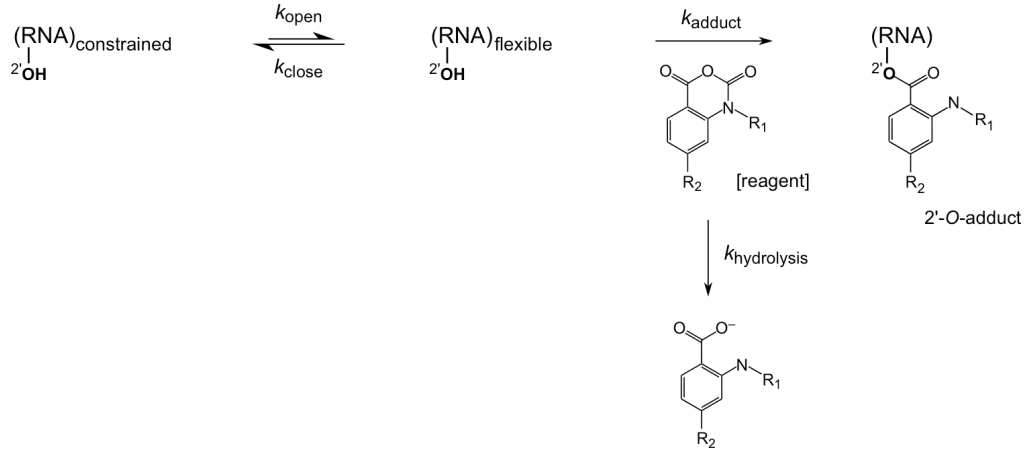
Supporting Information for

Slow Conformational Dynamics at C2'-endo Nucleotides in RNA

C.M. Gherghe, S.A. Mortimer, J.M. Krahn, N.L. Thompson and K.M. Weeks

Derivation of Eqn. 1. The electrophile-dependent reaction of RNA to form a 2'-*O*-adduct involves the following mechanism and four relevant rate constants:



Scheme 1

The observed rate of 2'-*O*-adduct formation is given by:

$$k_{obs} = \frac{k_{open} k_{adduct}[reagent]}{k_{open} + k_{close} + k_{adduct}[reagent]}$$

then,[1]

$$f \approx 1 - e^{-\frac{k_{obs}}{k_{hydrolysis}}(1-e^{-k_{hydrolysis}t})}$$

where f is the fraction 2'-*O*-adduct formed at any given nucleotide.

If the reaction is allowed to proceed until reagent hydrolysis is complete ($t \rightarrow \infty$), this equation simplifies to:

$$f \approx 1 - e^{-\frac{k_{obs}}{k_{hydrolysis}}}$$

as given in Eqn. 1.

IA, NMIA, 4NIA and 1M7 hydrolysis. Hydrolysis was followed by adding reagent (2.0 mM IA, 1.5 mM NMIA, 2.5 mM 4NIA, or 2.0 mM 1M7 in 300 μL DMSO) to 1.1× buffer [2.7 mL, 6.7 mM $MgCl_2$, 111 mM NaCl, 111 mM HEPES (pH 8.0)] equilibrated at 37 °C in a cuvette. Pseudo-first-order rates were obtained by monitoring the absorbance of the hydrolysis product (at 345 nm for 2-aminobenzoate, 360 nm for 2-methylaminobenzoate, 440 nm for 2-amino-4-nitrobenzoate, and 430 nm for 2-methylamino-4-nitrobenzoate).

RNA constructs. The C2'-endo hairpin RNA (Figure 2A) and the specificity domain of the RNase P RNA[2] were synthesized by *in vitro* transcription using a single stranded DNA (IDT) or a PCR-generated

template,[3] respectively. In both cases, the RNAs were embedded in the context of 5' and 3' structure cassette[4] sequences. RNAs were purified by denaturing polyacrylamide gel electrophoresis, excised from the gel, and recovered by electroelution and ethanol precipitation. Purified RNAs were resuspended in TE [10 mM Tris (pH 8.0), 1 mM EDTA] at concentrations of about 30 μM and stored at -20 °C.

SHAPE analysis. pAp-ethyl was 5'-end radiolabeled using γ-[^{32}P]-ATP, purified by denaturing gel electrophoresis, and resuspended in 1/2× TE. The pAp-ethyl (1 μL, 10000 cpm) was heated to 95 °C for 2 min, cooled on ice, mixed with 3 μL of 3.3× folding buffer [264 mM NaCl, 66 mM Hepes (pH 8.0), 16.5 mM $MgCl_2$], and incubated at 37 °C for 20 min. The pAp-ethyl solution was treated with reagent (1 μL; 100 mM; in anhydrous DMSO), allowed to react for 36 min (equal to five IA hydrolysis half-lives[1]). The no-reagent control contained 1 μL neat DMSO. The C2'-endo RNA (4 pmol) SHAPE experiments were performed as described for pAp-ethyl with the addition of a primer extension step. Modified RNA was recovered by ethanol precipitation [90 μL sterile H_2O, 5 μL NaCl (4 M), 1 μL glycogen (20 mg/mL), 400 μL ethanol; 30 min at -80 °C] and resuspended in 10 μL of TE. Analysis of the RNase P RNA was performed similarly except that the 3.3× folding buffer contained 333 mM Hepes (pH 8.0), 333 mM NaCl, and 33.3 mM $MgCl_2$.

Primer Extension. For the C2'-endo containing hairpin RNA, the primer extension reaction was performed using a 5'-[^{32}P]-label primer as described,[4] with the exception that the extension reaction was incubated at 45 °C for 1 min, 52 °C for 30 min, and 65 °C for 5 min. For the RNase P RNA, primer extension experiments were performed exactly as described,[3] with the exception that the DNA primers (5'-GAA CCG GAC CGA AGC CCG-3') were labeled with either VIC or NED for the (–) and (+) reagent experiments; dideoxy sequencing markers were generated using unmodified RNA and primers labeled with 6-FAM or PET. cDNA extension products were separated by capillary electrophoresis using an Applied Biosystems 3130 Genetic Analyzer capillary electrophoresis instrument.

Data Analysis. For pAp-ethyl experiments, band intensities were quantified by phosphorimaging (Molecular Dynamics Storm instrument).

For the C2'-endo RNA, individual band intensities for the (+) and (–) reagent reactions were integrated using SAFA,[5] as described.[4] Fraction adduct formation at each position was calculated as the band intensity divided by the full-length band. Reactivities were normalized to the intensity at position 80, a flexible nucleotide outside the RNA of interest. k_{open} was determined by fitting the fraction 2'-*O*-adduct using Eqn. 4 (given in the main text) with the addition of a plateau term:

$$f = 1 - e^{-k_{open}/k_{hydrolysis}} + A$$

where A accounts for the small amount of misfolded RNA in which the C2'-endo nucleotide exists in a uniformly reactive conformation. The addition of the A term improves the quality of the fit, but has no material effect on the estimate for k_{open}, within error. Least squares fitting was performed with Kaleidagraph (ver. 4.01).

For the RNase P RNA, raw traces from the ABI 3130 were processed using ShapeFinder[6] as described.[3] Overall reactivities for the IA and 1M7 experiments were normalized to intensities at positions 103 and 122; negative intensities were set to zero. Structure figures were composed with Pymol.[7]

Refinement of the RNase P Structure. The 3.15 Å resolution structure of RNase P was refined from the deposited model and structure factors (PDB accession 1NBS)[2] with careful treatment of nucleotide conformations. The new model includes several regions characterized by ambiguous experimental density in order to reduce overall map noise. Electron density maps from the original model contained large

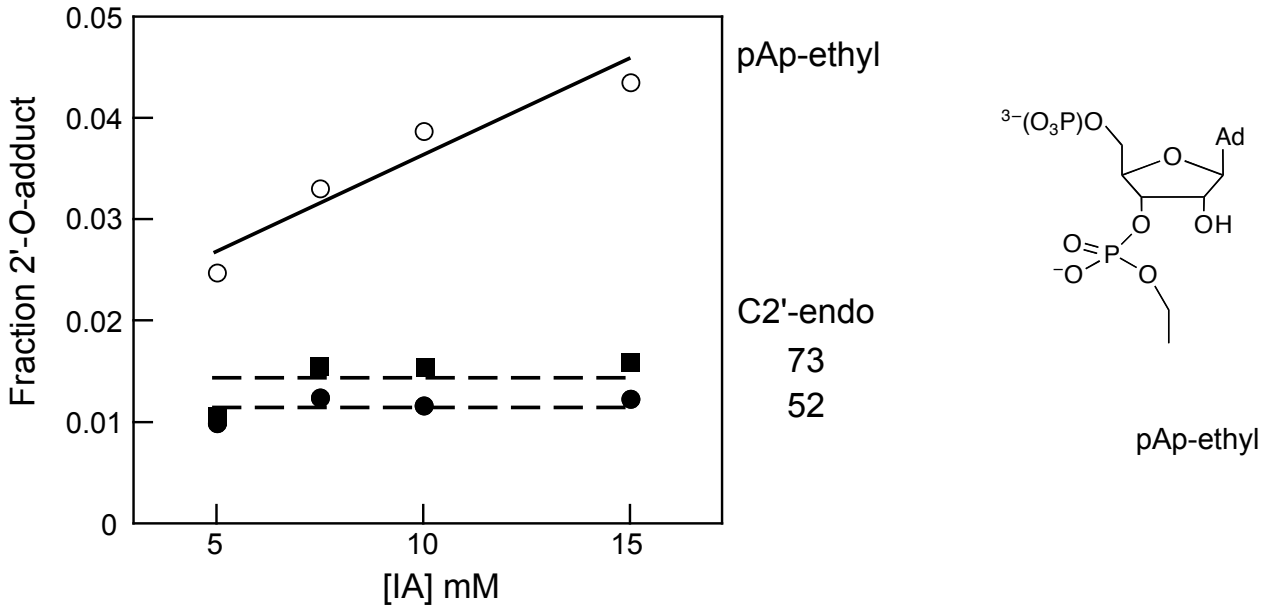
difference peaks in the |Fo|–|Fc| maps. Many of these sites had been modeled as Mg^{2+}, which do not adequately account for the density. Forty-seven sites were modeled as Sr^{2+} ions, which were a component of the crystallization solution. In addition, bulk solvent parameters were adjusted empirically, which helps to account for the relatively high level of strongly diffracting solution ions. Refinement with CNS 1.1 (ref. 8) resulted in a significant improvement in both the crystallographic R-factors and map interpretability. Some of the ion sites may in fact reflect partial occupancy of Pb^{2+}, or a mixture of ions, but strontium is a good fit to the density in most cases. The final model includes 11 chloride ions, 16 magnesium ions, and 15 waters based on charge and density considerations. At this point, the model could be extended to include 16 of the 40 missing nucleotides in poorly ordered regions. Inclusion of these nucleotides reduces overall noise in the electron density map.

This model was further improved by including non-crystallographic symmetry and geometric restraints to idealize ribose geometries in regions of the structure that formed regular, well-defined, A-form RNA helices. Refinement restraints for other regions were initially modified to allow unrestrained pucker conformations, and later restrained to either C3'-endo or C2'-endo based on electron density, geometric considerations, and analysis of the distance between the plane of the nucleobase and the phosphate using MolProbity.[9,10]

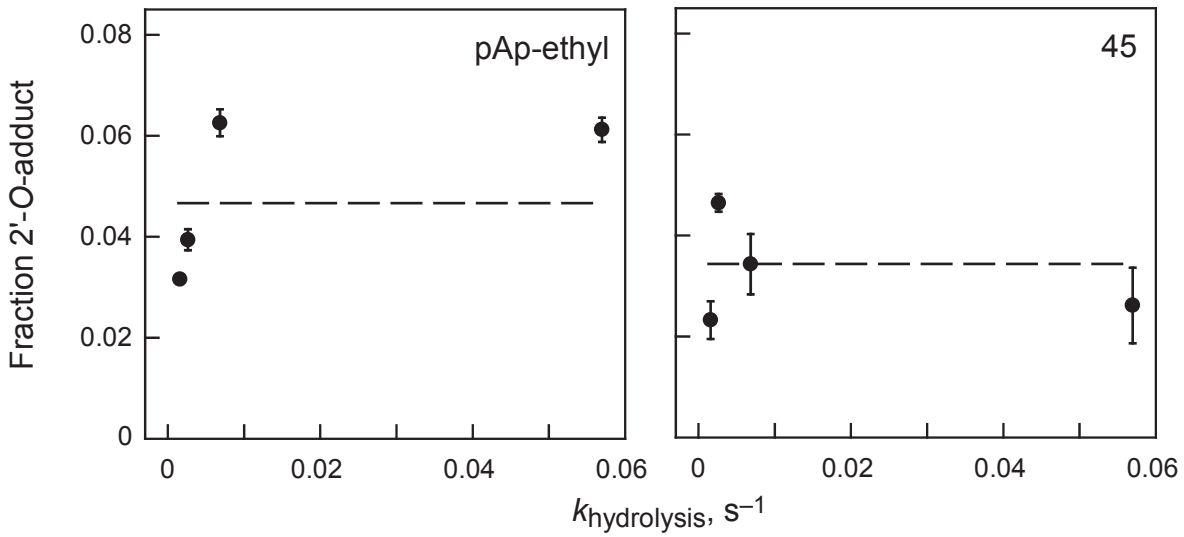
Improvement to the free R-factor[11] was used to guide the steps outlined above, using the original test set from the deposited structure factors in all R-free calculations. The working and free R-factors of 28.0 and 30.7% from the original model were improved to 20.3 and 24.1%, respectively. Nucleotides categorized as C2'-endo (in well-defined regions of the experimental electron density map and excluding those involved in crystal contacts) are 90, 130, 134, 148, 160, 167, 168, 191, 192, 196 (see Figure 4A).

References

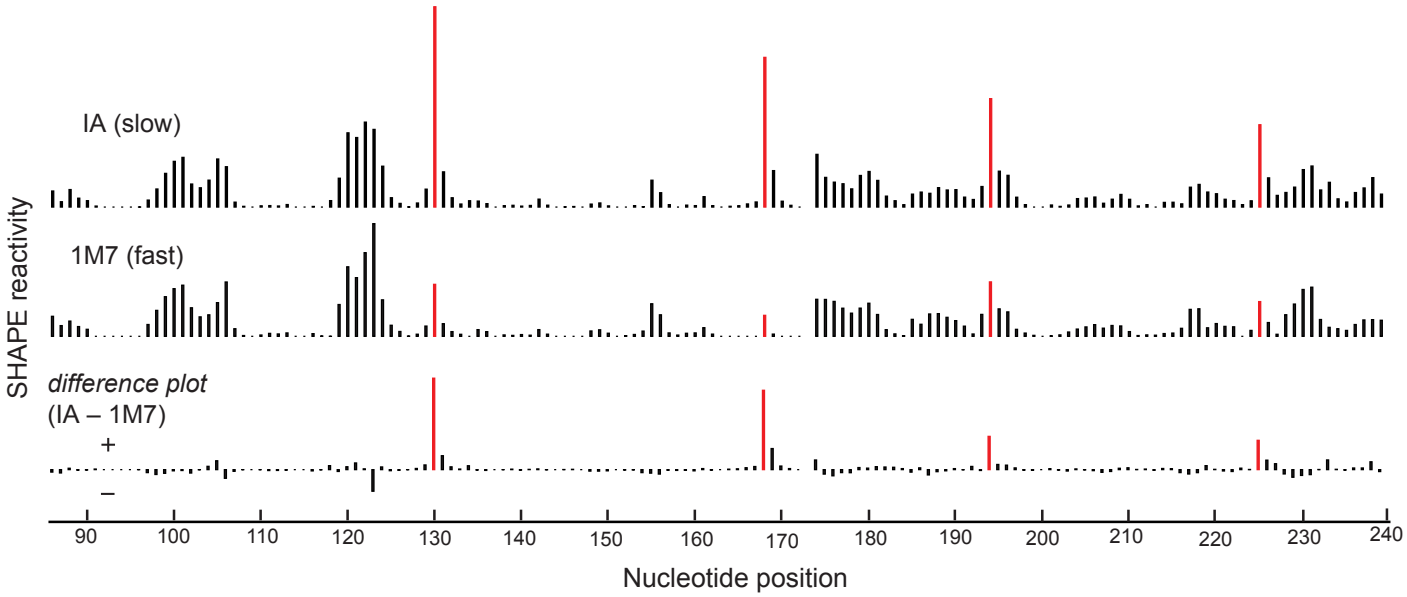
(1) Merino, E. J.; Wilkinson, K. A.; Coughlan, J. L.; Weeks, K. M. *J. Am. Chem. Soc.* **2005**, *127*, 4223-4231.
(2) Krasilnikov, A. S.; Yang, X.; Pan, T.; Mondragon, A. *Nature* **2003**, *421*, 760-764.
(3) Mortimer, S. A.; Weeks, K. M. *J. Am. Chem. Soc.* **2007**, *129*, 4144-4145.
(4) Wilkinson, K. A.; Merino, E. J.; Weeks, K. M. *Nature Protocols* **2006**, *1*, 1610-1616.
(5) Das, R.; Laederach, A.; Pearlman, S. M.; Herschlag, D.; Altman, R. B. *RNA* **2005**, *11*, 344-354.
(6) Wilkinson, K. A.; Gorelick, R. J.; Vasa, S. M.; Guex, N.; Rein, A.; Mathews, D. H.; Giddings, M. C.; Weeks, K. M. *PLoS Biology* **2008**, *6*, e96.
(7) DeLano, W. L. The Pymol Molecular Graphics System, DeLano Scientific, South San Francisco, CA, USA. www.pymol.org.
(8) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta. Cryst.* **1998**, *54*, 905-921.
(9) Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B.; Snoeyink, J.; Richardson, J. S.; Richardson, D. C. *Nucl. Acids Res.* **2007**, *35*, W375-383.
(10) Richardson, J. S.; Schneider, B.; Murray, L. W.; Kapral, G. J.; Immormino, R. M.; Head, J. J.; Richardson, D. C.; Ham, D.; Hershkovits, E.; Williams, L. D.; Keating, K. S.; Pyle, A. M.; Micallef, D.; Westbrook, J.; Berman, H. M. *RNA* **2008**, *14*, 465-481.
(11) Brunger, A. T. *Acta Cryst.* **1993**, *D49*, 24-36.
(12) Tolbert, B. S.; Kennedy, S. D.; Schroeder, S. J.; Krugh, T. R.; Turner, D. H. *Biochemistry* **2007**, *46*, 1511-1522.



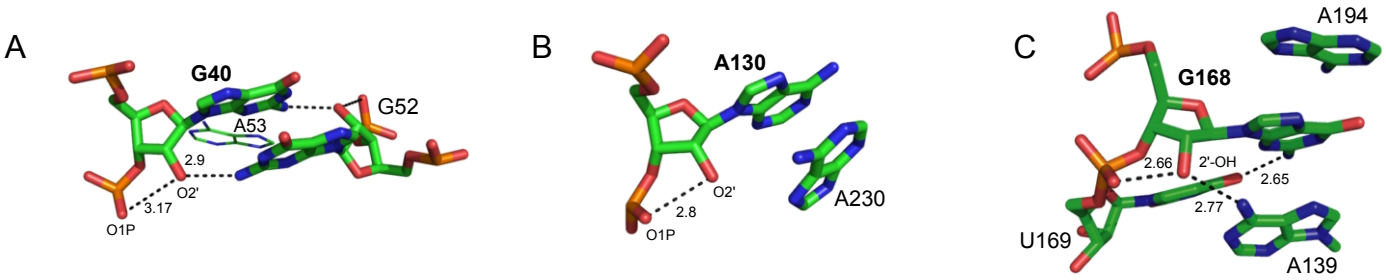
Supporting Figure 1. Concentration dependence for reaction at positions 52 and 73 in the C2'-endo RNA construct and for the (unconstrained) model nucleotide pAp-ethyl.



Supporting Figure 2. Absence of a dependence of fraction adduct formed as a function of $k_{hydrolysis}$ for reaction of the (unconstrained) model nucleotide pAp-ethyl and for nucleotide 45 in the loop of the C2'-endo construct.



Supporting Figure 3. Absolute SHAPE reactivities and difference plot for RNase P using slow (IA) and fast (1M7) reagents. Red columns indicate nucleotide positions that show the greatest difference in reactivity between the two reagents. Reactivites at all other positions are identical within experimental error.



Supporting Figure 4. Base stacking and hydrogen bonding interactions at C2'-endo nucleotides that undergo slow local conformational changes. (A) C40 from the C2'-endo RNA construct[12] and (B) A130 and (C) G168 in the RNase P RNA. Distances are in Å.