J|A|C|S
COMMUNICATIONS

Published on Web 06/18/2008

Slow Conformational Dynamics at C2′-endo Nucleotides in RNA

Costin M. Gherghe, Stefanie A. Mortimer, Joseph M. Krahn, Nancy L. Thompson, and Kevin M. Weeks*
Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290

Received April 11, 2008; E-mail: weeks@unc.edu

Local and global dynamics in folded RNAs occur over broad timescales spanning picoseconds to minutes.[1] Slow motions likely play predominant roles in governing RNA folding and ribonucleoprotein assembly reactions. However, slow local motions are extremely difficult to detect, especially for large RNAs with complex structures.

The local environment and degree of flexibility can be evaluated at nucleotide resolution for RNAs of any size using selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) chemistry.[2] RNA nucleotides exist in equilibrium between constrained (closed) and flexible (open) states. The 2′-OH group in flexible nucleotides preferentially adopts an open, reactive, conformation that facilitates reaction with electrophilic reagents to form a 2′-*O*-adduct (Figure 1). SHAPE experiments work well using electrophiles based on the isatoic anhydride (IA) scaffold.[2a,3] Positions that form 2′-*O*-adducts are detected by primer extension.[2]

IA derivatives both react with the RNA 2′-OH group and also undergo concurrent degradation by hydrolysis (Figure 1). 2′-OH reactivity is thus conveniently monitored by allowing a reaction to proceed until the reagent has been consumed, either by hydrolysis or reaction with RNA. At this end point, the fraction adduct at any nucleotide (f) is

$$f \approx 1 - e^{-(k_{obs}/k_{hydrolysis})} \quad (1)$$

where

$$k_{obs} = \frac{k_{open}k_{adduct}[reagent]}{k_{open} + k_{close} + k_{adduct}[reagent]} \quad (2)$$

and the rate of hydrolysis has been shown to be proportional to the rate of adduct formation,[2b,3] $k_{adduct}/k_{hydrolysis} = \beta$. These relationships lead to two limits. In limit 1, $k_{open} + k_{close} \gg k_{adduct}[reagent]$

$$f = 1 - e^{-k_{open}/(k_{open}+k_{close})\beta[reagent]} \quad (3)$$

For limit 2, $k_{open} + k_{close} \ll k_{adduct}[reagent]$

$$f = 1 - e^{-k_{open}/k_{hydrolysis}} \quad (4)$$

It should therefore be possible to monitor local nucleotide dynamics in RNA under conditions where limit 2 applies by varying the reactivity (or $k_{hydrolysis}$) of the hydroxyl-selective electrophile. IA has a hydrolysis half-life ($t_{1/2}$) of 430 s at 37 °C. Electron-withdrawing substituents at the cyclic amine (R_1) or in the benzene ring (R_2) enhance reagent reactivity. Compared to IA, *N*-methyl isatoic anhydride (NMIA), 4-nitroisatioc anhydride (4NIA), and 1-methyl 7-nitroisatoic anhydride (1M7)[3] have progressively shorter hydrolysis half-lives (table, Figure 1).

To investigate if distinct local nucleotide dynamics can be captured by varying the SHAPE electrophile, we focused on an important variation in RNA structure: the C2′-endo conformation. Although C2′-endo nucleotides are relatively rare, they are highly over-represented in important RNA tertiary interactions and in catalytic active sites.[4] Local structure at tandem G•A mismatches depends on the local sequence context.[5] Guanosine nucleotides in G•A pairs adopt the C2′-endo conformation in the sequences $(UGAA)_2$,[5a] and $(GGAU)_2$,[5b] the C3′-endo conformation typical of standard A-form helix geometry in $(CGAG)_2$,[5c] and a mixture of C2′-endo/C3′-endo conformations in

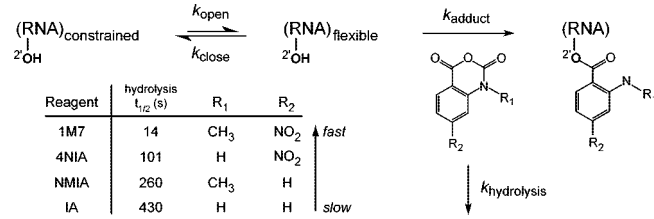


Figure 1. Mechanistic framework for RNA SHAPE chemistry.

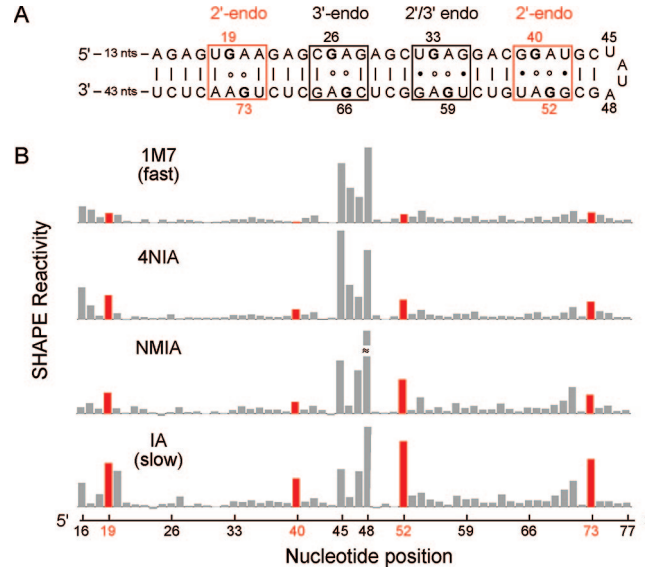


Figure 2. (A) The C2′-endo RNA construct contains tandem G•A base pairs that adopt either C2′- or C3′-endo ribose conformations. (B). Absolute SHAPE reactivities as a function of reagent electrophilicity.

$(UGAG)_2$.[5b] We constructed a simple hairpin RNA (termed the C2′-endo RNA) containing each of these sequences. Because the G•A mismatch-containing sequences are palindromic, there are two equivalent examples of each G•A pair in the RNA, including four total C2′-endo nucleotides (in red, Figure 2A).

When the C2′-endo RNA was subjected to SHAPE analysis using the fastest reagent, 1M7, flexible nucleotides in the apical loop (nts 45−48) are reactive, while positions constrained by base pairing are unreactive, regardless of the sugar pucker (top panel, Figure 2B). When an otherwise identical experiment was performed with IA, which reacts 30-fold more slowly, nucleotides in the apical loop were again reactive, while most of the base paired nucleotides were unreactive, similar to their reactivities with 1M7. In strong contrast, the four G nucleotides that adopt the C2′-endo conformation were highly reactive, even more so than some nucleotides in the flexible loop (Figure 2B, bottom panel; red bars at positions 19, 40, 52, 73). For the two reagents with intermediate reactivities, 4NIA and NMIA, the C2′-endo positions are moderately reactive (Figure 2B). Nucleotides constrained in these C2′-endo conformations are thus unreactive toward fast-reacting electrophiles but are highly reactive toward the slower reagents.

10.1021/ja802691e CCC: $40.75

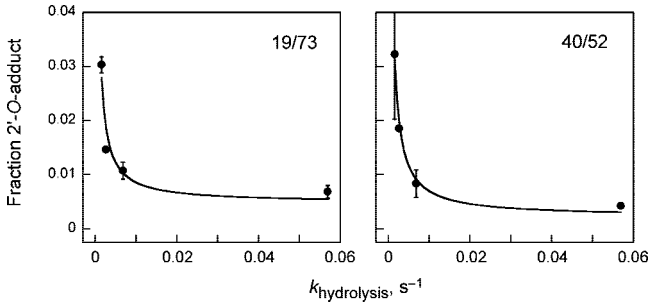


Figure 3. Determination of k_{open} for C2′-endo positions 19/73 and 40/52. Lines represent a fit to eq 4, with the addition of a plateau term; rate constants are 4 (±2) × 10^{-5} s^{-1}. Error bars indicate standard deviations from independent measurements.

Limit 1 (eq 3) predicts that adduct formation is a function of the equilibrium constant for formation of the open state [$k_{open}/(k_{open} + k_{close})$] and is independent of the reagent hydrolysis rate. Most nucleotides thus far analyzed by SHAPE, including in tRNA[2,6] and in an RNase P RNA,[3] are characterized by limit 1.

In contrast, the observation of a strong dependence of adduct formation on $k_{hydrolysis}$ suggests that limit 2 applies to the C2′-endo nucleotides in Figure 2. Limit 2 also implies (1) that the extent of reaction at C2′-endo nucleotides will be independent of reagent concentration and (2) that k_{obs} reports k_{open} (compare eqs 3 and 4). We analyzed the concentration dependence for reaction at positions 52 and 73 using isatoic anhydride and found, as predicted by limit 2, that adduct formation is independent of reagent concentration under conditions where reaction of the unconstrained model nucleotide, pAp-ethyl, showed a clear concentration dependence (Figure S1).

We therefore estimated k_{open} at the C2′-endo nucleotides at positions 19/73 and 40/52 by fitting the extent of 2′-*O*-adduct formation as a function of $k_{hydrolysis}$ to eq 4 (Figure 3). In both cases, k_{open} is 4 × 10^{-5} s^{-1}. In contrast, reactivities for both pAp-ethyl and flexible loop nucleotides are independent of $k_{hydrolysis}$ (Figure S2).

Critically, some C2′-endo nucleotides thus experience extraordinarily slow local dynamics to form conformations reactive toward isatoic anhydride-based electrophiles.

We next explored whether the differential reactivity between 1M7 and IA can be used to identify nucleotides that undergo slow conformational dynamics in an RNA with a complex structure, the specificity domain of the *Bacillus subtilis* ribonuclease P enzyme (RNase P).[7] After excluding nucleotides where the electron density was not well-defined or that participate in crystal contacts, we identified 10 C2′-endo nucleotides in the RNase P RNA (in color, Figure 4A).

For the vast majority of RNase P nucleotides, including all positions with C3′-endo conformations, SHAPE reactivities were identical for both the fast (1M7) and slow (IA) reagents (Figure S3). These nucleotides reflect limit 1. The 10 well-defined C2′-endo nucleotides fell into three categories: (i) most C2′-endo nucleotides are highly constrained and, as expected,[2a] unreactive toward both reagents (blue nucleotides, Figure 4A); (ii) one nucleotide is not constrained and is reactive toward both electrophiles (red nucleotide, Figure 4A); and (iii) two C2′-endo positions show large changes in reactivity (nts A130 and G168, circled nucleotides and red bars, Figure 4). Two other nucleotides showed smaller changes in reactivity but were in regions of the structure where experimental electron density was poorly defined (gray columns, Figure 4B). A similar distribution of reactive and unreactive C2′-endo nucleotides occurs in the *Tetrahymena* P5−P4−P6 domain using NMIA.[8]

While the C2′-endo conformation by itself clearly does not govern SHAPE reactivity, a distinct class of C2′-endo nucleotides experiences extraordinarily slow local dynamics. These nucleotides in both the simple C2′-endo RNA (Figure 2B) and the RNase P RNA (Figure 4B) share key characteristics: (1) the ribose group has the C2′-endo conformation, and (2) the nucleotide conformation is partially constrained by base stacking and hydrogen bonding interactions (Figure S4). These C2′-endo dynamics are orders of magnitude slower than for other local RNA conformational changes such as base opening reactions[9] and are also slower than folding processes that involve assembly of whole domains in large RNAs.[1a]

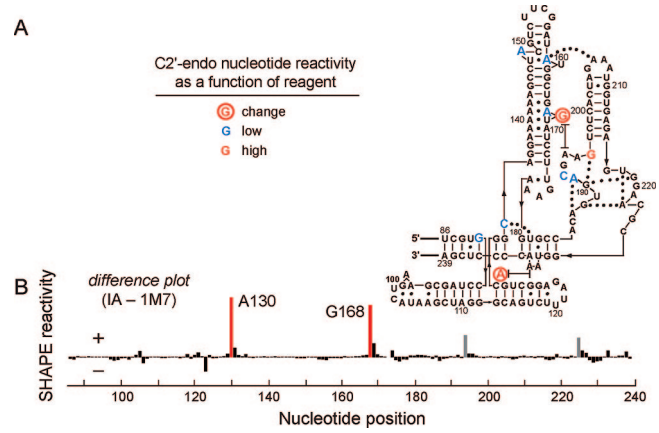


Figure 4. SHAPE reactivities at C2′-endo nucleotides in the specificity domain of RNase P. (A) Secondary structure showing SHAPE reactivities at the 10 well-defined C2′-endo nucleotides. (B) Differential SHAPE reactivities for slow (IA) versus fast (1M7) reagents.

An instructive precedent for slow conformational changes at a single residue in a biopolymer is the cis−trans isomerization of prolyl residues in proteins. Proline cis−trans conformations interconvert on the order of 10^{-2}−10^{-5} s^{-1},[10] and are thus comparable to the rates measured here for local dynamics at some C2′-endo nucleotides in RNA. The cis−trans isomerization can function as a molecular switch in biology.[10c]

We postulate that slow conformational dynamics at C2′-endo nucleotides also have the potential to function as rate-determining molecular switches and will play important, but currently unexplored, roles in RNA folding, ligand recognition, and catalysis.

Acknowledgment. We are indebted to Doug Turner and Blanton Tolbert for insightful discussions. This work was supported by grants from the NSF (MCB-0416941) and NIH (AI068462) to K.M.W.

Supporting Information Available: Methods and four figures. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) See Figure 12 of: (a) Buchmueller, K. L.; Hill, B. T.; Platz, M. S.; Weeks, K. M. *J. Am. Chem. Soc.* **2003**, *125*, 10850–10861. (b) Shajani, Z.; Varani, G. *Biopolymers* **2007**, *86*, 348–359.
(2) (a) Merino, E. J.; Wilkinson, K. A.; Coughlan, J. L.; Weeks, K. M. *J. Am. Chem. Soc.* **2005**, *127*, 4223–4231. (b) Wilkinson, K. A.; Merino, E. J.; Weeks, K. M. *J. Am. Chem. Soc.* **2005**, *127*, 4659–4667. (c) Wilkinson, K. A.; Merino, E. J.; Weeks, K. M. *Nat. Protoc.* **2006**, *1*, 1610–1616.
(3) Mortimer, S. A.; Weeks, K. M. *J. Am. Chem. Soc.* **2007**, *129*, 4144–4145.
(4) Richardson, J. S.; et al. *RNA* **2008**, *14*, 465–481.
(5) (a) Heus, H. A.; Wijimenga, S. S.; Hoppe, H.; Hilbers, C. W. *J. Mol. Biol.* **1997**, *271*, 147–158. (b) Tolbert, B. S.; Kennedy, S. D.; Schroeder, S. J.; Krugh, T. R.; Turner, D. H. *Biochemistry* **2007**, *46*, 1511–1522. (c) SantaLucia, J. J.; Turner, D. H. *Biochemistry* **1993**, *32*, 12612–12623.
(6) Wang, B.; Wilkinson, K. A.; Weeks, K. M. *Biochemistry* **2008**, *47*, 512–524.
(7) Krasilnikov, A. S.; Yang, X.; Pan, T.; Mondragon, A. *Nature* **2003**, *421*, 760–764.
(8) Vicens, Q.; Gooding, A. R.; Laederach, A.; Cech, T. R. *RNA* **2007**, *13*, 536–548.
(9) (a) Leroy, J. L.; Bolo, N.; Figueroa, N.; Plateau, P.; Gueron, M. *J. Biomol. Struct. Dyn.* **1985**, *2*, 915–939. (b) Spies, M. A.; Schowen, R. L. *J. Am. Chem. Soc.* **2002**, *124*, 14049–14053.
(10) (a) Grathwohl, C.; Wuthrich, K. *Biopolymers* **1981**, *20*, 2623–2633. (b) Schoetz, G.; Trapp, O.; Schurig, V. *Electrophoresis* **2001**, *22*, 2409–2415. (c) Lu, K. P.; Finn, G.; Lee, T. H.; Nicholson, L. K. *Nat. Chem. Biol.* **2007**, *3*, 619–629.

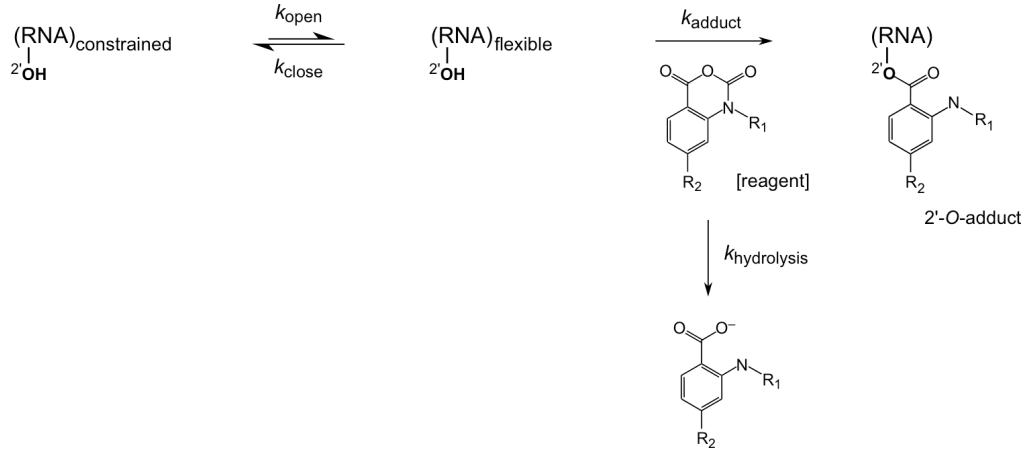
JA802691E

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

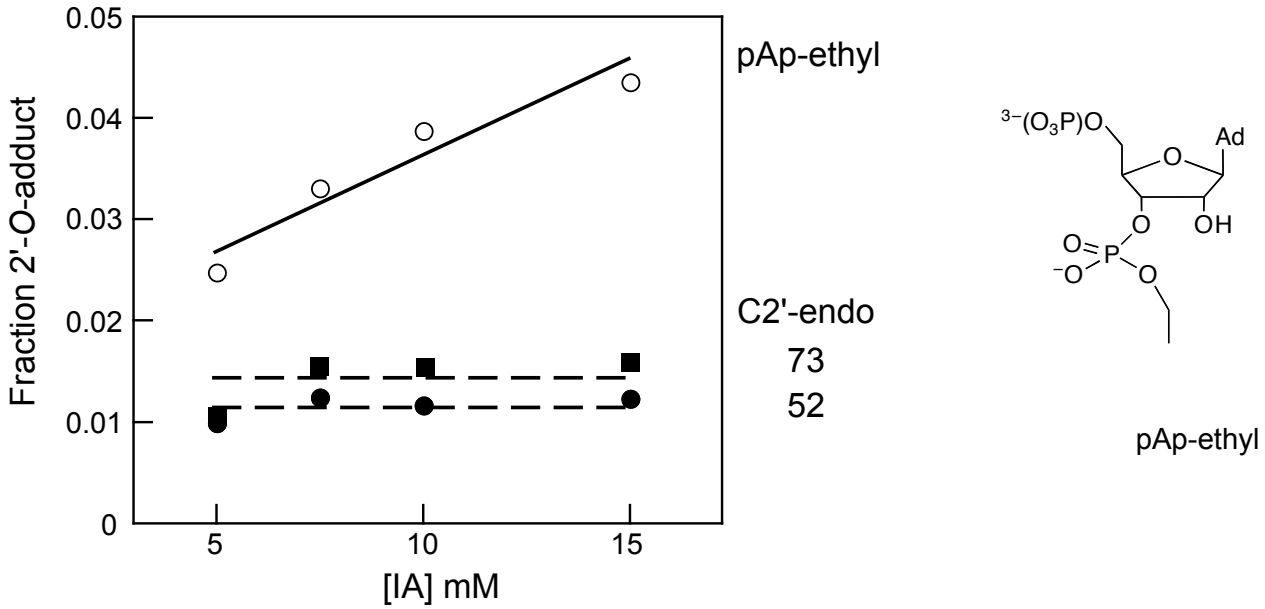
difference peaks in the |Fo|–|Fc| maps. Many of these sites had been modeled as Mg^{2+}, which do not adequately account for the density. Forty-seven sites were modeled as Sr^{2+} ions, which were a component of the crystallization solution. In addition, bulk solvent parameters were adjusted empirically, 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]

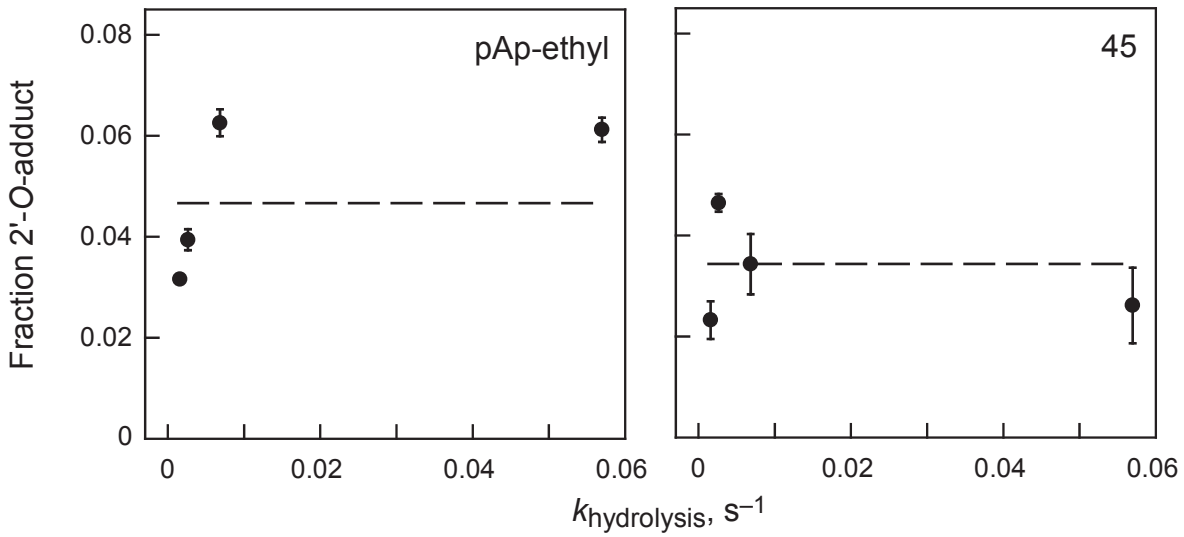
Improvement to the free R-factor[11] was used to guide the steps outlined above, using the original test set from the deposited structure factors in all R-free calculations. The working and free R-factors of 28.0 and 30.7% from the original model were improved to 20.3 and 24.1%, respectively. Nucleotides 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

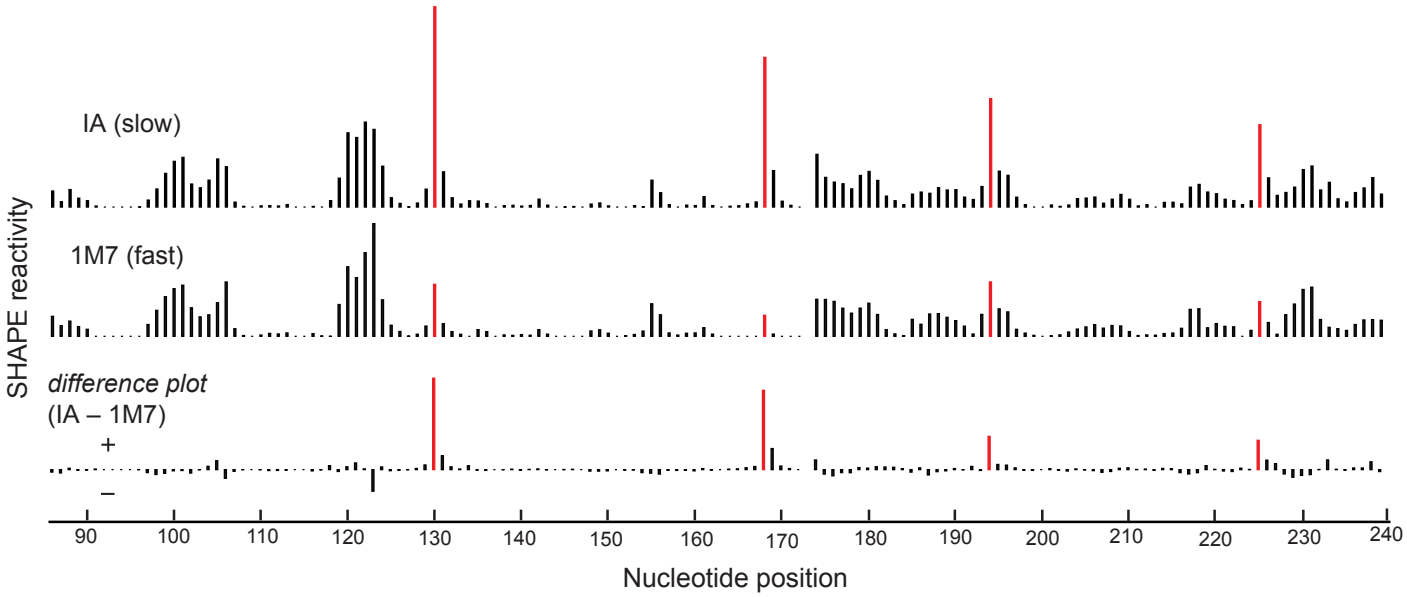
(1) Merino, E. J.; Wilkinson, K. A.; Coughlan, J. L.; Weeks, K. M. *J. Am. Chem. Soc.* **2005**, *127*, 4223-4231.
(2) Krasilnikov, A. S.; Yang, X.; Pan, T.; Mondragon, A. *Nature* **2003**, *421*, 760-764.
(3) Mortimer, S. A.; Weeks, K. M. *J. Am. Chem. Soc.* **2007**, *129*, 4144-4145.
(4) Wilkinson, 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.



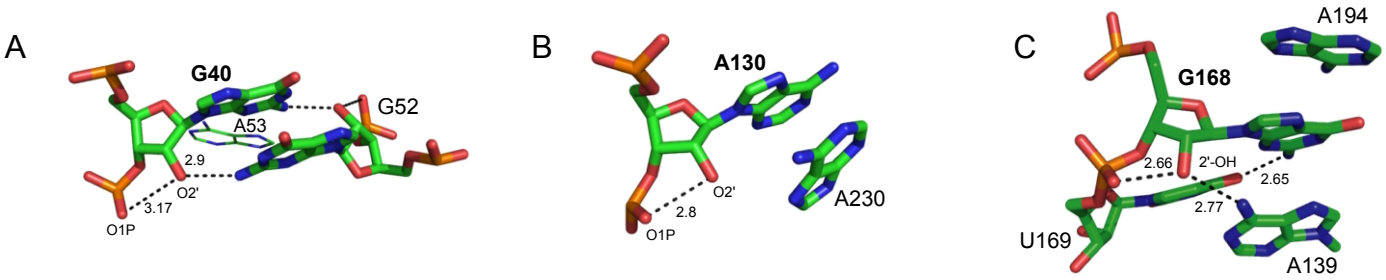
Supporting Figure 1. Concentration dependence for reaction at positions 52 and 73 in the C2'-endo RNA construct and for the (unconstrained) model nucleotide pAp-ethyl.



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.