J|A|C|S
COMMUNICATIONS

Published on Web 11/08/2008

Time-Resolved RNA SHAPE Chemistry

Stefanie A. Mortimer and Kevin M. Weeks*

Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290

Received August 4, 2008; E-mail: weeks@unc.edu

Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) chemistry is emerging as a powerful approach for quantitative analysis of the equilibrium structures of diverse biological RNAs.[1−3] SHAPE chemistry exploits the discovery that the nucleophilic reactivity of the ribose 2′-hydroxyl position is strongly gated by the underlying nucleotide flexibility (Figure 1). Flexible nucleotides preferentially adopt conformations that react with a hydroxyl-selective electrophile to form a 2′-*O*-adduct, while base paired or otherwise conformationally constrained nucleotides are unreactive. The electrophile is concurrently consumed by hydrolysis making a specific quench step unnecessary, provided the adduct-forming reaction with RNA is allowed to continue until all reagent is consumed.[1] SHAPE works extremely well for interrogating the equilibrium structure of RNA motifs from a few nucleotides to large catalytic RNAs and viral genomes spanning hundreds to thousands of nucleotides.[1−4]

To date, SHAPE chemistry has been based on electrophilic reagents derived from an isatoic anhydride (IA) scaffold. Reactivity can be tuned by varying the electron withdrawing potential of the functional groups on the IA scaffold. We have created a suite of reagents with hydrolysis half-lives from 14 to 430 s.[1,2,3c] In our view, the most reactive of these reagents, 1-methyl-7-nitroisatoic anhydride (1M7),[2] represents a near-ideal probe for nucleotide-resolution analysis of RNA structures at equilibrium. The intrinsic reactivity of 1M7 is insensitive to solution conditions like Mg^{2+} concentration and temperature. Complete degradation of the reagent requires 70 s: this moderate level of reactivity makes the reagent easy to handle in the laboratory but means that RNA structures are averaged on the minute time scale.[2]

A full understanding of structure−function relationships in RNA biology requires a nucleotide-resolution view of the time-resolved mechanisms by which RNA molecules fold, interconvert between distinct states, and function in ribonucleoprotein complexes.[5] The reactivity of IA derivatives is circumscribed by the electron withdrawing potential of the substituents that can be introduced into this reagent. The limits of this useful scaffold have likely been reached.

We therefore sought to develop a new chemical framework to interrogate RNA structure on the seconds time scale. Benzoyl cyanide (BzCN) reacts with hydroxyl functional groups to yield a stable ester. Formation of the stable cyanide ion leaving group renders the reaction irreversible (Figure 1). As with all SHAPE electrophiles, this scaffold should satisfy two important criteria: (1) react broadly with RNA 2′-hydroxyl groups in a structure-selective manner and (2) undergo rapid inactivation by hydrolysis with water (Figure 1).

We used BzCN to assess the structure of the specificity domain of *Bacillus subtilis* ribonuclease P (RNase P), whose structure has been well characterized using IA-based reagents.[2,3c] The RNase P RNA was treated with BzCN under conditions that stabilize the native tertiary fold (10 mM $MgCl_2$, 100 mM NaCl, pH 8.0) and also in the absence of Mg^{2+} where the RNA forms only its secondary structure. Sites of 2′-*O*-adduct formation were identified as stops to primer extension using fluorescently labeled DNA primers, resolved by capillary electrophoresis.[2,4] Absolute SHAPE reactivities observed for the RNase P RNA using BzCN are very similar to those obtained with 1M7. Both reagents react preferentially with loops and, in both cases, loops that participate in tertiary interactions are protected from forming 2′-*O*-adducts in the presence of Mg^{2+} (Figure 2). There is a very strong linear correlation (Pearson's $r \geq 0.80$) between 1M7 and BzCN reactivities, both in the presence and absence of Mg^{2+} (data not shown). Thus, like 1M7,[2] BzCN reactivity measures local nucleotide flexibility and is selective for the 2′-OH group. BzCN therefore meets the first of the two criteria for time-resolved SHAPE because it accurately reports the secondary and tertiary structure of the RNase P specificity domain in the presence and absence of Mg^{2+}.

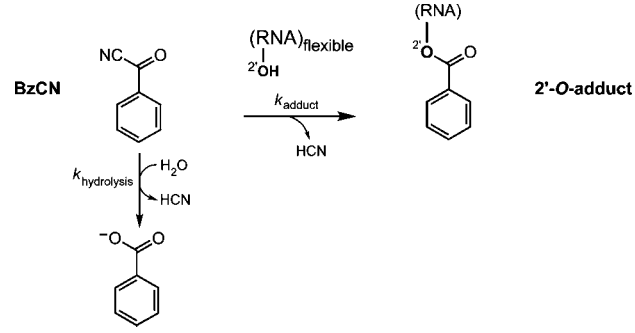

Figure 1. Mechanism of RNA SHAPE chemistry with BzCN. BzCN reacts both with 2′-hydroxyl groups at flexible positions in RNA to form a 2′-*O*-adduct and also undergoes inactivation by hydrolysis.

We next determined the rate of BzCN hydrolysis by monitoring the competition between hydrolysis versus 2′-*O*-adduct formation using the model nucleotide, 3′-phosphoethyl-5′-adenosine monophosphate (pAp-ethyl) (Figure 3A). Hydrolysis is rapid and was quenched at defined time points with dithiothreitol (see filled circle, Figure 3B). The rate constant for BzCN hydrolysis was determined by fitting the fraction adduct formed as a function of reaction time (Figure 3B). BzCN undergoes hydrolysis with a half-life of 0.25 s at 37 °C (Figure 3A).

The reaction between BzCN and RNA is complete in ∼1 and 2.5 s at 37 and 25 °C (Figure 3). BzCN is thus an ideal reagent for following RNA folding on the seconds time scale.

We then used BzCN to analyze the folding behavior of the RNase P RNA in a time-resolved SHAPE experiment. The RNase P RNA (154 nt) forms several classes of tertiary structure modules: a U-shaped T-loop motif that docks into the P7−P10 helices through an A-minor interaction, numerous stacking interactions (including A130 on A230 in P9), two large internal loops (J11/12 and J12/11) which form extensive noncanonical interactions, and a tetraloop-receptor interaction (involving L12 and P10.1).[6]

10.1021/ja8061216 CCC: $40.75

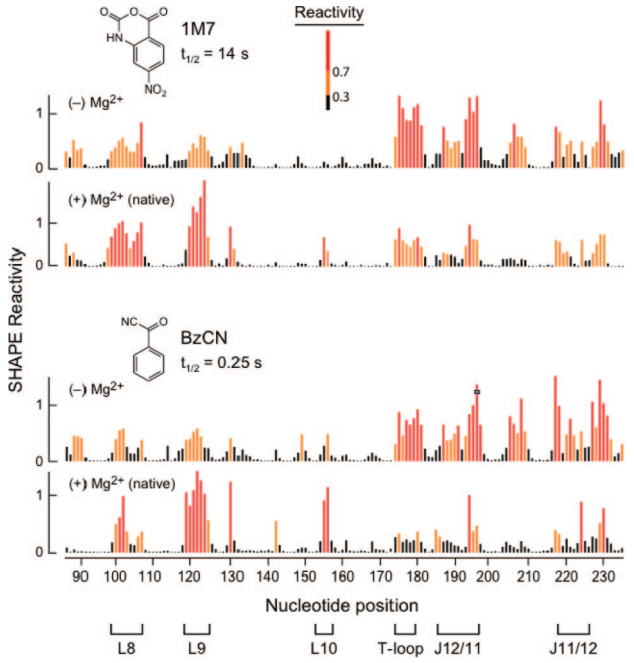

Figure 2. Histograms comparing absolute SHAPE reactivities for the RNase P specificity domain in the presence and absence of Mg^{2+} obtained using 1M7 (top) and BzCN (bottom). Loop (L*x*) and joining structure (J*x*) landmarks are emphasized.

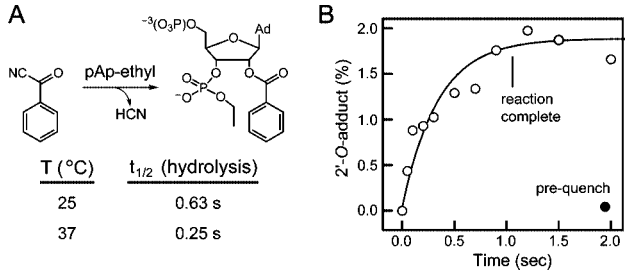

Figure 3. Reaction half-life of BzCN. (A) Reaction of BzCN with pAp-ethyl. (B) Time-dependent reaction of BzCN with pAp-ethyl at 37 °C. Half-life of BzCN in water was determined by fitting the fraction 2′-*O*-adduct formed to $1 - \exp[(0.02)(\exp(-k_{\text{hydrolysis}}t) - 1)]$.

Tertiary folding of the RNase P RNA was initiated by adding $MgCl_2$ (to 10 mM) to a solution of RNA (0.5 μM) preincubated at 37 °C in 100 mM NaCl (at pH 8.0). Aliquots were removed at time intervals between 5 and 1200 s and added directly to BzCN. This experiment provides 1 s snapshots of RNase P structure over 20 min (Figure 4A). Within 5 s after addition of Mg^{2+}, there is a rapid ordering of the RNA secondary structure, observed as a decrease in SHAPE reactivity throughout the RNA. We focused our analysis on the tertiary folding that occurs after this step. After data normalization, significant changes in SHAPE reactivity occur at 16 nucleotides in the RNA. These nucleotides fell into two kinetically distinct categories (Figure 4B). The first category is a fast phase that folds with a rate constant of 0.06 s^{-1}; the second category is a slower phase that folds with a rate constant of 0.004 s^{-1}. The nucleotides that fall into the first, fast, phase also exhibit conformational changes consistent with the slow phase, suggesting that the slow phase involves a whole-scale tightening of the RNA structure (see Supporting Information, Table S1).

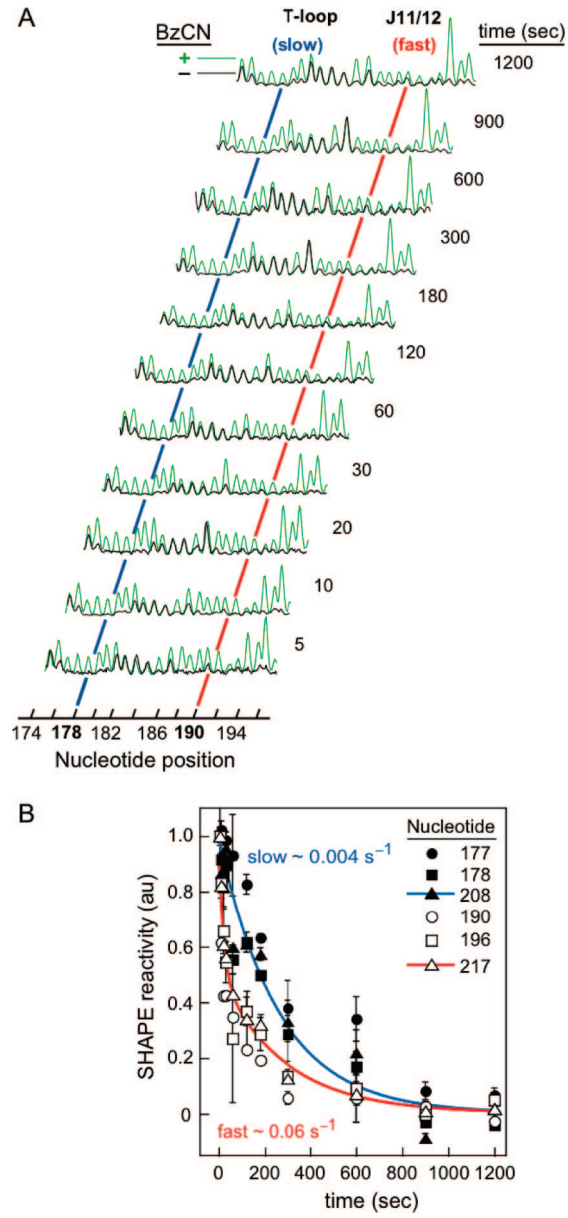

Figure 4. Formation of tertiary interactions in the RNase P specificity domain. (A) Processed capillary electrophoresis traces obtained for the reaction of BzCN with the RNase P specificity domain from 5 s to 20 min. The difference in peak intensity between the (+) and (−) BzCN channels reports the extent of 2′-*O*-adduct formation at each point. Nucleotides involved in tertiary interactions show a decrease in reactivity over time. Nucleotide 178 undergoes this transition slowly, while nucleotide 190 undergoes this transition more rapidly (blue and red lines, respectively). (B) Folding rate constants monitored by time-resolved SHAPE. All nucleotides show a decrease in SHAPE reactivity as a function of time, indicating an increase in structural constraints as the RNA forms tertiary interactions. Nucleotides folding in the fast and slow phases are indicated by open and closed symbols, respectively.

The time-resolved SHAPE data make possible modeling of kinetic intermediates in the folding pathway for the RNase P RNA. Consistent with the two categories of reactivity changes, we observe two distinct folding steps. The first tertiary folding step is characterized by significant protections in the structural module involving the J11/12 and J12/11 loops (residues 185−196 and 217−225) and also in the interaction of A229 with C134 and C232 (in red, Figure 5). The second, slower, folding step involves

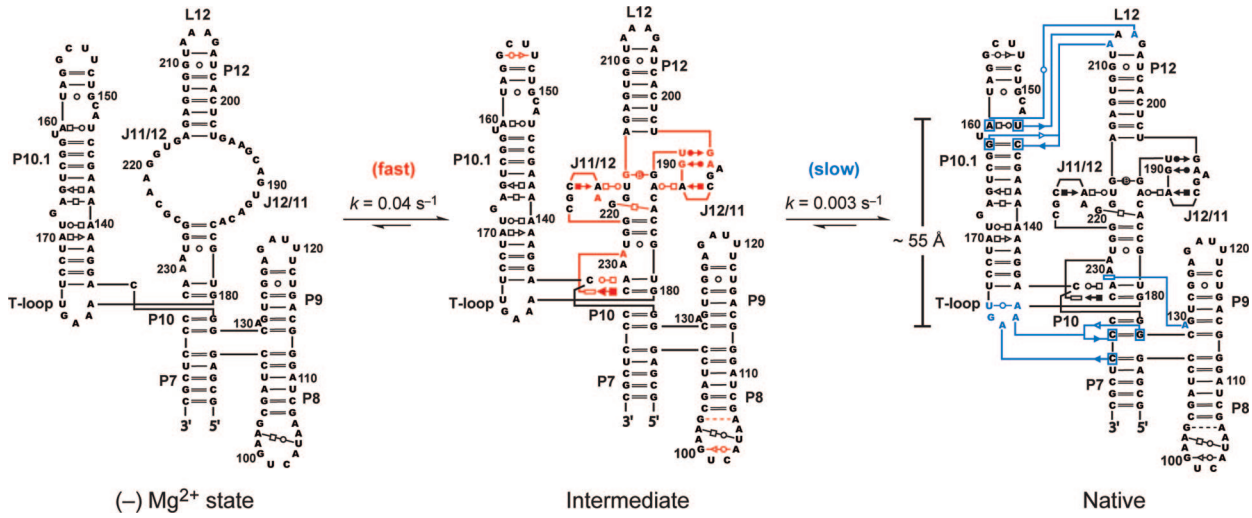

Figure 5. Mechanism for folding of the RNase P specificity domain. Fast-forming tertiary interactions found in the intermediate are red, interactions that form in the rate-determining step are in blue. Structure is drawn to approximate the arrangement of helices in three-dimensional space.[9]

formation of the core of the specificity domain: docking of the T-loop (U175-A179) at P7−P10, stacking of A230 on A130 in P9, and docking of the GAAA tetraloop (P12) into its receptor in P10.1 (in blue, Figure 5).

There are several critical features of this folding mechanism. Although it is becoming clear that RNA folding is not obligatorily hierarchical,[1b,3b,7] in the case of the RNase P RNA, tertiary structure folding does clearly follow the near-complete formation of the secondary structure. Second, tertiary interactions involving local interactions between the J11/12 and J12/11 loops and involving a base stacking interaction at A229 form first (in red, Figure 5). Third, the rate-limiting folding step, despite being on the order of minutes, does not involve the breaking of non-native contacts. Instead, the rate-determining step involves formation of tertiary interactions located 55 Å apart (in blue, Figure 5).

A three-state folding mechanism for the RNase P specificity domain was also developed using equilibrium and kinetic approaches.[8] Time-resolved SHAPE yielded a highly detailed, nucleotide resolution, view of folding in this RNA in a single, very concise, set of experiments (Figure 4A).

Time-resolved SHAPE with BzCN makes it possible to monitor the pathways for formation of individual sets of tertiary interactions in complex RNAs on timescales as short as a few seconds. Because BzCN is self-quenching, yields nucleotide resolution information at every position simultaneously, and affords 1 s snapshots of RNA structure, time-resolved SHAPE should find wide application for developing detailed models for time-dependent processes involving RNA.

Acknowledgment. We are indebted to Dorothy Erie and Scott Kennedy for use of and extensive assistance with quench-flow instrumentation. The NSF supported this work (MCB-0416941 to K.M.W.).

Supporting Information Available: Experimental methods and one table. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (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.
(2) Mortimer, S. A.; Weeks, K. M. *J. Am. Chem. Soc.* **2007**, *129*, 4144–4145.
(3) (a) Badorrek, C. S.; Gherghe, C. M.; Weeks, K. M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13640–13645. (b) Wang, B.; Wilkinson, K. A.; Weeks, K. M. *Biochemistry* **2008**, *47*, 3454–3461. (c) Gherghe, C. M.; Mortimer, S. A.; Krahn, J. M.; Thompson, N. L.; Weeks, K. M. *J. Am. Chem. Soc.* **2008**, *130*, 8884–8885. (d) Gherghe, C. M.; Shajani, Z.; Wilkinson, K. A.; Varani, G.; Weeks, K. M. *J. Am. Chem. Soc.* **2008**, *130*, 12244–12245.
(4) (a) Wilkinson, K. A.; Gorelick, R. J.; Vasa, S. M.; Guex, N.; Rein, A.; Mathews, D. H.; Giddings, M. C.; Weeks, K. M. *PLoS Biol.* **2008**, *6*, e96. (b) Duncan, C. D. S.; Weeks, K. M. *Biochemistry* **2008**, *47*, 8504–8513.
(5) Methods for probing the kinetics of RNA folding include DMS and hydroxyl radical footprinting: (a) Hennelly, S. P.; Antoun, A.; Ehrenberg, M.; Gualerzi, C. O.; Knight, W.; Lodmell, J. S.; Hill, W. E. *J. Mol. Biol.* **2005**, *346*, 1243–1258. (b) Tijerina, P.; Mohr, S.; Russell, R. *Nat. Protoc.* **2007**, *2*, 2608–2623. (c) Scalvi, B.; Sullivan, M.; Chance, M. R.; Brenowitz, M.; Woodson, S. A. *Science* **1998**, *279*, 1940–1943. (d) Shcherbakova, I.; Brenowitz, M. *Nat. Protoc.* **2008**, *3*, 288–302.
(6) Krasilnikov, A. S.; Yang, X.; Pan, T.; Mondragon, A. *Nature* **2003**, *421*, 760–764.
(7) Wu, M.; Tinoco, I. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11555–11560.
(8) (a) Qin, H.; Sosnick, T. R.; Pan, T. *Biochemistry* **2001**, *40*, 11202–11210. (b) Baird, N. J.; Westhof, E.; Qin, H.; Pan, T.; Sosnick, T. R. *J. Mol. Biol.* **2005**, *352*, 712–722. (c) Baird, N. J.; Fang, X.; Srividya, N.; Pan, T.; Sosnick, T. R. *Q. Rev. Biophys.* **2007**, *40*, 113–161.
(9) Lescoute, A.; Westhof, E. *Nucleic Acids Res.* **2006**, *34*, 6587–6604.

JA8061216

Supporting Information for

Time-Resolved RNA SHAPE Chemistry

Stefanie A. Mortimer and Kevin M. Weeks*

*correspondence, weeks@unc.edu

Benzoyl cyanide 2′-*O*-adduct formation and hydrolysis. Adduct formation between [^{32}P]-labeled 3′-phosphoethyl-5′-adenosine (pAp-ethyl)[1] and benzoyl cyanide (BzCN, Sigma-Aldrich) was followed by adding 10% (vol/vol) BzCN (60 mM final, in DMSO; 35.2 μL) to 1.1× reaction buffer [235.4 μL, 11.1 mM $MgCl_2$, 111 mM NaCl, 111 mM Hepes (pH 8.0)] using a chemical quench-flow instrument[2] (KinTek, Model RQF-3; drive syringes were 1.0 and 10 mL for the BzCN reagent and buffer reaction loops, respectively) and quenched with 1 vol 250 mM dithiothreitol [at the following reaction times (in sec): 0.02, 0.07, 0.15, 0.25, 0.40, 0.60, 0.80, 1.0, 1.3, 1.7, 2.5]. Each quenched reaction was precipitated by addition of 10 vol isobutanol, resolved by gel electrophoresis (30% polyacrylamide; 29:1 acrylamide:bisacrylamide; 0.4 mm × 28.5 cm × 23 cm; 30 W, 45 min), and quantified by phosphorimaging. We observed a single adduct band under all conditions, consistent with a mechanism in which BzCN reacts at the 2′-OH group and not significantly with other positions in the ribose or base moieties.

Rate constants for BzCN hydrolysis were obtained using an equation that accounts for parallel reaction of BzCN by 2′-*O*-adduct formation (k_{adduct}) and hydrolysis ($k_{hydrolysis}$):[3]

$$\text{Fraction product} = 1 - exp[(k_{adduct}/k_{hydrolysis})(e^{-k_{hydrolysis}t} - 1)]$$

(see Figure 3 in main text). The term $k_{adduct}/k_{hydrolysis}$ was obtained from the fraction adduct formed at long time points, where (as $t \rightarrow \infty$) $k_{adduct}/k_{hydrolysis} = -\ln(1 - \text{fraction product}) = 0.02$ for all experiments.

Structure-Selective RNA Modification. The RNase P specificity domain RNA[4] was synthesized as described.[1] RNA (60 pmol) in 60 μL sterile water was heated at 95 °C for 2 min, cooled on ice, treated with 36 μL of 3× no-Mg^{2+} folding buffer [333 mM NaCl, 333 mM Hepes (pH 8.0)], and incubated at 37 °C for 5 min. Tertiary structure folding was initiated by adding 12 μL of 10× $MgCl_2$ (100 mM). After mixing, 9 μL of this solution was removed (at 5, 10, 15, 20, 30, 60, 120, 180, 300, 600, 900, and 1200 s) and added directly to 1 μL 10× BzCN (600 mM in DMSO). No-reagent control reactions were added to 1 μL neat DMSO. Modified RNA was recovered by ethanol precipitation [90 μL sterile water, 5 μL NaCl (5 M), 1 μL glycogen (20 mg/mL), 400 μL ethanol; 30 min at –80 °C] and resuspended in 10 μL of TE [10 mM Tris (pH 8.0), 1 mM EDTA].

Primer Extension. The general procedure was that outlined previously.[1,5,6] Briefly, a fluorescently labeled DNA primer (5′ VIC- or NED-labeled GAA CCG GAC CGA AGC CCG; 3 μL, 0.3 μM) was annealed to the RNA (10 μL, from the previous step) by heating at 65 °C (6

min) and 35 °C (5 min). Reverse transcription buffer and Superscript III were added and the reactions incubated at 52 °C for 30 min. Primer extension reactions were quenched by adding an equal volume of a mixture of sodium acetate (1.5 M, pH 5.2) and EDTA (40 mM, pH 8.0) and the resulting cDNAs were recovered by ethanol precipitation, washed twice with 70% ethanol, dried by vacuum for 10 min, and resuspended in 10 μL de-ionized formamide. Dideoxy sequencing markers were generated using unmodified RNA and primers labeled with unique fluorophores (6-FAM or PET, 0.6 μM), and by adding 1 μL of 2′,3′-dideoxycytosine (10 mM) or 2′,3′-dideoxyguanosine (0.25 mM) triphosphate after addition of reverse transcription buffer. cDNA extension products were separated by capillary electrophoresis using an Applied Biosystems 3130 DNA sequencing instrument.

Data Analysis. Raw traces from the ABI 3130 were processed using ShapeFinder.[7] Data sets were normalized by excluding the 2% most reactive nucleotides (3 total) and dividing by the average intensity of the next 8% most reactive nucleotides (12 total). After data normalization, time-dependent changes in reactivity were judged to be significant if they were 0.2 SHAPE units or greater. Sixteen nucleotides exhibited significant reactivity changes by this criterion; the average change in SHAPE reactivity was 0.4. Folding rates for individual nucleotides were obtained by normalizing intensities (I) to the first time point and fitting to either a single

$$I = A + (1 - A)e^{-k_2 t}$$

or double

$$I = A + (1 - A - B)e^{-k_1 t} + Be^{-k_2 t}$$

exponential. Nucleotides 130, 175, 176, 177, 178, 179, 180, 206 and 208 were fit as a single exponential and nucleotides 189, 190, 195, 196, 217, 220, and 229 were fit to the double exponential (Table S1). Secondary structure for the RNase P specificity domain was adapted from Lescoute and Westhof.[8]

References

(1) Mortimer, S. A.; Weeks, K. M. *J. Am. Chem. Soc.* **2007**, *129*, 4144-4145.
(2) Holmes, S. F.; Forster, E. J.; Erie, D. A. *Methods Enzymol.* **2003**, *371*, 71-81.
(3) Merino, E. J.; Wilkinson, K. A.; Coughlan, J. L.; Weeks, K. M. *J. Am. Chem. Soc.* **2005**, *127*, 4223-4231.
(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. *Nature Protocols* **2006**, *1*, 1610-1616.
(6) Wilkinson, K. A.; Gorelick, R. J.; Vasa, S. M.; Guex, N.; Rein, A.; Mathews, D. H.; Giddings, M. C.; Weeks, K. M. *PLoS Biol.* **2008**, *6*, e96.
(7) Vasa, S. M.; Guex, N.; Wilkinson, K. A.; Weeks, K. M.; Giddings, M. C. *RNA* **2008**, *14*, in press.
(8) Lescoute, A.; Westhof, E. *Nucleic Acids Res.* **2006**, *34*, 6587-6604.

Nucleotide	fast k_1 (s^{-1})	slow k_2 (s^{-1})	Folding domain
A130	—	0.006 ± 0.002	
U175	—	0.006 ± 0.001	T-loop
G176	—	0.004 ± 0.002	T-loop
A177	—	0.002 ± 0.0005	T-loop
A178	—	0.004 ± 0.0008	T-loop
A179	—	0.004 ± 0.001	T-loop
G180	—	0.0030 ± 0.001	T-loop
U189	0.06 ± 0.03	0.002 ± 0.003	J11/12
G190	0.07 ± 0.03	0.004 ± 0.004	J11/12
A195	0.05 ± 0.04	0.002 ± 0.006	J11/12
G196	0.04 ± 0.01	0.003 ± 0.003	J11/12
A206	—	0.002 ± 0.001	GAAA Tetraloop-Receptor
A208	—	0.004 ± 0.0009	GAAA Tetraloop-Receptor
G217	0.05 ± 0.02	0.003 ± 0.002	J12/11
A221	0.08 ± 0.07	0.005 ± 0.005	J12/11
A229	0.02 ± 0.04	0.001 ± 0.001	

Table S1. Folding rate constants for nucleotides in the RNase P specificity domain involved in forming tertiary interactions.