Supporting Information for:

Fingerprinting Non-Canonical and Tertiary RNA Structures by Differential SHAPE Reactivity

Kady-Ann Steen, Greggory M. Rice, and Kevin M. Weeks*
Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599-3290

* Correspondence, weeks@unc.edu

MATERIALS AND METHODS

RNA Constructs. DNA templates for the aptamer domains of the *Escherichia coli* thiamine pyrophosphate (TPP) riboswitch[1], *Vibrio vulnificus* adenine riboswitch[2], *Thermotoga maritima* lysine riboswitch[3], and the specificity domain of the *Bacillus subtilis* RNase P ribozyme[4], each imbedded within 5′ and 3′ structure cassette flanking sequences,[5] were generated by PCR as described.[6] RNAs were generated by *in vitro* transcription [1 mL; 40 mM Tris (pH 8.0), 10 mM $MgCl_2$, 10 mM dithiothreitol, 2 mM spermidine, 0.01% (v/v) Triton X-100, 4% (w/v) poly(ethylene) glycol 8000, 2 mM each NTP, 50 µL PCR-generated template, 0.1 mg/mL T7 RNA polymerase; 37 °C; 4 h]. RNAs were separated by denaturing polyacrylamide gel electrophoresis [8% polyacrylamide, 7 M urea, 29:1 acrylamide:bisacrylamide, 0.4 mm × 28.5 cm × 23 cm; 32 W, 1.5-2.5 h (depending on the length of the RNA)], excised from the gel, recovered by overnight passive elution at 4 °C, and precipitated with ethanol. The purified RNAs were resuspended in 50 µL TE and stored at -20 °C.

RNA Folding and SHAPE Modification. The adenine, lysine, and TPP riboswitch RNA constructs (5 pmol in 5 µL 1/2× TE) were denatured at 95 °C for 2 min, cooled on ice, treated with 3 µL 3.3× folding buffer [333 mM HEPES (pH 8.0), 333 mM NaCl (333 mM KCl for lysine riboswitch), 33.3 mM $MgCl_2$], and incubated at 37 °C for 10 min. The corresponding ligand (1 µL; 50 µM adenine, lysine, or TPP) was added, and samples were incubated at 37 °C for 20 min. No added ion experiments with the TPP riboswitch (shown in Figure 5) were performed identically except that the folding buffer contained only HEPES (pH 8.0). For RNase P, 5 pmol in 6 µL 1/2× TE were denatured at 95 °C for 2 min, cooled on ice, treated with 3 µL 3.3× folding buffer [333 mM HEPES (pH 8.0), 333 mM NaCl, 200 mM $MgCl_2$], and incubated at 37 °C for 20 min. After folding, RNAs were treated with reagent [1 µL; 80 mM for the RNase P, adenine, and TPP riboswitch RNAs, 50 mM for the lysine riboswitch); in anhydrous DMSO] and allowed to react at 37 °C for 3 min [1-methyl-6-nitroisatoic anhydride (1M6)], 7 min [1-methyl-6-bromoisatoic anhydride (1M6Br)], 22 min [N-methylisatoic anhydride (NMIA)], or 26 min [1-methyl-6-methylisatoic anhydride (1M6M)]. No-reagent control reactions were performed with 1 µL neat DMSO. The RNAs were recovered by ethanol precipitation and resuspended in 10 µL 1/2× TE.

Primer Extension and Data Analysis. The general procedure was outlined previously.[5,6] Resuspended RNA from the modification reaction was added to a fluorescently labeled DNA primer (5′-VIC-labeled GAA CCG GAC CGA AGC CCG; 3 µL, 0.3 µM) and annealed at 65 °C for 6 min and then cooled on

ice. Reverse transcription buffer [6 μL; 167 mM Tris (pH 8.3), 250 mM KCl, 10 mM $MgCl_2$, 1.67 mM each dNTP] and Superscript III (1 μL, 200 units) were added, and the reaction was incubated at 45 °C for 2 min, 52 °C for 20 min, then 65 °C for 5 min. For the lysine riboswitch, the primer was allowed to anneal at 65 °C for 6 min and then at 35 °C for 5 min prior to cooling on ice; reverse transcription buffer was added as above, and the reaction was incubated at 40 °C for 2 min, 52 °C for 1 h, then 65 °C for 5 min. All reactions were quenched with 4 μL 50 mM EDTA. The cDNAs were recovered by ethanol precipitation, washed twice with 70% ethanol, dried and resuspended in 10 μL deionized formamide. Dideoxy sequencing ladders were produced using unlabeled, unmodified RNA, annealing a fluorescently labeled DNA primer (5′-NED-labeled GAA CCG GAC CGA AGC CCG; 3 μL, 0.3 μM) and by adding 1 μL 2′,3′-dideoxycytosine triphosphate (10 mM for the adenine and TPP riboswitches and the RNase P RNA; 30 mM for the lysine riboswitch) before addition of Superscript III. The cDNA fragments were separated by capillary electrophoresis using Applied Biosystems 3130 or 3500 DNA sequencing instruments. Raw capillary electrophoresis traces were analyzed using ShapeFinder.[7] SHAPE reactivities for each experiment were obtained by subtracting the no-reagent background integrated areas from the plus-reagent integrated areas. Datasets were normalized by excluding the top 2% of the reactive nucleotides, and dividing all nucleotides by the average of the next 10% of reactive nucleotides. Each experiment was performed in triplicate and the mean SHAPE reactivity for each nucleotide for each reagent was determined.

Differential SHAPE Analysis. After obtaining mean, normalized SHAPE reactivities for both the 1M6 and NMIA reactions, 1M6 reactivities were subtracted from the NMIA reactivities to obtain the mean difference at each nucleotide position. In all cases (Figures 2, 5 and S3), differential reactivities with a mean absolute reactivity difference of ≥ 0.3 SHAPE units *and* a p-value < 0.05 (calculated using the Student's t-test) were taken to be statistically significant.

We emphasize that there are clearly many subtle electrostatic and chemical features of RNA structure that can contribute to the 2'-hydroxyl reactivity of SHAPE reagents. However, to a first-order approximation, preferential reactivities towards NMIA and 1M6 report nucleotides that experience slow local dynamics (and are typically C2'-endo) or that present one face of the nucleobase for favorable stacking interactions, respectively.

Electronic Structure Calculations. Models of SHAPE reagents and nucleotides were created separately using Avogadro.[8] For each RNA nucleotide type, we focused on the most common ribose pucker (C3'- or C2'-endo) as observed in the set of 18 nucleotides preferentially reactive towards 1M6 among the four crystal structures analyzed in this work (Figures 2 and S3). A simplified nucleic acid model system consisting of a nucleoside with a phosphate at the 3′-position and an alcohol at the ribose 5′-position was created for each of the four RNA nucleobase types. Phosphate charges were neutralized with the addition of a hydrogen atom to simplify gas-phase optimization. Nucleotide and SHAPE reagent complexes were based on models of the 1M6 stacking interaction with C24 in the TPP riboswitch.[1] The orientation of the SHAPE reagent was chosen to maximize overlap between the ring systems and substituents between the molecules, which has been shown previously to strengthen stacking interactions.[9] Structures were optimized using the default approach implemented in the Gaussian 09 package[10] in gas phase using the M06-2X/6-311G* functional and basis set. The M06-2X functional has been shown to be sufficiently robust to model stacking interactions in aromatic systems[11] and to consistently recover the CCSD(T) CBS

π-π interaction energy.[12] The 6-311G* basis set was chosen for its computational efficiency for a system of this size. Once models were optimized, a high-level single-point energy calculation was performed using M06-2X/6-311+G(2d,p) to obtain accurate structure energies. Interaction energies for different SHAPE complexes were calculated by subtracting the self-consistent field method (SCF) energies of the individual nucleotide and SHAPE structures from that of the complex to obtain the energy difference at infinite distance: $E_{\text{interaction}} = E_{\text{complex}} - E_{\text{reagent}} - E_{\text{nucleoside}}$.

REFERENCES

(1) Serganov, A.; Polonskaia, A.; Phan, A. T.; Breaker, R. R.; Patel, D. J. *Nature* **2006**, *441*, 1167-1171.

(2) Serganov, A.; Yuan, Y. R.; Pikovskaya, O.; Polonskaia, A.; Malinina, L.; Phan, A. T.; Hobartner, C.; Micura, R.; Breaker, R. R.; Patel, D. J. *Chem. Biol.* **2004**, *11*, 1729-1741.

(3) Serganov, A.; Huang, L.; Patel, D. J. *Nature* **2008**, *455*, 1263-1267.

(4) Krasilnikov, A. S.; Yang, X.; Pan, T.; Mondragon, A. *Nature* **2003**, *421*, 760-764.

(5) Wilkinson, K. A.; Merino, E. J.; Weeks, K. M. *Nat. Protoc.* **2006**, *1*, 1610-1616.

(6) Steen, K. A.; Malhotra, A.; Weeks, K. M. *J. Am. Chem. Soc.* **2010**, *132*, 9940-9943.

(7) Vasa, S. M.; Guex, N.; Wilkinson, K. A.; Weeks, K. M.; Giddings, M. C. *RNA* **2008**, *14*, 1979-1990.

(8) *Avogadro: an open-source molecular builder and visualization tool. Version 1.0.1.*

(9) Florián, J.; Šponer, J.; Warshel, A. *J. Phys. Chem. B* **1999**, *103*, 884-892.

(10) Frisch, M. J., et al.; Revision B.01 ed.; Gaussian, Inc.: Wallingford CT, 2009.

(11) Churchill, C. D.; Wetmore, S. D. *Phys. Chem. Chem. Phys.* **2011**, *13*, 16373-16383.

(12) Rutledge, L. R.; Wetmore, S. D. *Can. J. Chem.* **2010**, *88*, 815-830.

(13) Hansch, C.; Leo, A. *Substituent constants for correlation analysis in chemistry and biology*; Wiley-Interscience: NY, 1979.

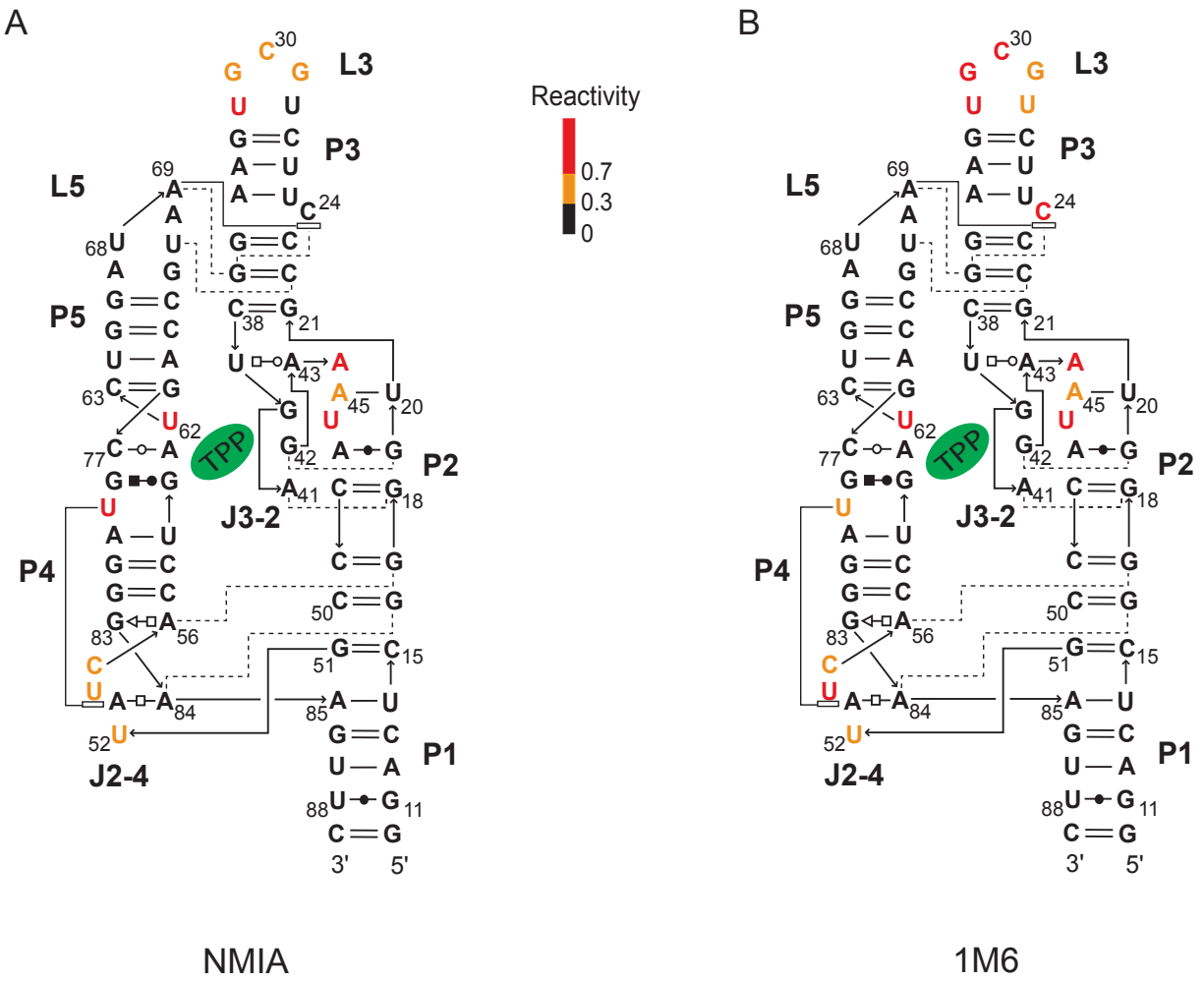
(14) Gherghe, C. M.; Mortimer, S. A.; Krahn, J. M.; Thompson, N. L.; Weeks, K. M. *J. Am. Chem. Soc.* **2008**, *130*, 8884-8885.

FIGURE LEGENDS

Figure S1. Absolute SHAPE reactivities for the ligand-bound TPP riboswitch resulting from reaction with (A) NMIA and (B) 1M6 superimposed on secondary structure models of the folded RNA.

Figure S2. Effect of varying the electron-withdrawing substituent on the SHAPE reactivity at nucleotides with one face of the base available for a stacking interaction. Increasing the electron-withdrawing ability of the functional group results in a more electrophilic reactive center and changes in the overall electrostatic profile of the reagent. (A) Electrostatic potential maps[8] for each reagent. Sites of substituent substitution and reactivity are shown with an R and an arrow, respectively. (B) Correlation between SHAPE reactivity at C24 (which presents an available open stacking face; see Figure 3B) and A45 (which has pre-existing stacking interactions at both nucleobase faces) in the TPP riboswitch and the electron withdrawing potential as measured by the Hammett coefficient (σ_m)[13] of the reagent R-group. The observation of an increase in reactivity as a function of R-group electron withdrawing potential with C24 but not with A45 supports a model in which enhanced reactivity at C24 reflects favorable stacking interactions rather than simply a general increase in reagent reactivity.

Figure S3. Nucleotides showing differential reactivities illustrated in the context of the three-dimensional structures (left) and the corresponding differential reactivity plots (right) for the (A) adenine riboswitch (PDB ID 1y26[2]), (B) lysine riboswitch (PDB ID 3dil[3]), and (C) specificity domain of RNase P (PDB ID 1nbs[4,14]). Nucleotides displaying strong differential reactivities are numbered explicitly.

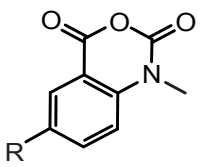


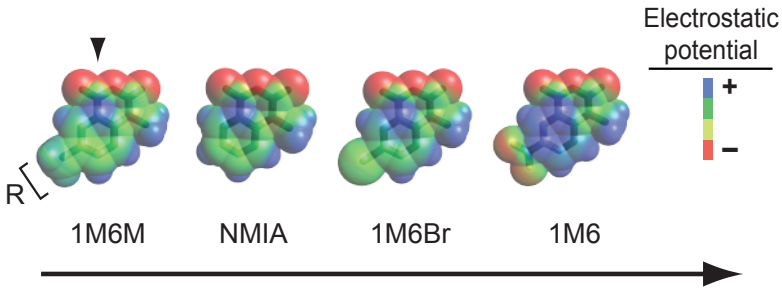
Steen et al., Figure S1

A

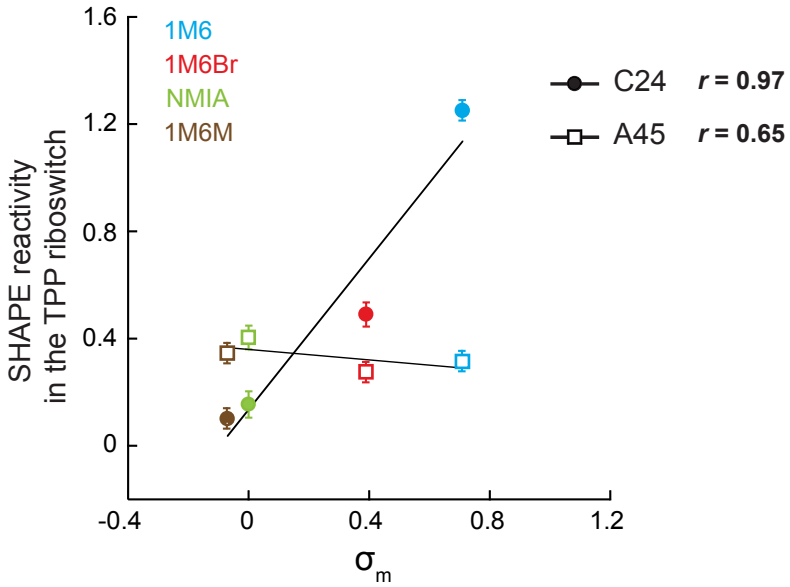
Reagent	hydrolysis $t_{1/2}$ (s)	R
1M6M	304	CH_3
NMIA	260	H
1M6Br	76	Br
1M6	31	NO_2

Increasingly electron-withdrawing

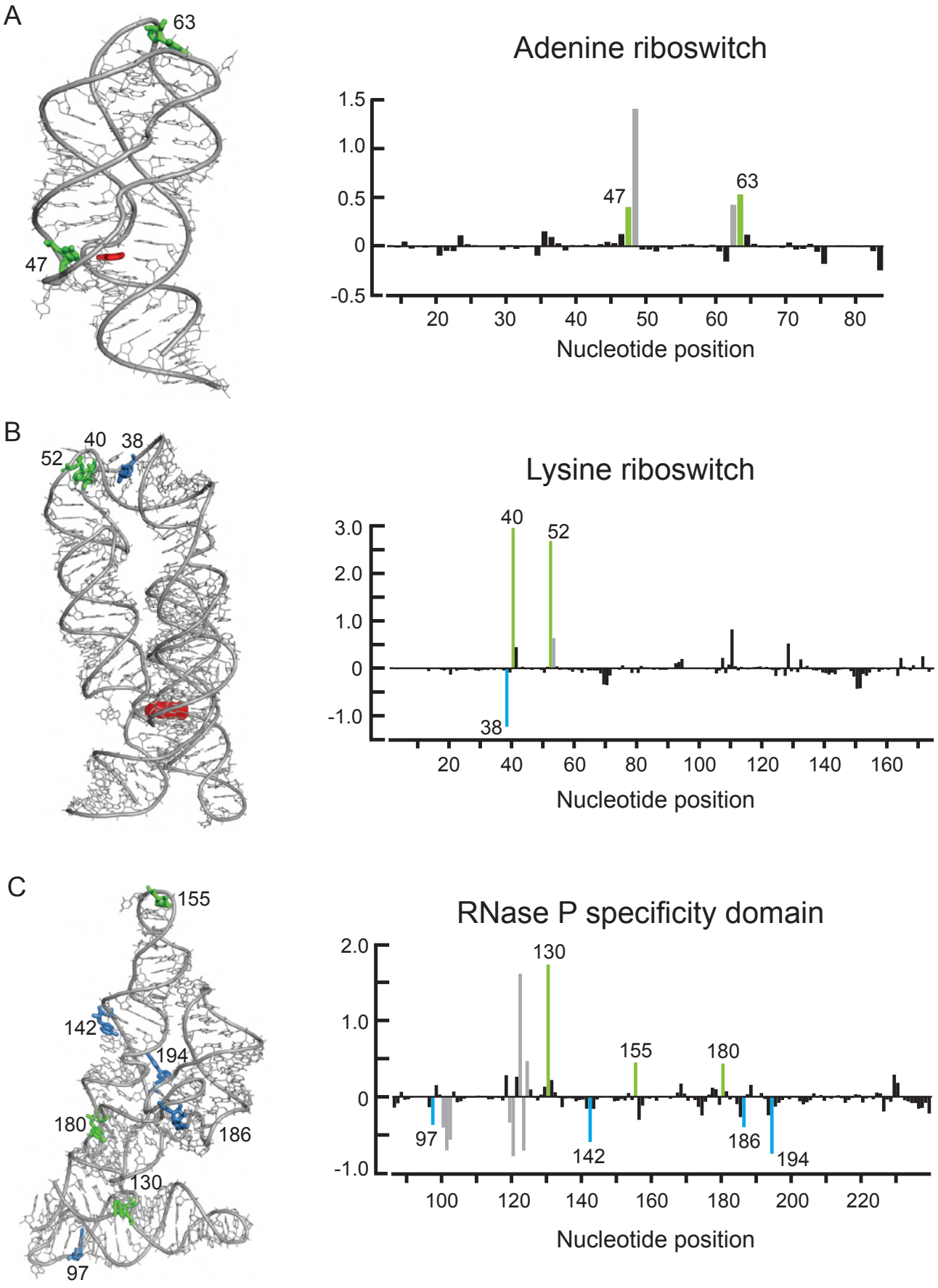




B



Steen et al., Figure S2



Steen et al., Figure S3