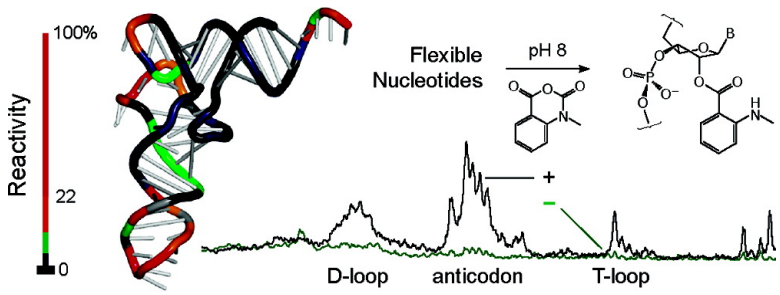
Article

RNA Structure Analysis at Single Nucleotide Resolution by Selective 2'-Hydroxyl Acylation and Primer Extension (SHAPE)

Edward J. Merino, Kevin A. Wilkinson, Jennifer L. Coughlan, and Kevin M. Weeks

J. Am. Chem. Soc., **2005**, 127 (12), 4223-4231• DOI: 10.1021/ja043822v • Publication Date (Web): 04 March 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 12 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 03/04/2005

RNA Structure Analysis at Single Nucleotide Resolution by Selective 2′-Hydroxyl Acylation and Primer Extension (SHAPE)

Edward J. Merino, Kevin A. Wilkinson, Jennifer L. Coughlan, and Kevin M. Weeks*

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

Received October 11, 2004; E-mail: weeks@unc.edu

Abstract: The reactivity of an RNA ribose hydroxyl is shown to be exquisitely sensitive to local nucleotide flexibility because a conformationally constrained adjacent 3′-phosphodiester inhibits formation of the deprotonated, nucleophilic oxyanion form of the 2′-hydroxyl group. Reaction with an appropriate electrophile, *N*-methylisatoic anhydride, to form a 2′-O-adduct thus can be used to monitor local structure at every nucleotide in an RNA. We develop a quantitative approach involving Selective 2′-Hydroxyl Acylation analyzed by Primer Extension (SHAPE) to map the structure of and to distinguish fine differences in structure for $tRNA^{Asp}$ transcripts at single nucleotide resolution. Modest extensions of the SHAPE approach will allow RNA structure to be monitored comprehensively and at single nucleotide resolution for RNAs of arbitrary sequence and structural complexity and under diverse solution environments.

Introduction

Most RNA molecules function correctly in biology only once they fold back on themselves to form complex and difficult-to-predict three-dimensional structures.[1] In addition, the biology of many RNAs requires that specific regions have local structural flexibility or the ability to transition from one conformation to another. Understanding the function of almost all RNAs thus requires an accurate and, ideally, comprehensive knowledge of the base-paired secondary structure and identification of nucleotides involved in tertiary interactions.

Traditional chemical and enzymatic reagents[2] and backbone-based cleavage[3] are useful for mapping RNA secondary structure, especially when multiple reagents are used in concert. However, traditional approaches for monitoring base pairing and local nucleotide flexibility suffer from two broad shortcomings. (1) Reagents are leaky in the sense that the absolute magnitude of discrimination between single-stranded and base-paired regions can be small and also because helix termini and some base pairs (especially G−U pairs) are often reactive.[4] (2) The information obtained by existing approaches is also sparse because a given reagent typically reacts with only a subset of the four RNA nucleotides[2] or backbone sites.[3] Because the chemistry of each sparsely reactive structure-selective reagent is idiosyncratic, multiple reagents must be used to interrogate all positions in an RNA, and it is difficult to ascribe a single, unifying set of rules for the classes of structures that react versus those that do not.

An alternate approach for mapping RNA structure exploits the observation that the chemical reactivity of the 2′-ribose position is strongly modulated by the adjacent 3′-phosphodiester anion.[5,6] Acylation of synthetic 2′-amine-substituted nucleotides to form the 2′-amide product is strongly gated by the underlying local nucleotide flexibility.[5,7,8] Flexible nucleotides in RNA are better able to reach a facilitated transition state in which the 3′-phosphodiester becomes appropriately positioned with the 2′-amine to exert a catalytic effect.[6] 2′-Amine acylation thus robustly detects essentially all base-paired RNA secondary structures and many tertiary interactions when the sites of 2′-adduct formation are mapped by primer extension.[5,8]

We have found that this 2′-ribose-based chemistry is significantly more straightforward to implement than traditional chemical or enzymatic approaches for monitoring RNA secondary and tertiary structure. However, introduction of an artificial 2′-amine group makes an experiment more complex, prevents this chemistry from being used in vivo, and may perturb some tertiary interactions involving the 2′-ribose position. We considered that the proximity of a 3′-phosphodiester anion might modulate the reactivity of the 2′-hydroxyl group normally present in RNA. Since every nucleotide has a 2′-hydroxyl, structural information is, in principle, obtainable for every

(1) (a) Brion, P.; Westhof, E. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 113−137. (b) Gesteland, R. F.; Cech, T. R.; Atkins, J. F. *The RNA World*, 2 ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1999. (c) Tinoco, I.; Bustamante, C. *J. Mol. Biol.* **1999**, *293*, 271−281.
(2) Ehresmann, C.; Baudin, F.; Mougel, M.; Romby, P.; Ebel, J.-P.; Ehresmann, B. *Nucleic Acids Res.* **1987**, *15*, 9109−9128.
(3) (a) Soukup, G. A.; Breaker, R. R. *RNA* **1999**, *5*, 1308−1325. (b) Helm, M.; Brule, H.; Degoul, F.; Cepanec, C.; Leroux, J. P.; Giege, R.; Florentz, C. *Nucleic Acids Res.* **1998**, *26*, 1636−1643. (c) Wittberger, D.; Berens, C.; Hammann, C.; Westhof, E.; Schroeder, R. *J. Mol. Biol.* **2000**, *300*, 339−352.
(4) Mathews, D. H.; Disney, M. D.; Childs, J. L.; Schroeder, S. J.; Zuker, M.; Turner, D. H. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7287−7292.
(5) Chamberlin, S. I.; Weeks, K. M. *J. Am. Chem. Soc.* **2000**, *122*, 216−224.
(6) Chamberlin, S. I.; Merino, E. J.; Weeks, K. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14688−14693.
(7) (a) John, D. M.; Weeks, K. M. *Chem. Biol.* **2000**, *7*, 405−410. (b) John, D. M.; Weeks, K. M. *Biochemistry* **2002**, *41*, 6866−6874. (c) John, D. M.; Merino, E. J.; Weeks, K. M. *J. Mol. Biol.* **2004**, 611−619.
(8) Chamberlin, S. I.; Weeks, K. M. *Biochemistry* **2003**, *42*, 901−909.

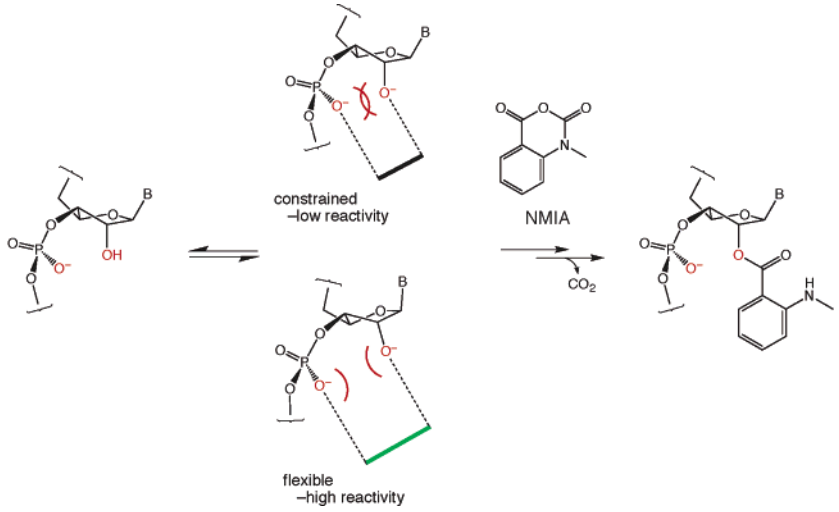


Figure 1. Reaction of a ribose 2′-hydroxyl with NMIA via formation of an oxyanion nucleophile. Nucleotides constrained in the RNA C3′-endo and other conformations are unreactive due to destabilization (red overlap symbol) of the oxyanion by the proximal 3′-phosphodiester.

position in an RNA. Formation of a bulky 2′-O-adduct could then be detected as an impediment to cDNA synthesis by a reverse transcriptase enzyme.[5,8]

In this work, we show that selective acylation of the ribose 2′-hydroxyl position by *N*-methylisatoic anhydride (NMIA) is exquisitely sensitive to local nucleotide flexibility. We then use selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) to analyze comprehensively local structure in yeast $tRNA^{Asp}$ transcripts at nucleotide resolution in a single series of experiments.

Results

A 2′-Hydroxyl Reactive Reagent. Since prior work showed that the nucleophilicity of a 2′-ribose substituent is modulated by the proximal 3′-group,[6] we initially monitored reactivity of ATP and 3′-deoxy-ATP toward diverse reagents[9,10] potentially able to react with hydroxyl groups under mild conditions in aqueous solution. The most promising reagent was *N*-methylisatoic anhydride (NMIA), which liberates CO_2 to form a 2-methylaminobenzoic acid nucleotide 2′-ester[10] (Figure 1). The 2′-O-adduct is readily resolved as a low mobility band by gel electrophoresis (Figure 2A). ATP reacts 2.5-fold more rapidly than 3′-deoxy-ATP to form the monoadduct (Figure 2B) and, once modified, reacts slowly to form a second adduct with NMIA.

An electrophile capable of reacting with a hydroxyl group in aqueous solution will face competition from the analogous hydrolysis reaction (Figure 3A). To measure the rate of NMIA hydrolysis, we monitored formation of the 2-methylaminobenzoate product by fluorescence. 2-Methylaminobenzoate was excited selectively at 375 nm and emission detected at 440 nm (Figure 3B). NMIA hydrolysis shows good first-order kinetics with a rate constant of 0.031 min^{-1} (at 25 °C, Figure 3C). Given this rate constant for reagent hydrolysis, we are able to calculate rates for ribose 2′-O-adduct formation explicitly.

NMIA Reactivity with Nucleotides is Modulated by the 3′-Substituent. Attack of a 2′-hydroxyl at the anhydride moiety

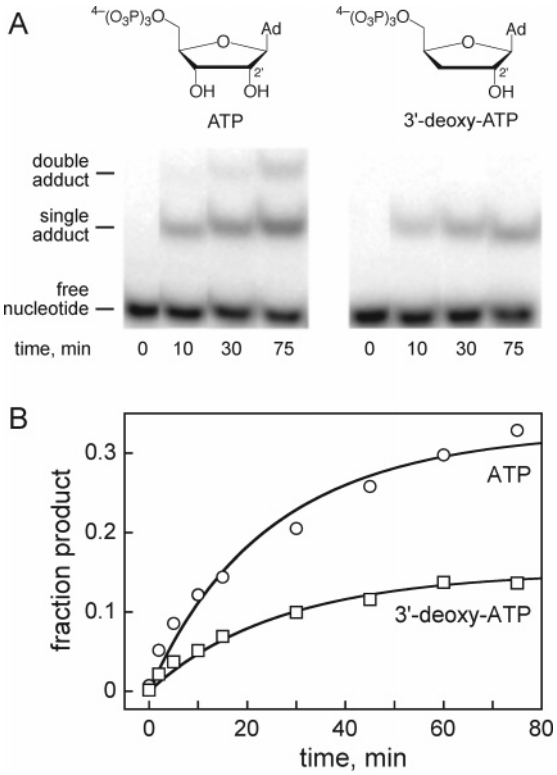


Figure 2. Structure-selective reaction of NMIA with ribose hydroxyl groups. (A) Visualizing nucleotide 2′-ester formation with ATP and 3′-deoxy-ATP. NMIA-modified nucleotides migrate with reduced mobility by polyacrylamide gel electrophoresis. (B) Kinetic analysis of nucleotide adduct formation; lines represent least-squares fits to eq 1.

of NMIA requires formation of the significantly more nucleophilic 2′-oxyanion prior to reaching the ester-forming transition state (see first step in Figure 1). Formation of the 2′-oxyanion would be destabilized if the negatively charged 3′-phosphodiester were constrained to be adjacent to the 2′-position (see red overlap symbols in Figure 1). Because formation of the oxyanion involves loss of a proton, the hydroxyl pK_a is a good indicator for the relative nucleophilicity of a 2′-hydroxyl. The pK_a for deprotonation of the 2′-hydroxyl of adenosine is 12.2, and this pK_a increases to 13.4 for adenosine-3′-phosphate.[11] The relative nucleophilicity of a 2′-hydroxyl is thus strongly modulated by

(9) (a) de Belder, A. N.; Wik, K. O. *Carbohydr. Res.* **1975**, *44*, 251−257. (b) Stearns, N. A.; Prigent-Richard, S.; Letourneur, D.; Castellot, J. J. *Anal. Biochem.* **1997**, *247*, 348−356. (c) Cardinaud, R.; Baker, B. R. *J. Med. Chem.* **1970**, *13*, 467−470.

(10) (a) Moorman, A. R.; Abeles, R. H. *J. Am. Chem. Soc.* **1982**, *104*, 6785−6786. (b) Hiratsuka, T. *Biochim. Biophys. Acta* **1983**, *742*, 496−508.

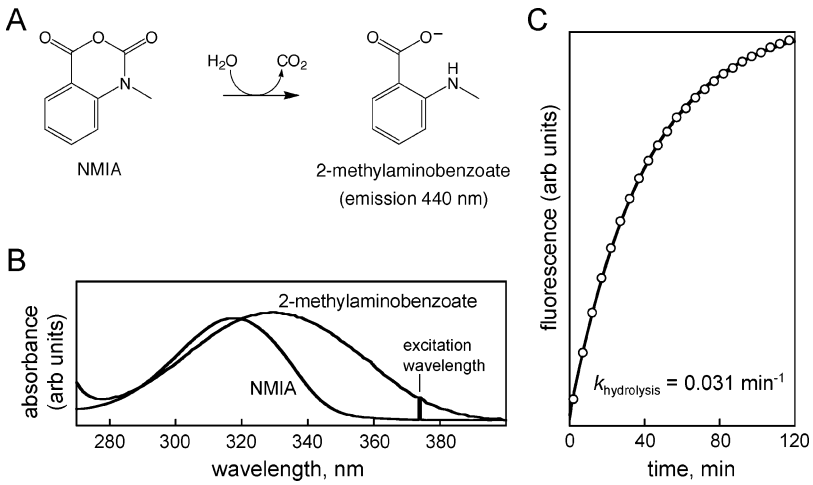


Figure 3. NMIA hydrolysis. (A) Reaction of NMIA with water to form the fluorescent 2-methylaminobenzoate. (B) Formation of the 2-methylaminobenzoate product is accompanied by a shift in absorbance maximum from 320 to 335 nm; product fluorescence was thus detected by exciting at 375 nm (heavy line) to eliminate inner filter effects. (C) Pseudo-first-order hydrolysis of NMIA (25 °C); every tenth data point is shown.

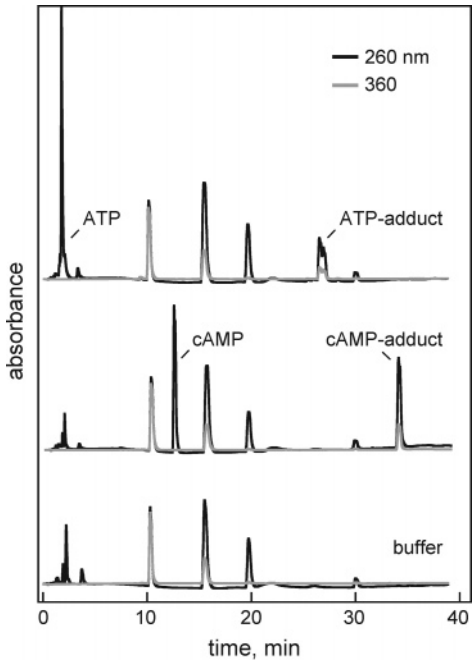


Figure 4. End point analysis to determine rate of reaction of ATP and cAMP with NMIA. Reaction endpoints were quantified by HPLC. Fraction ATP or cAMP product was obtained by integration of the indicated peaks. Assignment of nucleotide starting material, products, and buffer components is straightforward because all components absorb at 260 nm, whereas only 2-methylaminobenzoate products are detected at 360 nm.

neighboring functional groups at the 3′-position. Formation of the oxyanion also presumably induces secondary changes in ribose conformation to yield complex relationships between nucleotide structure and the 2′-hydroxyl pK_a. We, therefore, explored reactivity for an instructive set of nucleotides (Figure 5) whose functional group or conformation at the 3′-position had the potential to inhibit or facilitate 2′-ester formation with NMIA (see Figure 1).

Two methods were used to obtain pseudo-first-order rate constants for adduct formation (k_{adduct}) for reaction of the nucleotide analogues with NMIA. Most analogues could be radiolabeled and were resolved from their reaction products by electrophoresis (Figure 2A). Rates were then obtained by fitting complete reaction profiles (Figure 2B). For the remaining analogues, product formation was determined as the fraction 2′-O-adduct formation after complete NMIA hydrolysis, using HPLC (Figure 4; see Experimental Section). Rates were obtained by both methods for ATP and 2′-deoxy-ATP and agreed within error.

All analogues contained either a 5′-mono- or triphosphate to block reaction at the 5′-hydroxyl. The k_{adduct} values for ATP and AMP are similar at 0.015 and 0.022 min^{-1}, respectively, indicating that reaction of the 2′-hydroxyl with NMIA is not significantly modulated by the number of phosphate moieties at the 5′-position (Figure 5E). When 2′,3′-dideoxy-ATP (Figure 5F) was treated with NMIA, no new species were detected by HPLC, confirming that adduct formation requires a free ribose hydroxyl group.

We evaluated the reactivity of two reference analogues, pAp-ddC and pAp-ethyl (Figure 5A), designed to mimic the 2′-hydroxyl reactivity of unstructured regions in larger RNAs. Both analogues possess an authentic 3′-phosphodiester group and a single reactive 2′-hydroxyl. The pAp-ddC and pAp-ethyl react with k_{adduct} values of 0.0010 and 0.0018 min^{-1}, respectively. The slightly faster reactivity of the pAp-ethyl analogue shows that simple steric effects make, at most, a small contribution to 2′-hydroxyl reactivity at flexible nucleotides.

We then evaluated a series of nucleotide analogues designed to assess the impact of a 3′-phosphodiester on the reactivity of the adjacent 2′-hydroxyl (Figure 5B). When the 3′-phosphodiester in pAp-ddC or pAp-ethyl is replaced with either a 3′-methoxy (3′-O-methyl-ATP) or a 3′-hydrogen (3′-deoxy-ATP), k_{adduct} increases by 3- or 6-fold, respectively. Analogously, constraining the 3′-phosphodiester away from the 2′-hydroxyl in a 3′,5′-cyclic linkage (cAMP, Figure 5C) yields an analogue that reacts at a rate of 0.018 min^{-1}, 10−18-fold faster than that of the RNA-like analogues, pAp-ddC and pAp-ethyl.

Finally, the inhibitory effect of a proximal charged phosphate moiety is obvious from comparing reactivities for 3′- versus 2′-deoxy-ATP (Figure 5B,D). The 5′-triphosphate is closer in distance to the 3′-hydroxyl than the 2′-hydroxyl, and correspondingly, 2′-deoxy-ATP reacts 12-fold slower than 3′-deoxy-ATP (k_{adduct} is 0.0005 versus 0.0060 min^{-1}). We infer

(11) Velikyan, I.; Acharya, S.; Trifonova, A.; Földesi, A.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **2001**, *123*, 2893−2894.

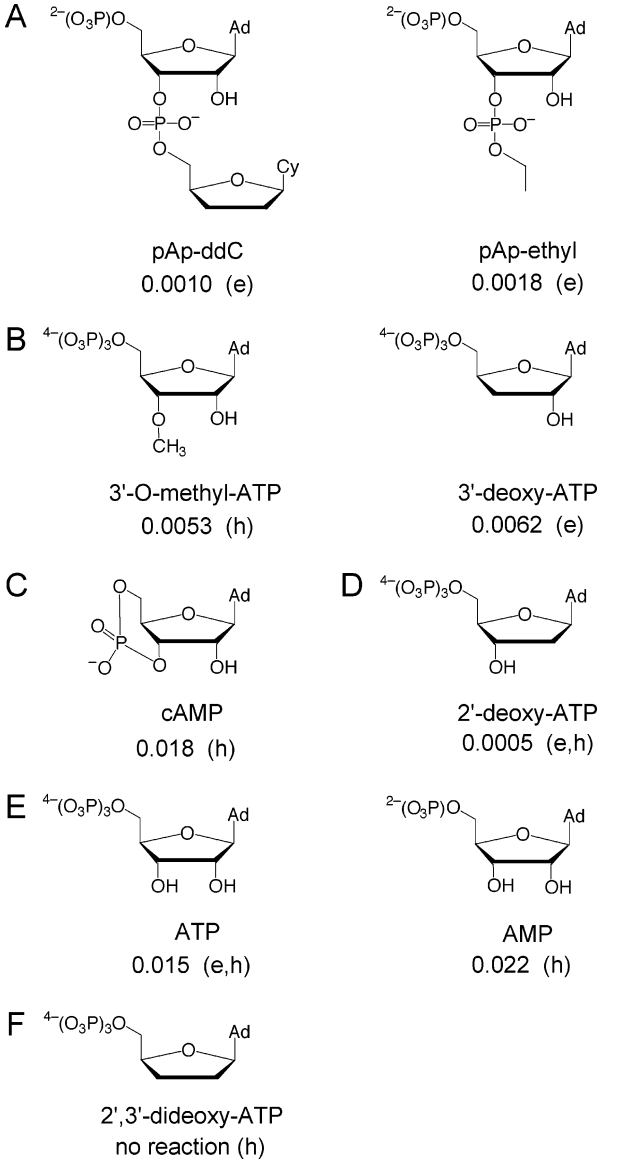


Figure 5. Structure−reactivity analysis for reaction of nucleotide analogues with NMIA. For each nucleotide or analogues series (A−F), k_{adduct} is reported in min^{-1} at 13 mM NMIA and 25 °C. Rates were obtained from complete profiles visualized by gel electrophoresis (e) or by single time point analysis using HPLC (h); a subset of nucleotides was evaluated by both methods.

that the 5′-triphosphate more strongly destabilizes oxyanion formation at the 3′-position than at the more distant 2′-hydroxyl (see 2′-deoxy-ATP in Figure 5D).

In sum, these data support the model (Figure 1) that the reactivity of a ribose hydroxyl group is decreased by the proximity of *any* phosphate group. Reactivity can be recovered either by replacing a phosphate diester with an uncharged analogue or by increasing the distance between the hydroxyl and phosphate groups.

Structure-Selective 2′-Hydroxyl Reactivity in an Oligonucleotide. We extended these studies on the relationship between constrained local structure and 2′-hydroxyl reactivity by evaluating reactivity in an 11-mer oligonucleotide (Figure 6). This RNA was synthesized with a 5′-monophosphate radiolabel and a 3′-dideoxycytidine nucleotide so that all 10 remaining hydroxyl groups lie in internal linkages and adjacent to 3′-phosphodiester groups. When the 5′-labeled oligonucleotide

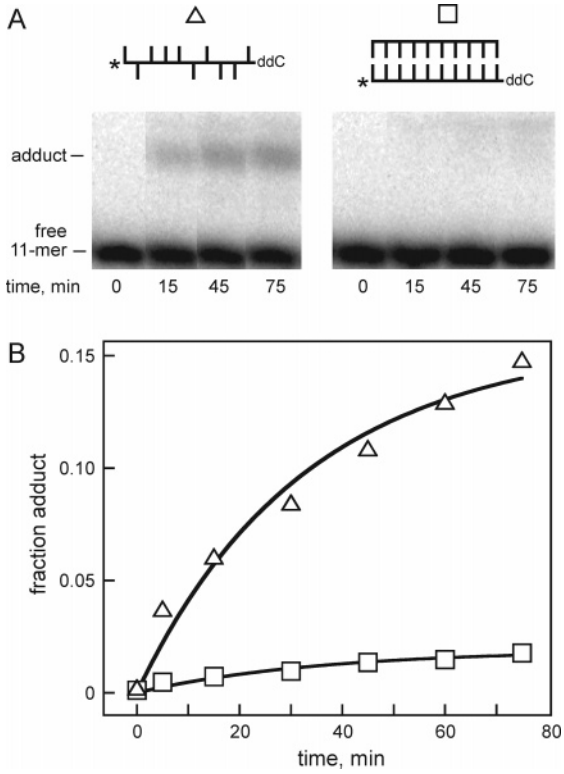


Figure 6. 2′-Hydroxyl acylation by NMIA is selective for unpaired nucleotides. (A) 2′-Adduct formation with the oligonucleotide (5′-UAC-UAACACC-ddC) and the corresponding duplex visualized by gel electrophoresis. Site of [^{32}P]-label is indicated with an asterisk. (B) Quantitative analysis of 2′-O-adduct formation averaged over the 10 hydroxyl groups in the labeled oligonucleotide. Rates for the single strand (triangles) and duplex (squares) were obtained by fitting to eq 1.

was treated with NMIA, a slowly migrating adduct formed, whose overall synthesis occurs at a rate of 0.0048 min^{-1} (left panel in Figure 6A). Since this oligonucleotide contains 10 hydroxyl groups, each reacts, on average, at a rate of (0.0048/10) or ~0.0005 min^{-1} (triangles in Figure 6B), or within 2-fold of the rate for the pAp-ddC reference dinucleotide (0.0010 min^{-1}, Figure 5A). The slightly slower average rate may reflect stronger stacking and other constraining interactions in the longer oligonucleotide.

When the 11-mer oligonucleotide was hybridized to its complement, the rate of adduct formation decreased by 10-fold to 0.0005 min^{-1} (compare triangles and squares in Figure 6B). Residues at RNA helix termini experience fraying[12] and are more flexible than internal nucleotides. Thus, the calculated average rate at a single position (0.0005/10 or 5×10^{-5} min^{-1}) must represent an upper limit for reaction of hydroxyl groups constrained by base pairing.

The differential reactivity for single-stranded versus helical oligonucleotides (Figure 6) supports two important conclusions. First, the average reactivity at an individual site in the 11-mer oligonucleotide is within a factor of 2 for that of the pAp-ddC reference dinucleotide. The similar rate supports the interpretation that the analogue studies reproduce key mechanistic features of 2′-O-adduct-forming reactions at flexible sites in large RNAs. Second, because ribose 2′-hydroxyl groups project outward from an RNA duplex, the differential reactivity of single-stranded

(12) (a) Preisler, R. S.; Mandal, C.; Englander, S. W.; Kallenbach, N. R.; Frazier, J.; Miles, H. T.; Howard, F. B. *Biopolymers* **1984**, *23*, 2099−2125. (b) Snoussi, K.; Leroy, J.-L. *Biochemistry* **2001**, *40*, 8898−8904.

versus duplex RNA (Figure 6B) rules out solvent accessibility as the primary explanation for structure-sensitive 2′-hydroxyl reactivity.

A Structure Cassette for Analyzing tRNA^{Asp} Conformation. We apply selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) to explore the structure of *Saccharomyces cerevisiae* tRNA^{Asp} transcripts. The well-studied L-shaped tRNA^{Asp} structure[5,13,14] (Figure 7A) encompasses interactions that comprise any folded RNA, including double-stranded helices, single-stranded loops and connecting regions, and a network of tertiary interactions. The secondary structure shown (Figure 7A) emphasizes the three-dimensional fold of the RNA in which (i) the anticodon and D-stems and (ii) the acceptor and T-stems stack. These two pairs of coaxially stacked helices are then held together in the tRNA architecture by tertiary interactions between the D- and T-loops and involving the joining regions at positions U8−A9 and G45−U48.

The tRNA^{Asp} transcript was placed in the center of an RNA "structure cassette" to facilitate analysis of 2′-O-adduct formation by primer extension (Figure 7B). Each stem-loop structure in the cassette contains a stable UUCG tetraloop[15] to enforce the designed fold and eliminate alternate foldings with tRNA^{Asp}. The hairpins that comprise the flanking 5′ and 3′ cassette structures have three functions. (1) The 3′-most sequence (green in Figure 7B) provides the DNA primer binding site required to initiate reverse transcription. (2) The 3′-linker (blue) between the primer binding site and the first position in the tRNA allows the reverse transcriptase enzyme to become fully processive prior to reaching the region of structural interest. This linker also prevents nontemplated primer extension products from masking structural information at the 3′-end of the RNA. (3) The 5′-portion of the cassette (red in Figure 7B) displaces the abundant full length extension product that would otherwise obscure structural information at the 5′-end of the RNA. Overall, the design of the structure cassette balances the requirements that it fold autonomously, not interact with the internal RNA, and, yet, efficiently bind a primer DNA.

SHAPE Footprinting of tRNA^{Asp}. The tRNA^{Asp} construct was folded under strongly native conditions (100 mM NaCl, 10 mM $MgCl_2$)[5,13] at 35 °C. NMIA was added and allowed to react with the RNA for five hydrolysis half-lives (54 min). Sites of 2′-esterification (Figure 1) were identified as stops to primer extension by reverse transcriptase. The presence of a 2′-O-adduct causes the reverse transcriptase to stop exactly one nucleotide prior to the modified base.

Primer extension products were resolved in a sequencing gel and converted to intensity versus position plots (Figure 8A). When tRNA^{Asp} was incubated under the standard reaction conditions in the absence of NMIA, primer extension yields a strong full-length product and only minor bands of intermediate length (green trace, Figure 8A). When an otherwise identical SHAPE experiment was performed in the presence of NMIA, a strong and reproducible pattern of 2′-O-adducts is detected by primer extension (compare black (+NMIA reaction) with green (−NMIA) traces in Figure 8A). The fraction full-length extension product is 70% of that for the no reaction control, indicating that NMIA reacts with single-hit kinetics.[16] Sites of adduct formation were identified by comparison with a dideoxy sequencing lane (blue trace, Figure 8A).

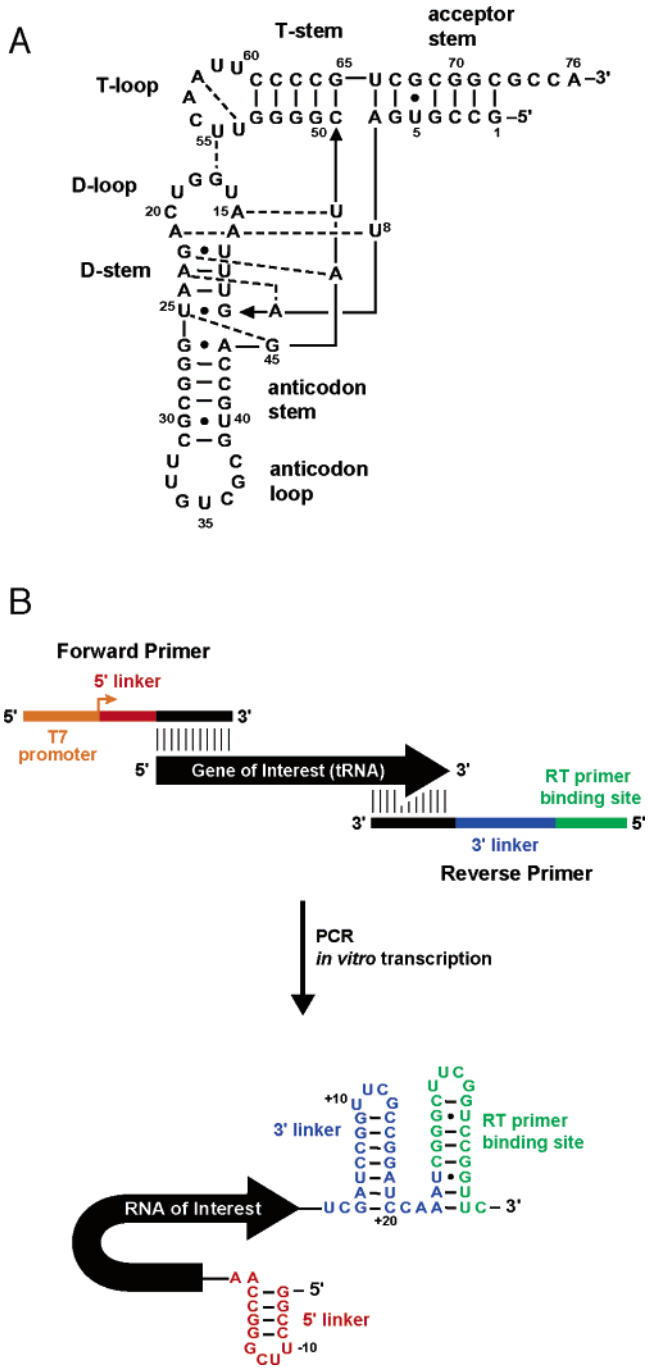


Figure 7. The tRNA^{Asp} transcript and structure cassette. (A) Secondary structure of tRNA^{Asp}. Structure is drawn to suggest arrangement of RNA helices in three-dimensional space. Tertiary interactions are emphasized with dashed lines. The RNA transcript does not contain post-transcriptional modifications, and the native tRNA^{Asp} U1−A72 base pair was changed to a G−C pair to permit comparison with prior work.[5] By convention,[14] nucleotides in yeast tRNA^{Asp} are numbered 1−76, with number 47 omitted. (B) The RNA structure cassette can be appended to any RNA of interest by transcription from a PCR-generated template.

(13) (a) Perret, V.; Garcia, A.; Puglisi, J.; Grosjean, H.; Ebel, J. P.; Florentz, C.; Giege, R. *Biochimie* **1990**, *72*, 735−743. (b) Romby, P.; Moras, D.; Dumas, P.; Ebel, J. P.; Giege, R. *J. Mol. Biol.* **1987**, *195*, 193−204.

(14) (a) Westhof, E.; Dumas, P.; Moras, D. *J. Mol. Biol.* **1985**, *184*, 119−145. (b) Westhof, E.; Dumas, P.; Moras, D. *Acta Crystallogr. A* **1988**, *44*, 112−123.

(15) (a) Tuerk, C.; Gauss, P.; Thermes, C.; Groebe, D. R.; Gayle, M.; Guild, N.; Stormo, G.; d'Aubenton-Carafa, Y.; Uhlenbeck, O. C.; Tinoco, I.; Brody, E. N.; Gold. L. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1364−1368. (b) Cheong, C.; Varani, G.; Tinoco, I. *Nature* **1990**, *346*, 680−682.

(16) Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Methods Enzymol.* **1986**, *130*, 132−181.

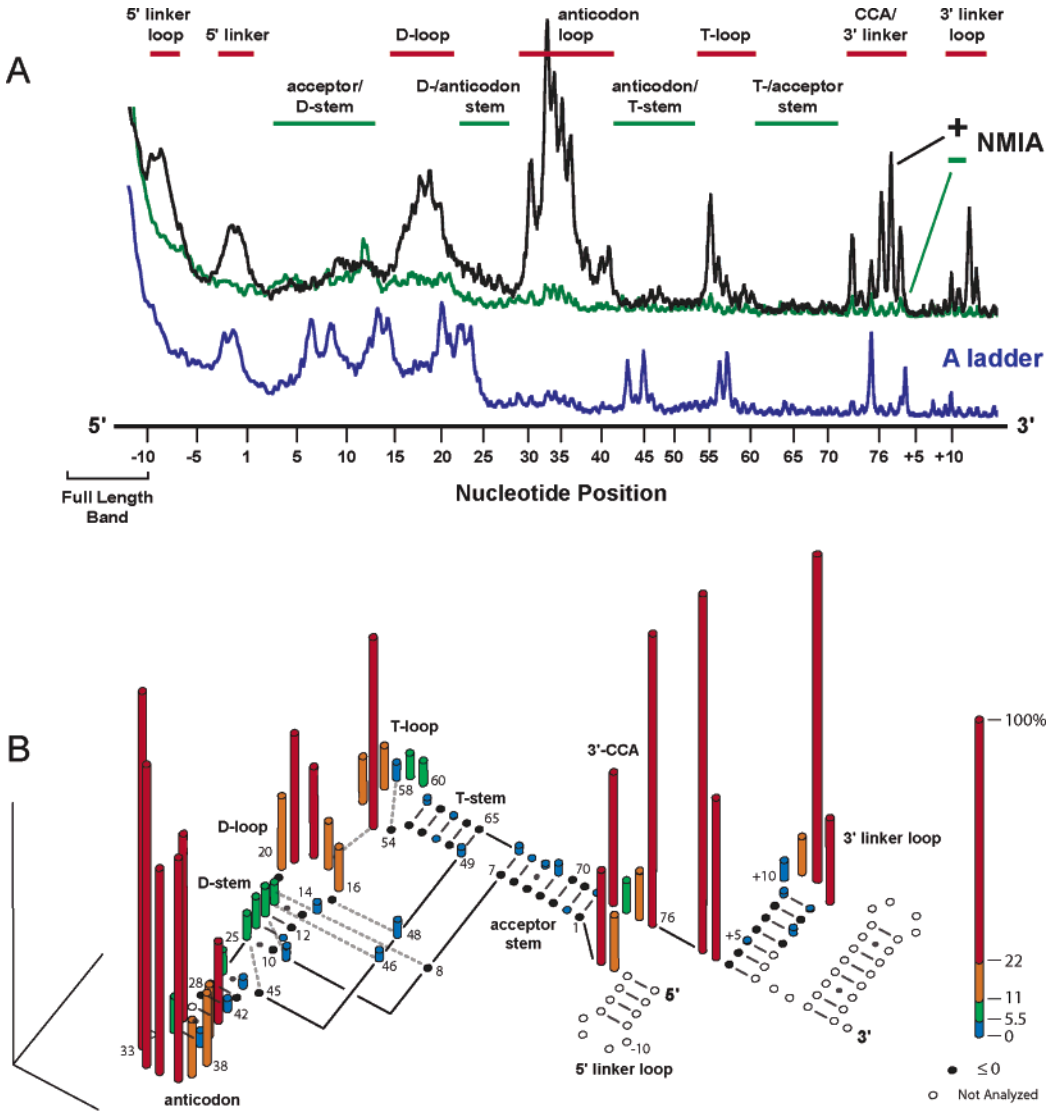


Figure 8. SHAPE analysis of tRNA^{Asp}. (A) Profile for 2′-O-adduct formation as a function of nucleotide position, visualized by primer extension. Reactions performed in the presence and absence of NMIA are shown in black and green, respectively. Regions that react selectively with or are excluded from reaction with NMIA are emphasized with red or green bars, respectively. A dideoxy sequencing marker is shown in blue and is one nucleotide longer than corresponding positions in the (+) and (−) NMIA lanes. (B) Absolute 2′-hydroxyl reactivities superimposed on a secondary structure for the tRNA^{Asp} construct. Column heights are proportional to the band intensities shown in part A (black trace) minus background.

Regions that showed the strongest reactivity relative to background are emphasized with red bars in Figure 8A, and all reactivities are superimposed on a secondary structure for the RNA transcript in Figure 8B. Comparing the most reactive nucleotide, U33 in the anticodon loop, with unreactive residues in, for example, the T- and acceptor stems indicates that the dynamic range of the SHAPE experiment spans 30−50-fold. 2′-Hydroxyl reactivity observed in the 5′- and 3′-linkers is exactly that expected for proper folding of the cassette, indicating the structure cassette folds as designed (Figure 8B).

When superimposed on the secondary structure for tRNA^{Asp} and the flanking structure cassette, absolute 2′-hydroxyl reactivities correspond closely to whether a given nucleotide is constrained. Nucleotides that do not participate in either base pairing or tertiary interaction are most reactive (red and orange columns in Figure 8B). Conversely, NMIA modification is significantly reduced in each of the structures that comprise the four canonical tRNA helices (green, blue, and black columns in Figure 8B). Tertiary interactions, including interaction of G45 with the G10−U25 pair, the A15−U48 Levitt pair, the A14−A21−U8 base triple, and the U54−A58 intraloop interaction, are unreactive.

2′-Ribose-based NMIA chemistry is clearly generic to all RNA nucleotides. Each nucleotide (A, G, C, and U) shows strong reactivity in at least one part of the tRNA^{Asp} transcript, and all four residues show protection at other positions when constrained by base pairing or a tertiary interaction (Figure 8B, see Figure 7A for nucleotide assignments).

Discussion

2′-Hydroxyl Acylation with NMIA Scores Local Nucleotide Flexibility. RNA SHAPE footprinting reproduces the canonical structure of tRNA^{Asp}. Reactive nucleotides lie almost exclusively in the conserved T-, D-, and anticodon loops and at the 3′-CCA terminus (red and orange columns in Figure 8B). Of the five noncanonical G−U or G−A pairs in tRNA^{Asp}, only

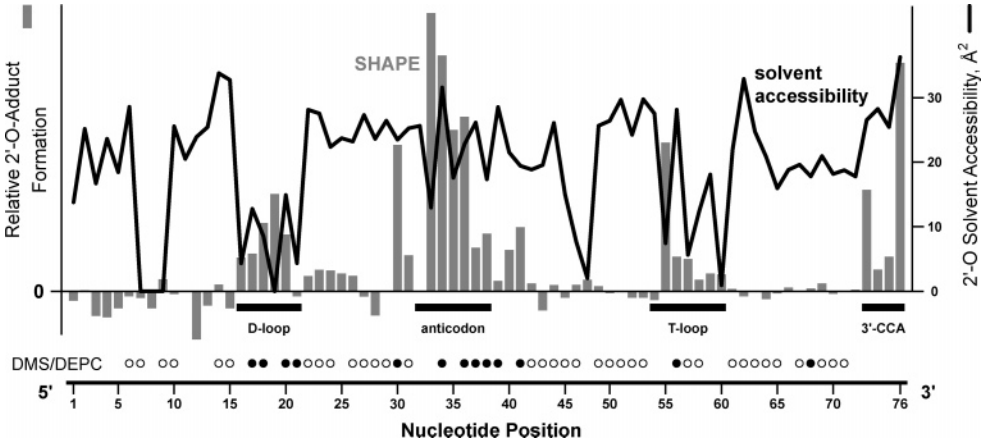


Figure 9. RNA SHAPE reactivity is not correlated with solvent accessibility. Relative 2′-O-adduct formation (gray bars) was calculated from the data shown in Figure 8A. Solvent accessibilities (heavy lines) were computed[19] assuming a probe radius of 1.4 Å. For comparison, results from chemical mapping studies on $tRNA^{Asp}$ transcripts using DMS and DEPC[13a] are shown with spheres; reactive and unreactive positions are emphasized with filled and open spheres, respectively.

one, in the anticodon stem, is reactive. SHAPE footprinting thus appears to be more sensitive to constraints imposed by noncanonical pairing than traditional chemistries because G−U and G−A pairs are typically reactive toward DMS and CMCT.[4,13]

NMIA reactivity also detects tertiary interactions. Essentially, all nucleotides that form higher-order interactions (dashed lines in Figure 8B) are unreactive at 35 °C. This is a striking result given that the unreactive structures are constrained by diverse local tertiary interactions, including Hoogsteen paring, base triples, and cross-loop interactions.

Finally, the SHAPE approach is also uniquely information-rich in the sense that the absolute reactivity amplitudes appear to be meaningful. First, nucleotides in the T- and D-loops, while generally reactive, are less so than those in the anticodon loops, consistent with formation of significant tertiary interactions in the former loops. Within the T-loop, the cross-loop U54−A58 interaction is reported by low reactivities in an otherwise reactive region. Second, nucleotides in the C−G-rich T- and acceptor stems show reactivities that are measurably lower than those of many nucleotides in the D- and anticodon stems (Figure 8B). Third, inspection of the local patterns of reactivity within each of the loops in $tRNA^{Asp}$, at the 3′-CCA region, and in the 3′-cassette (the +10 − +14 loop) shows a reproducible trend in which the center-most loop nucleotides are more reactive than nucleotides that abut constrained positions in the adjacent duplexes. These data together emphasize that SHAPE mapping reveals significant fine gradations in local nucleotide flexibility.

NMIA Modification Does *Not* Score Solvent Accessibility. We compared the relative reactivities obtained by SHAPE with calculated molecular surface areas at the ribose 2′-position, assuming a solvent-sized probe and using the crystallographic coordinates[14] for $tRNA^{Asp}$ (Figure 9). Consistent with the results obtained with the 11-mer oligonucleotide (Figure 6), 2′-O-adduct formation in $tRNA^{Asp}$ does not correlate with solvent accessibility. For example, the entire anticodon stem and loop structures, spanning nucleotides 26−44, have roughly uniform high solvent accessibility (≥20 $Å^2$ per ribose 2′-O atom) but feature both unreactive and highly reactive positions (compare gray bars with black line in Figure 9). Analogously, both the D- and T-loops include residues with low calculated static solvent accessibility but high 2′-hydroxyl reactivity; conversely, the acceptor and T-stems (positions 49−53 and 60−72) are solvent accessible but unreactive (Figure 9). 2′-Hydroxyl reactivity is thus uncorrelated with static solvent accessibility, strongly supporting models (Figure 1) that emphasize local conformational changes as the primary determinant of 2′-hydroxyl reactivity.

We also compare SHAPE with chemical mapping experiments[13a] employing the traditional base-reactive reagents, dimethyl sulfate (DMS) and diethylpyrocarbonate (DEPC). The correlation between nucleotides reactive by SHAPE and with DMS and DEPC is very strong (compare gray bars and filled circles in Figure 9). This is a striking conclusion considering that DMS and DEPC react largely at nucleophilic positions in purine bases,[2] while NMIA reacts at the RNA backbone and, again, strongly supports the model (Figure 1) that local nucleotide flexibility is the primary determinant for reactivity by SHAPE chemistry. The major difference occurs at position 68, where the guanosine residue in a G−U pair is reactive toward DMS, but not with the more selective SHAPE chemistry (Figure 9).

The direct comparison between traditional and SHAPE chemistries also emphasizes important advantages of SHAPE. First, all nucleotides are interrogated in a single reaction so there are no or fewer regions lacking structural information. Second, instead of simply classifying positions as reactive or not, SHAPE lends itself naturally to quantitative analysis (Figures 8B and 9).

2′-Ribose Chemistry, Local Nucleotide Flexibility, and the Influence of the 3′-Phosphodiester Anion. Reactions of 2′-hydroxyl groups (Figure 10A) and of 2′-amine substitutions (Figure 10B) in RNA to form ester or amide linkages, respectively, are both strongly sensitive to local nucleotide flexibility (Figures 5 and 8B and ref 6). We postulate that the physical basis for this effect in both cases stems predominantly from the influence of the proximal 3′-phosphodiester anion on reactivity at the 2′-position. The contribution of the 3′-phosphodiester potentially includes direct charge−charge electrostatic effects, differential hydrogen bonding, changes to the pK_a of the 2′-group, and generalized electrostatic and dielectric effects in the local environment.

However, the detailed reaction pathways for the 2′-ester- versus the 2′-amide-forming reactions are quite different. Reaction of a 2′-hydroxyl to form a 2′-ester proceeds via

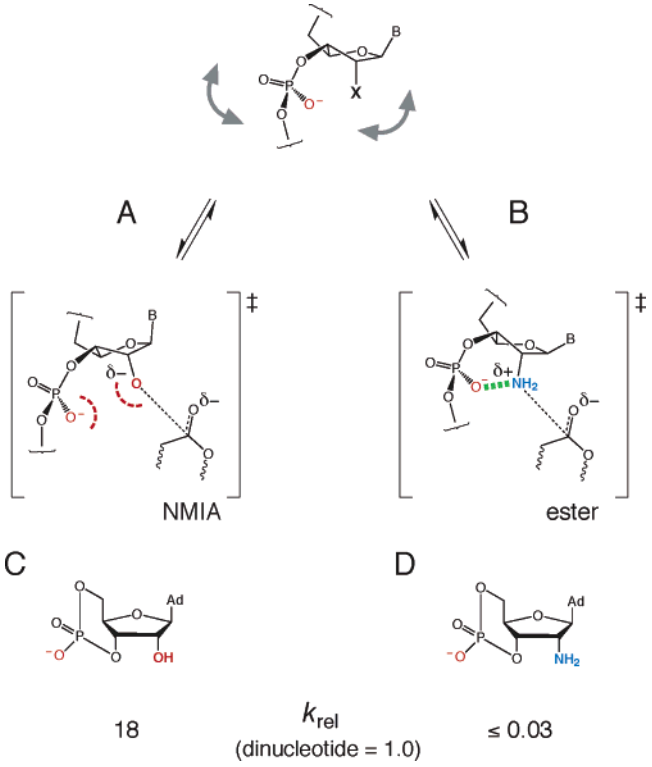


Figure 10. Distinct contributions of the 3′-phosphodiester anion for reaction of a flexible (gray arrows) ribose 2′-hydroxyl versus 2′-amine. Reaction pathways for (A) 2′-hydroxyl and (B) 2′-amine acylation. (C and D) Relative reactivities of cAMP analogues. Reference dinucleotides were $pA^{2'-OH}p$-ddC (Figure 5A) and $pU^{2'-NH_2}pU$,[6] respectively.

formation of an oxyanion in a *pre-equilibrium* step (Figure 1) and yields negative charge on the nucleophile in the transition state (Figure 10A). In contrast, 2′-amide formation involves *direct* attack of the unperturbed, 2′-NH_2 form[6,17] of the nucleophile and a transition state in which there is a partial positive charge on the amine nucleophile (Figure 10B).

A proximal 3′-phosphodiester anion thus *destabilizes* formation of the 2′-oxyanion formed in a pre-equilibrium step (see interference symbol in Figure 10A) but *stabilizes* the amine nucleophile (dashed green line in the transition state in Figure 10B). These contrasting effects are supported directly by nucleotide analogue studies (Figure 5 and ref 6) and especially by the cAMP analogues, which constrain the 3′-phosphodiester anion to be distant from a 2′-hydroxyl or 2′-amine nucleophile. Because the 3′-phosphodiester group electrostatically destabilizes the 2′-oxyanion, cAMP reacts (18-fold) faster than a reference dinucleotide (Figure 10C). In contrast, reaction of the 2′-amine, whose transition state is stabilized by the proximal anion, becomes undetectable (at least 30-fold slower) for 2′-amino-cAMP (Figure 10D). Flexible nucleotides sample a broad ensemble of conformations (gray arrows in Figure 10), distinct constellations are able to *both* minimize the electrostatic influence of the 3′-phosphodiester and facilitate 2′-oxyanion formation or, alternately, to align the 3′-group and facilitate the amide-forming transition state.

Perspective. SHAPE mapping takes advantage of the strong linkage between reactivity at the 2′-hydroxyl position and destabilization of nucleophilic chemistry at this site by the adjacent 3′-phosphodiester anion (Figures 1, 5, and 10). SHAPE footprinting affords a comprehensive and nucleotide resolution view of base pairing and tertiary interactions in an RNA that is generic to all four nucleotides (Figure 8). SHAPE reactivity amplitudes also measure quantitatively subtle differences in the degree to which a given nucleotide is constrained. RNA structure mapping chemistries based on reaction at the 2′-hydroxyl will likely make possible facile nucleotide resolution analysis of RNA as a function of ion environment, temperature, ligand binding, and conformational switches.

Experimental Section

General. Nucleotides (see Figure 5) ATP, AMP, 2′-deoxy-ATP, 3′-5′-cyclic-adenosine monophosphate (cAMP), and adenosine 3′-monophosphate were obtained from Sigma-Aldrich (St. Louis, MO); 2′,3′-dideoxy-ATP and 3′-(*O*-methyl)-ATP were from Trilink Biotech (San Diego, CA); 3′-(2′,3′-dideoxycytidine)-adenosine 5′-monophosphate and the 5′-UACUAACACC-ddC oligo and its complement (Figure 6) were synthesized by Dharmacon (Boulder, CO), and adenosine-3′-(*O*-ethyl)-phosphate was the gift of J. Chattopadhyaya (University of Uppsala). Radiolabeled (α-^{32}P) ATP and 2′- and 3′-deoxy-ATP were from Amersham. The pAp-ethyl and pAp-ddC were synthesized by 5′-[^{32}P]-labeling the appropriate precursor using polynucleotide kinase and were purified by polyacrylamide gel electrophoresis.[6]

Nucleotide Analogue and Model Duplex Reaction Kinetics. Nucleotide analogue reactions were performed in 100 mM potassium phosphate (pH 8.0), 250 mM NaCl, supplemented with 10% (v/v) NMIA (13 mM final, Molecular Probes) in dimethyl sulfoxide at 25 °C. For [^{32}P]-labeled nucleotides, reaction aliquots (10 000 cpm/μL) were quenched by addition to 1 vol 250 mM dithiothreitol and 2 vol formamide. Reaction products were resolved by gel electrophoresis (30% polyacrylamide; 29:1 acrylamide:bisacrylamide; 0.4 mm × 28.5 cm × 23 cm; 30 W; 1 h) and visualized by phosphorimaging (Molecular Dynamics). Reaction rates were obtained using an equation that accounts for parallel reaction of NMIA by 2′-O-adduct formation (k_{adduct}) and by hydrolysis ($k_{hydrolysis}$):

$$\text{fraction product} = 1 - \exp\left\{\frac{k_{adduct}}{k_{hydrolysis}}(e^{-k_{hydrolysis}t} - 1)\right\} \quad (1)$$

Nonradiolabeled analogues (10 μL, 1 mM in nucleotide) were allowed to react for five half-lives, lyophilized to dryness, resuspended in 600 μL of gradient buffer [100 mM potassium phosphate (pH 6)], and resolved by HPLC (Waters Xterra C-18 column, gradient to 50% (v/v) methanol) over 40 min. Nucleotides typically eluted within the first 15 min, while 2′-O-adducts eluted at ~35 min. In the limit of complete NMIA hydrolysis, eq 1 simplifies to $k_{adduct}/k_{hydrolysis} = -\ln(1 - \text{fraction product})$. Thus, the absolute rate of 2′-O-adduct formation is uniquely determined by the fraction product at the reaction endpoint.

Reaction of the 5′-[^{32}P]-labeled 11-mer oligonucleotide (0.55 μM, in 90 μL) was followed for both the single strand and duplex (complement concentration = 14 μM). RNAs were heated to 90 °C and allowed to cool to ambient temperature over 20 min. Reactions were initiated by addition of 10 μL NMIA (130 mM in DMSO), and aliquots were quenched by addition of an equal volume of formamide, 100 mM DTT, and 40 μM unlabeled 11-mer (to liberate the radiolabeled oligo from the duplex). Reaction aliquots were resolved on a (30%) denaturing polyacrylamide gel.

NMIA Hydrolysis. Reactions were initiated by adding 1 μL of 10 mM NMIA to reaction buffer (2 mL) in a fluorescence cuvette. Pseudo-first-order formation of the hydrolysis product, 2-methylaminobenzoate, was monitored by fluorescence (excitation 375 nm, emission 440 nm, 15 s time points).

(17) Verheyden, J. P. H.; Wagner, D.; Moffatt, J. G. *J. Org. Chem.* **1971**, *36*, 250−254.

Synthesis of tRNA^{Asp} Construct. A DNA template for transcription of tRNA^{Asp} in the context of the structure cassette was generated by PCR [1.5 mL; containing 20 mM Tris (pH 8.4), 50 mM KCl, 2.5 mM $MgCl_2$, 200 μM each dNTP, 500 nM each forward and reverse primer, 5 pM template, and 0.025 units/μL of Taq polymerase; denaturation at 94 °C, 45 s; annealing 55 °C, 30 s; and elongation 72 °C, 90 s; 34 cycles]. Primer and tRNA^{Asp} coding sequences are indicated in Figure 7. The PCR product was recovered by ethanol precipitation and resuspended in 300 μL of HE [10 mM Hepes (pH 8.0), 1 mM EDTA]. Transcription reactions (1.5 mL, 37 °C, 3 h) contained 40 mM Tris (pH 8.0), 10 mM $MgCl_2$, 10 mM DTT, 2 mM spermidine, 0.01% (v/v) Triton X-100, 4% (w/v) poly(ethylene) glycol 8000, 2 mM each NTP, 75 μL of PCR-generated template, and 0.1 mg/mL of T7 RNA polymerase. The RNA product was purified by denaturing (8%) polyacrylamide gel electrophoresis, excised from the gel, and recovered by electroelution and ethanol precipitation. The purified RNA ($\sim$2.5 μmol) was resuspended in 100 μL of HE.

Structure-Sensitive RNA Modification. RNA (20 pmol) in 6 μL of sterile water was heated at 95 °C for 3 min, quickly cooled on ice, treated with 3 μL of folding buffer [333 mM NaCl, 333 mM Hepes (pH 8.0), 33.3 mM $MgCl_2$], and incubated at 37 °C for 20 min. The RNA solution was treated with NMIA (1 μL, 130 mM in anhydrous DMSO, 35 °C), allowed to react for 54 min (equal to five NMIA half-lives), and placed on ice. Control reactions contained 1 μL of DMSO in place of NMIA. In prequench reactions, the RNA was added after NMIA was allowed to degrade by hydrolysis. Modified RNAs were subjected to primer extension without further purification.

Primer Extension. Modified RNA (2 μL, 4 pmol) was added to 5′-[^{32}P]-radiolabeled reverse transcription DNA primer (5′-^{32}P-GAA CCG GAC CGA AGC CCG, $\sim$0.2 pmol in 11 μL of H_2O), heated to 65 °C (6 min), and incubated at 35 °C (20 min). Reverse transcription buffer [6 μL; 167 mM Tris (pH 8.3), 250 mM KCl, 10 mM $MgCl_2$, 1.67 mM each dNTP, 16.7 mM DTT] was added; the RNA was heated to 52 °C; Superscript III (Invitrogen, 1 μL, 200 units) was added and mixed by gentle pipetting, and reactions were incubated at 52 °C for 5 min. Primer extension reactions were quenched by addition of 1 μL of 4 M NaOH, heating (95 °C for 5 min), subsequent addition of neutralizing gel loading solution (29 μL; 40 mM Tris/borate, 276 mM Tris/HCl, 5 mM EDTA, 0.01% (w/v) each xylene cyanol and bromophenol blue, 73% (v/v) formamide), and heating (95 °C, for an additional 5 min). Dideoxy sequencing markers were generated using unmodified RNA and by adding 1 μL of 10 mM dideoxy nucleotide triphosphate after addition of reverse transcription buffer. The cDNA extension products were separated by denaturing electrophoresis [90 mM Tris/borate, 2 mM EDTA, 7 M urea, 8% (29:1) acrylamide: bisacrylamide); 100 W, 5 min; 72 W, 2.5 h; 0.75 mm $\times$ 31 cm $\times$ 38.5 cm] and visualized by phosphorimaging. Two-dimensional gel images were converted to plots of intensity versus position using ImageQuant. Individual band intensities for the (+) and (−) NMIA reactions were integrated using SAFA.[18]

Acknowledgment. The work was supported by an NSF Career Award (MCB-9984289) to K.M.W., and we are indebted to the authors of SAFA for sharing this program prior to publication.

JA043822V

(18) Das, R.; Laederach, A.; Pearlman, S.; Herschlag, D.; Altman, R. B. SAFA: Semi-automated Footprinting Analysis Software for High-Throughput Quantification of Nucleic Acid Footprinting Experiments. *RNA* **2005**, *11*, 344−354.
(19) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379−400.