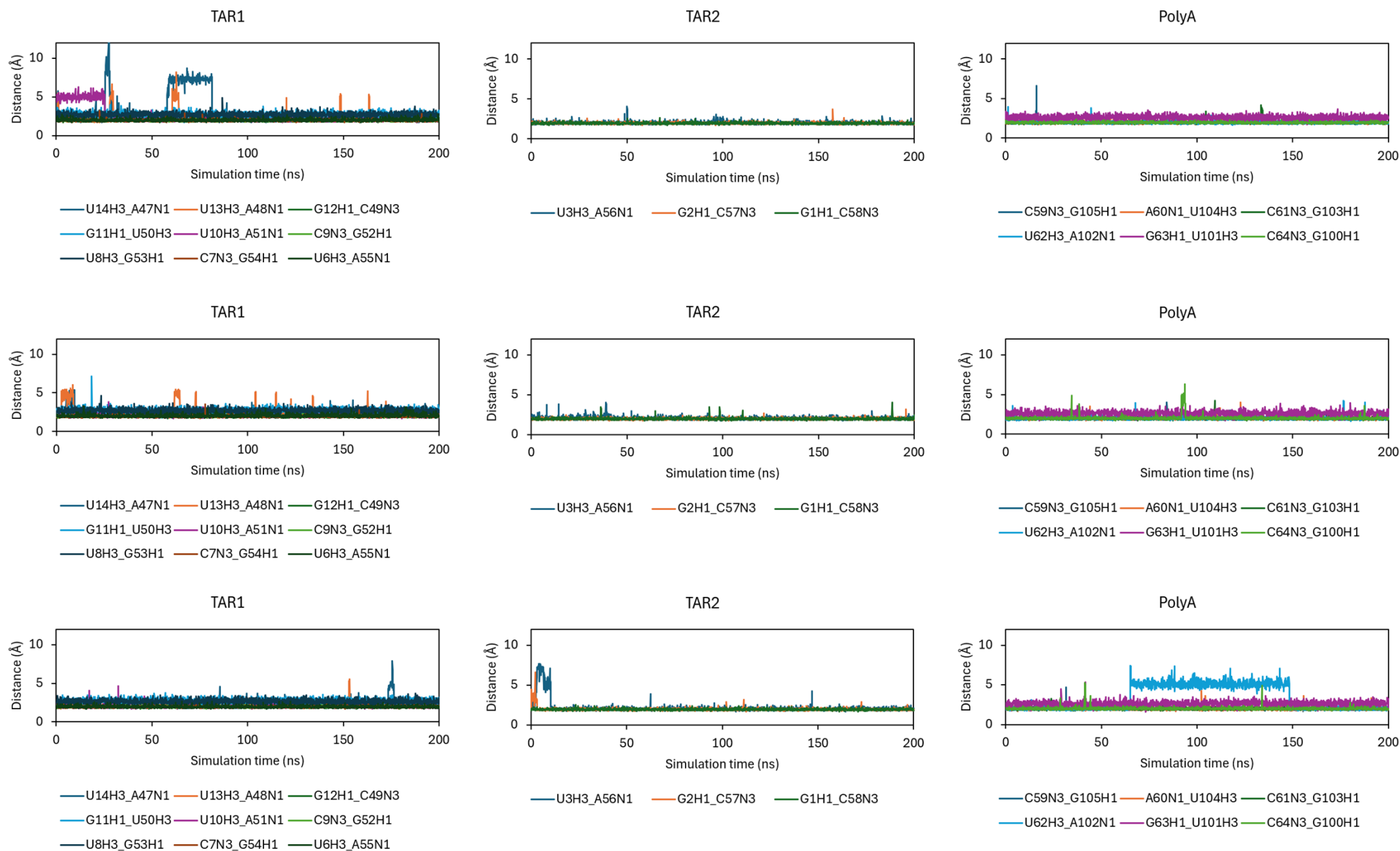
Assigned signals were indicated by labels. Residues for UUCG and GAAA loops are indicated by Ua-Gd and Ge-Ah. G0 indicated the 5'-G residue corresponding to the Cap structure of G3GGU.



Supplemental Fig. S6. Hydrogen bond distances during MD simulations of TP-G1G-46. For each model, 3 structures with lowest energies were analyzed.



Supplemental Fig. S7. Hydrogen bond distances during MD simulations of TP-G2GC-47. For each model with C mismatch, 3 structures with lowest energies were analyzed.



Supplemental Fig. S8. Hydrogen bond distances during MD simulations of TP-G2GU-47. For each model with U mismatch, 3 structures with lowest energies were analyzed.

Supplemental Table S1. NMR constraints and structural statistics

	TP-G1G-46	TP-G2GC-47	TP-G2GU-47
Number of experimental restraints			
Distance restraints	169	131	148
Intraresidue	28	25	31
Sequential	46	31	40
Medium range	11	8	6
Long range	36	16	20
Hydrogen bonding ^a	48	48	48
additional restraints		3	3
Dihedral restraints	400	386	389
Planarity for base pairs	20	20	20
Heavy-atoms r.m.s. deviation for 10 structures (Å) ^b			
All	2.477 ± 1.2162	3.576 ± 1.7298	3.109 ± 1.2531
All (pairwise)	2.926 ± 0.8173	3.153 ± 1.0005	3.201 ± 0.7370
Backbone	2.273 ± 1.2050	3.638 ± 1.8421	2.967 ± 1.2056
Backbone (pairwise)	2.897 ± 0.8766	3.076 ± 1.1317	3.057 ± 0.8201
TAR stem (6-14, 19-27)	0.487 ± 0.2241	0.721 ± 0.3103	0.449 ± 0.1953
TAR stem (pairwise)	0.575 ± 0.1792	0.801 ± 0.2872	0.531 ± 0.1009
PolyA stem (31-36, 41-46)	0.336 ± 0.1317	0.309 ± 0.1370	0.367 ± 0.1330
PolyA stem (pairwise)	0.383 ± 0.0938	0.350 ± 0.0953	0.347 ± 0.0691
r.m.s.d. around the ideal values			
bonds (Å)	0.0027 ± 0.00004	0.0024 ± 0.00004	0.0024 ± 0.00004
angle (°)	0.6809 ± 0.0069	0.6172 ± 0.0033	0.6515 ± 0.0050

^aDistance restraints between imino protons in the bulged regions; G3-U5, U5-U7 and U7-G54 for TP-G2GC-47, and G3-U4, U4-U7 and U7-G54.

^bAveraged r.m.s.d. between an average structure and the 10 converged structures were calculated. The converged structures did not contain experimental distance violation of >0.5 Å or dihedral violation >5°.