Supporting Information for:

Direct mapping of higher-order RNA interactions by SHAPE-JuMP

Thomas W. Christy[1,2], Catherine A. Giannetti[1], Gillian Houlihan[3], Matthew J. Smola[1], Greggory M. Rice[1], Jian Wang[4], Nikolay V. Dokholyan[4,5], Alain Laederach[6], Philipp Holliger[3], and Kevin M. Weeks[1,*]

[1] Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290; [2] Curriculum in Bioinformatics and Computational Biology, University of North Carolina, Chapel Hill, North Carolina 27599; [3] MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK; [4] Departments of Pharmacology, and Biochemistry and Molecular Biology, Penn State University College of Medicine, Hershey, PA 17033, USA; [5] Departments of Chemistry, and Biomedical Engineering, Pennsylvania State University, University Park, PA 16802; [6] Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599

* correspondence, weeks@unc.edu

One supporting table and six figures.
Supporting Datasets available at www.weekslab.com.

Table S1: RMSD values for modeled RNAs as a function of refinement step

	P546 domain	VS ribozyme	RNase P CD	group II intron
Step 1: Base pairing[a]	39	31	32	55
Step 2: + Junctions[a]	28	25	27	46
Step 3: + Tertiary (centroid)[a]	20	21	14	33
Step 3: + Tertiary (restraints)[b]	20	17	15	30

[a] RMSD values calculated relative to the reference structure, based on the centroid of the lowest energy cluster from lowest 1% of energy models.

[b] RMSD values calculated based on the model with the shortest distance range, based on SHAPE-JuMP restraints.

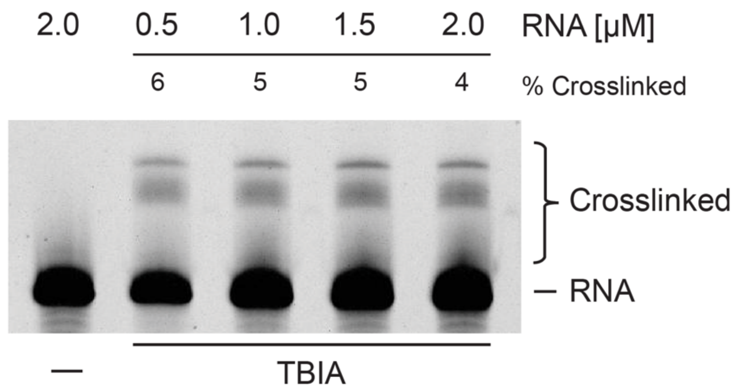


Figure S1: TBIA crosslinking in the Varkud satellite ribozyme. Crosslinking in the VS ribozyme visualized by denaturing gel electrophoresis as a function of RNA concentration, at constant TBIA concentration. Percent crosslinked calculated as the fraction of low mobility products relative to total RNA in each lane. RNA concentrations were normalized to 200 ng prior to gel loading. No reagent control (–) is shown at left.

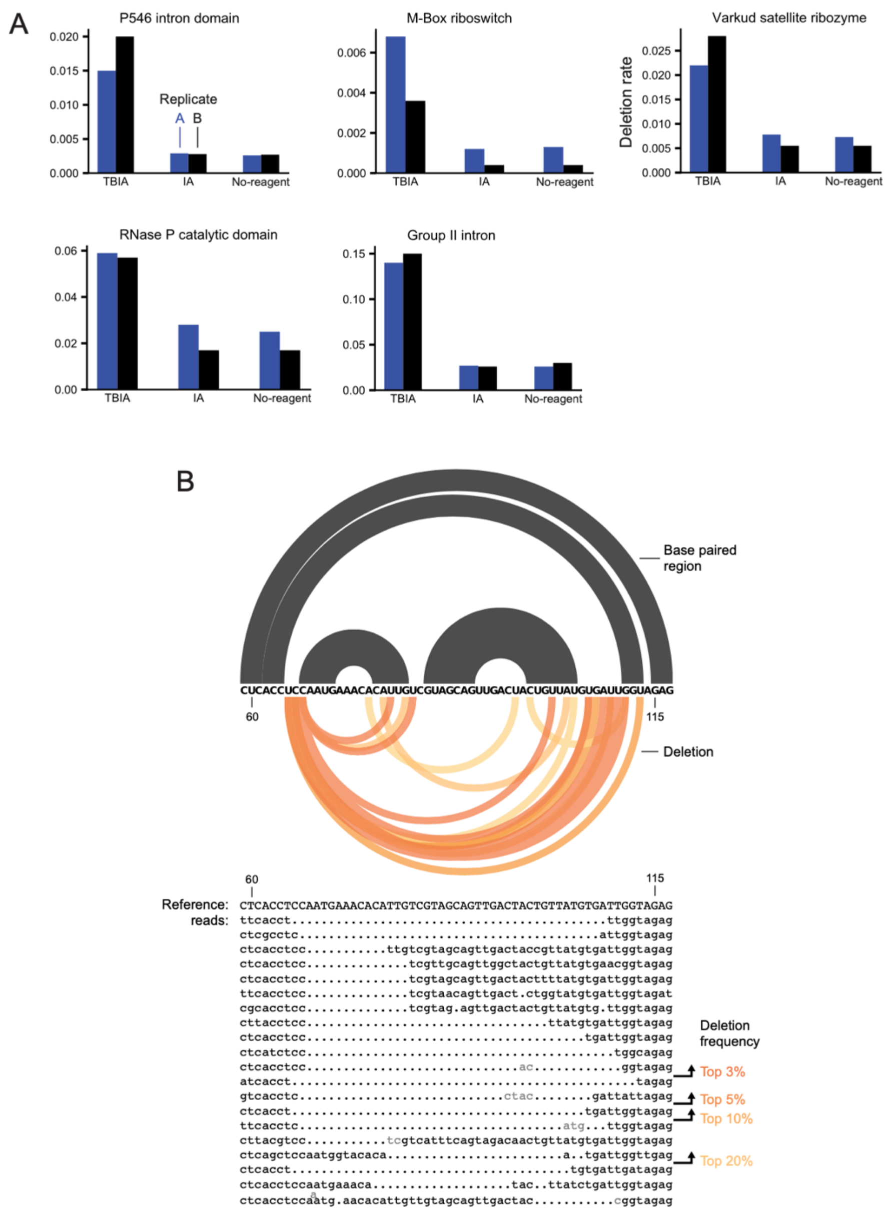


Figure S2: SHAPE-JuMP deletion rates for tested RNAs and representative SHAPE-JuMP deletion data. (**A**) Deletions rates for independent replicates. Rates, measured as total number of deletions divided by total number of aligned reads, are shown for each of two replicates for the no-reagent, mono-adduct (IA), and crosslinker (TBIA) experiments. (**B**) Representative SHAPE-JuMP reads for the 3-4-5 junction region in the VS ribozyme. Sample of 20 reads (lowercase) from a TBIA treated sample containing deletions within junction 3-4-5, aligned to reference sequence (uppercase). Inserts identified in alignments denoted with gray letters. Arc plot shows secondary structure (black arcs) for junction 3-4-5. Deletions in aligned reads are mapped on to the secondary structure (orange arcs). Deletions are colored by their relative frequency, with the darkest corresponding to the most frequent.

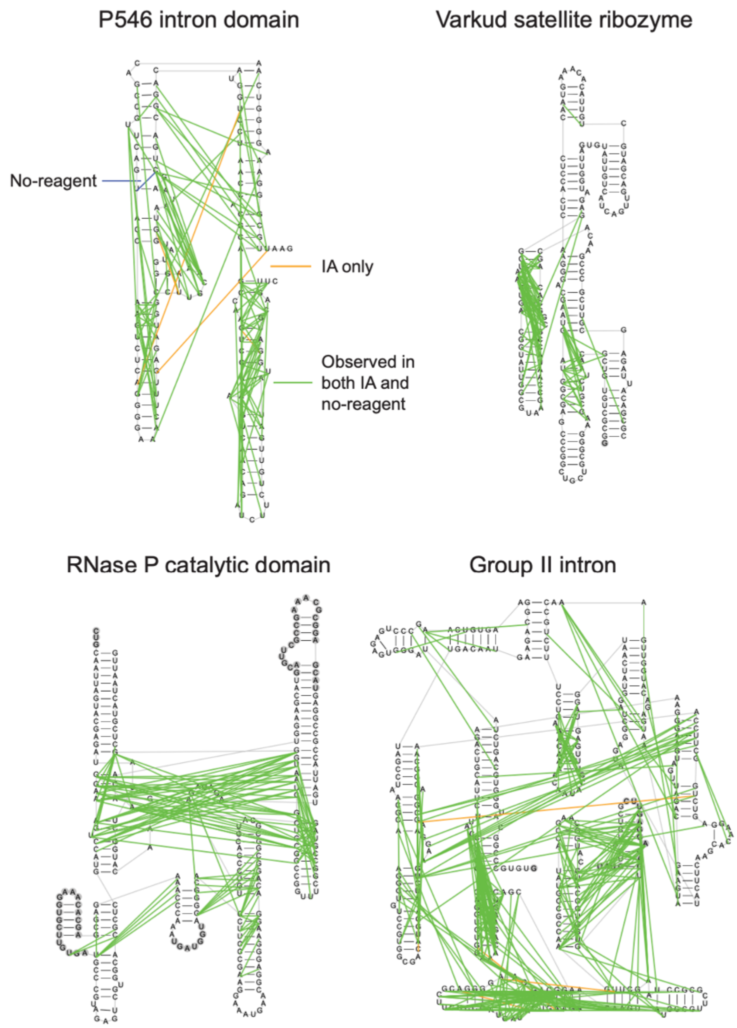
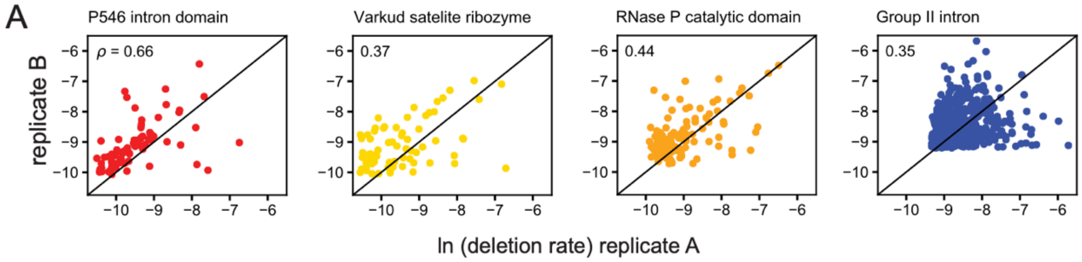


Figure S3: Background deletion signal is primarily due to intrinsic properties of RT-C8. Observed deletions for no-reagent and monoadduct (IA) samples for the most frequent 5% of deletions, mapped onto the reference secondary structures for tested RNAs (green lines). These schematics emphasize no-reagent and IA treatments yield nearly identical background signals.



B

	Deletion intersection with other replicate (%)			
Replicate	P546	VS	RNase P	Group II
A	87	89	80	87
B	94	79	69	82

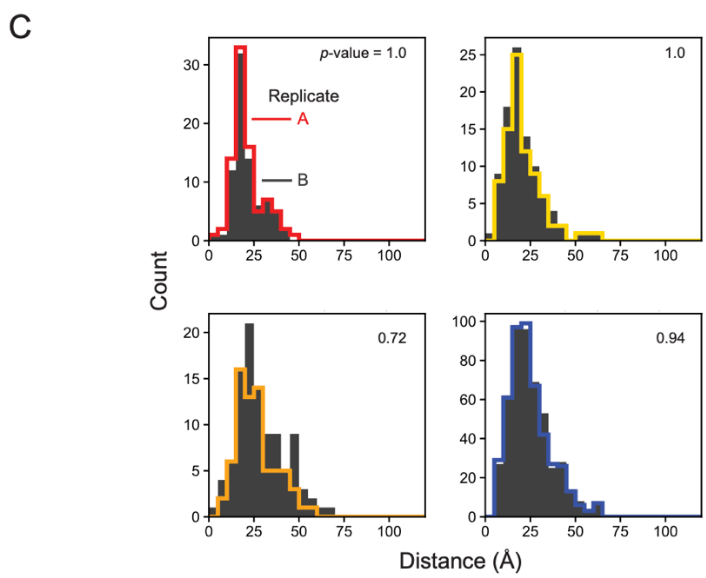


Figure S4: Correlations between deletion rates, locations and through-space distances for replicate experiments. (**A**) Deletion rates for the most frequent 3% of deletions in replicate samples. Spearman correlations are reported. (**B**) Deletion positions between replicates for the most frequent 3% of deletions. Deletion intersections correspond to the percent of contacts that occur in both replicates A and B. (**C**) Distributions of through-space distances for the 3% most frequent SHAPE-JuMP deletions for replicates A and B (colored lines and gray shading). Colors correspond to the RNAs shown in panel (A). Similarities in distribution were assessed by the Kolmogorov-Smirnov test and reported as *p*-values; high *p*-values indicate similar distributions.

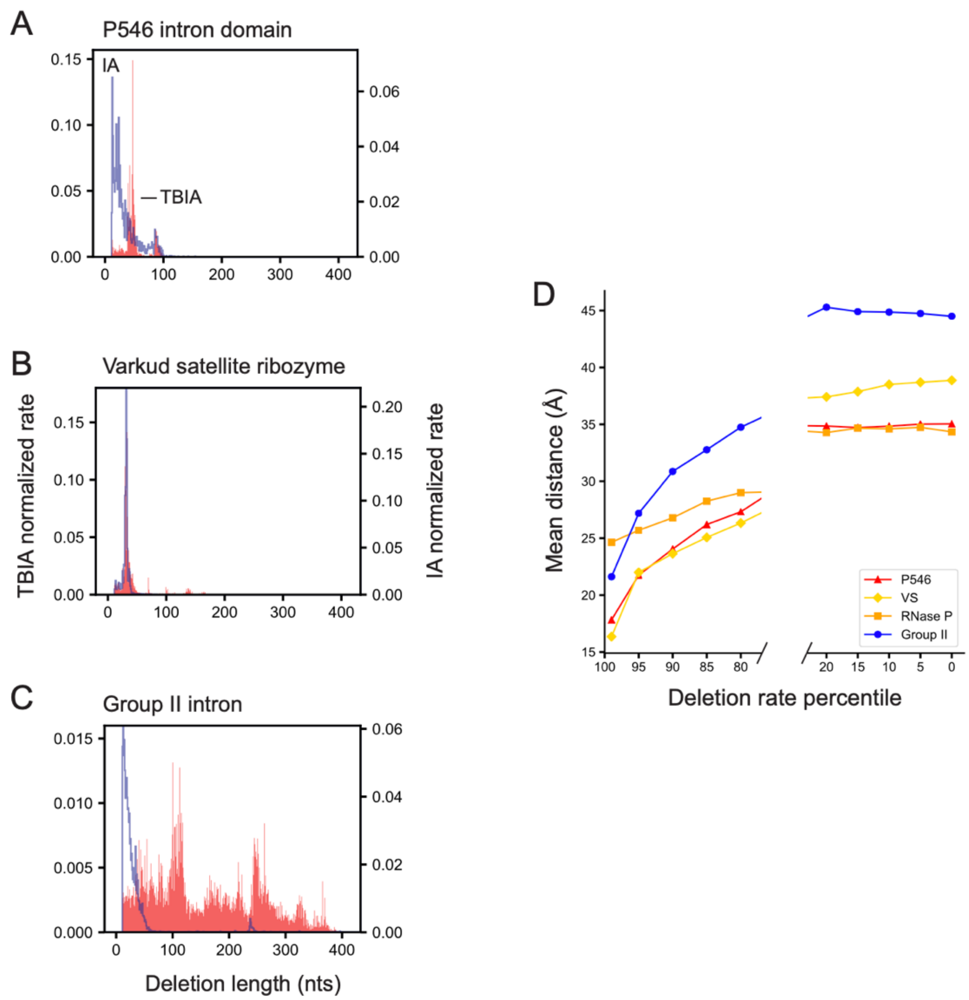


Figure S5: Measured deletion lengths as a function of reagent and frequency. Deletion lengths observed for (**A**) P546 intron domain, (**B**) VS ribozyme, (**C**) and group II intron, shown for crosslinker (TBIA) and monoadduct (IA) experiments. Deletion rates are normalized to sum to 1. (**D**) Mean internucleotide through-space distances for TBIA-induced SHAPE-JuMP interactions as a function of deletion rate threshold. Deletion rates are organized into percentiles where the 99th percentile corresponds to the most frequent 1% of deletions.

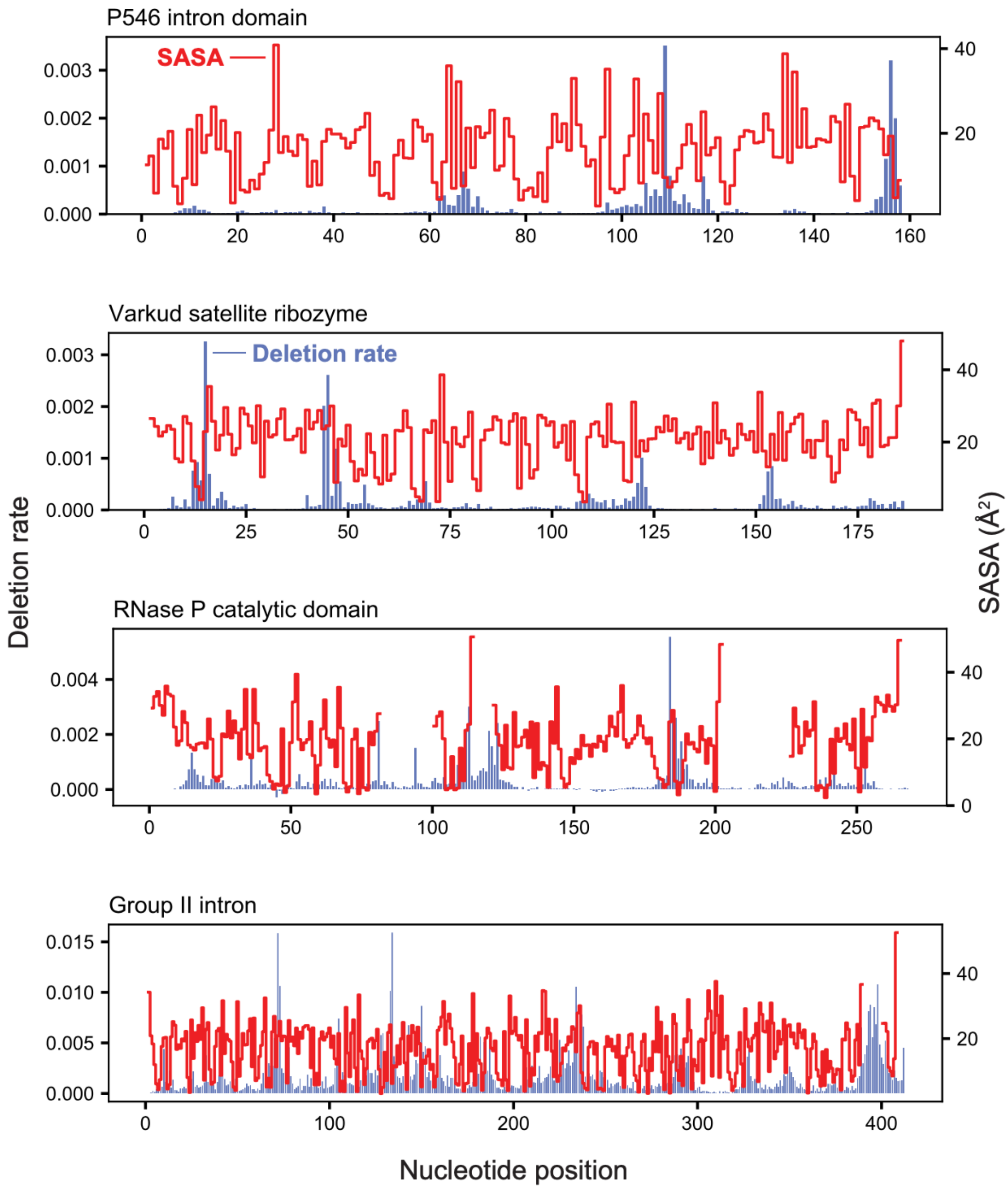


Figure S6: SHAPE-JuMP detects interactions in both solvent-accessible and inaccessible RNA regions. Deletion rates at each nucleotide for each RNA are plotted in blue (left axis) and solvent accessible surface areas (SASA) for each nucleotide 2'-hydroxyl group are in red (right axis).