JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Fingerprinting Noncanonical and Tertiary RNA Structures by Differential SHAPE Reactivity

Kady-Ann Steen, Greggory M. Rice, and Kevin M. Weeks*

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

Supporting Information

ABSTRACT: Many RNA structures are composed of simple secondary structure elements linked by a few critical tertiary interactions. SHAPE chemistry has made interrogation of RNA dynamics at single-nucleotide resolution straightforward. However, de novo identification of nucleotides involved in tertiary interactions remains a challenge. Here we show that nucleotides that form noncanonical or tertiary contacts can be detected by comparing information obtained using two SHAPE reagents, *N*-methylisatoic anhydride (NMIA) and 1-methyl-6-nitroisatoic anhydride (1M6). Nucleotides that react preferentially with NMIA exhibit slow local nucleotide dynamics and usually adopt the less common C2′-endo ribose conformation. Experiments and first-principles calculations show that 1M6 reacts preferentially with nucleotides in which one face of the nucleobase allows an unhindered stacking interaction with the reagent. Differential SHAPE reactivities were used to detect noncanonical and tertiary interactions in four RNAs with diverse structures and to identify preformed noncanonical interactions in partially folded RNAs. Differential SHAPE reactivity analysis will enable experimentally concise, large-scale identification of tertiary structure elements and ligand binding sites in complex RNAs and in diverse biological environments.

RNA molecules are involved in nearly every aspect of cellular information transfer.[1] Information is encoded in both the RNA primary sequence and its three-dimensional structure. Higher-order RNA structures are typically composed of secondary structure elements held together by a few key tertiary interactions,[2,3] including long-range stacking, loop-loop and loop-helix contacts, and pseudoknots. Regions of an RNA that contain significant tertiary structures ultimately have numerous important functional roles.

Nucleotides that participate in either base-pairing or stable higher-order tertiary structure interactions can be detected by protection from solution-phase chemical probing reagents.[4] Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) has emerged as an especially informative approach for probing RNA structure and dynamics.[4,5] SHAPE chemistry exploits the discovery that the reactivity of the ribose 2′-hydroxyl is highly sensitive to local nucleotide flexibility (Figure 1A). Flexible nucleotides sample many conformations, a few of which preferentially react with hydroxyl-selective electrophilic reagents to form 2′-*O*-adducts (Figure 1A). The strong

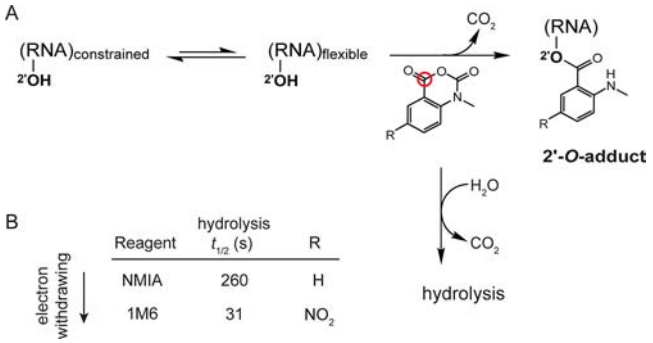


Figure 1. RNA SHAPE chemistry. (A) Mechanism in the context of the concurrent hydrolysis reaction. The red circle denotes the reactive center of the reagent. (B) SHAPE reagents and hydrolysis half-lives.

relationship between SHAPE reactivity and molecular motion[6] makes it possible to use this chemistry to achieve accurate secondary structure predictions, to monitor RNA dynamics and folding, and to explore RNA−protein interactions.[7−10]

However, it is not obvious from chemical reactivity of a nucleotide whether a given constraining interaction is a base-pairing or tertiary interaction. Nucleotides involved in tertiary interactions often have unusual backbone or stacking geometries,[3,11] adopt the syn conformation,[12] or undergo conformational changes on long time scales.[8,9] There is an unmet need for RNA chemical probing technologies that identify nucleotides involved in tertiary interactions. We report an approach for rapid "fingerprinting" of noncanonical and tertiary RNA structures using structure-selective SHAPE chemical reactivities.

We initially screened potential tertiary structure-selective SHAPE reagents using the aptamer domain of the thiamine pyrophosphate (TPP) riboswitch. The TPP riboswitch has been extensively characterized by crystallography,[13,14] in-solution dynamics,[15,16] and SHAPE chemistry.[17] This RNA contains many tertiary structure features that are common to highly structured RNAs, including base stacking, long-range docking interactions, and tight backbone turns.

Two reagents, *N*-methylisatoic anhydride (NMIA) and 1-methyl-6-nitroisatoic anhydride (1M6), proved promising. NMIA, one of the first reagents used in the SHAPE approach,[18] reacts slowly with RNA and can be used to identify nucleotides that undergo local conformational changes on long time scales (Figure 1B).[8] These nucleotides are usually in the relatively rare

Received: April 26, 2012
Published: August 1, 2012

   dx.doi.org/10.1021/ja304027m | *J. Am. Chem. Soc.* 2012, 134, 13160−13163

C2′-endo conformation and, in some cases, govern the folding of entire RNA domains.[9] The second reagent, 1M6, differs from NMIA by a single nitro group on the double ring system (Figure 1B). Addition of the electron-withdrawing group increases the electrophilicity of the reactive center (Figure 1A, red circle), and consequently, 1M6 reacts more rapidly than NMIA. Second, the nitro group significantly changes the electronic distribution of the reagent ring system. As we will show below, this allows 1M6 to stack with RNA nucleobases.

When the folded, ligand-bound TPP riboswitch was allowed to react with NMIA, the observed reactivities agreed with the known structure for the ligand-bound TPP riboswitch (Figure 2A and Figure S1A in the Supporting Information). When this

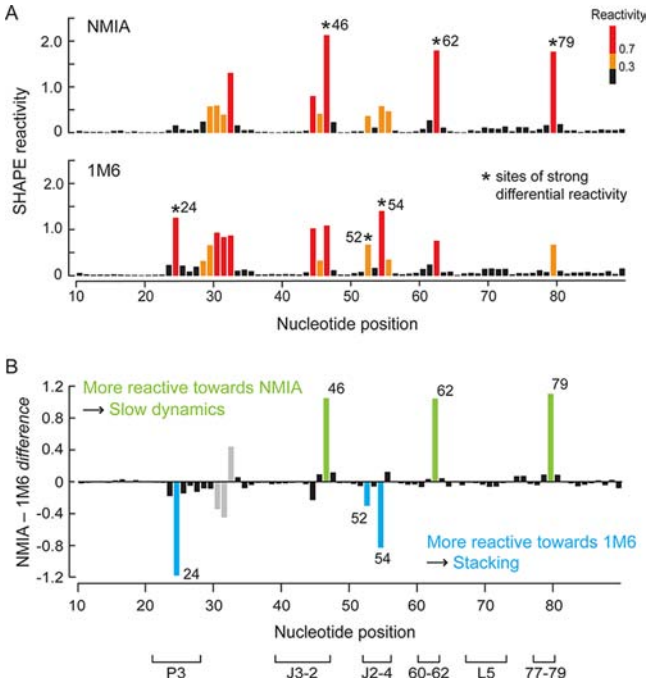


Figure 2. SHAPE analysis of the ligand-bound state of the TPP riboswitch. (A) Absolute SHAPE reactivities resulting from the reactions with (top) NMIA and (bottom) 1M6. Columns are colored by nucleotide reactivities. Asterisks indicate sites of strong differential reactivity. (B) Differential SHAPE reactivities calculated by subtracting the 1M6 profile from that of NMIA. Nucleotides exhibiting statistically significant differential reactivities (absolute reactivity difference of ≥0.3 SHAPE units and p-value of <0.05, calculated using Student's t test) are colored in green for NMIA and blue for 1M6. Gray columns represent nucleotides with differential reactivity that are involved in crystal contacts or have poorly defined electron density.

RNA was treated with 1M6, the overall SHAPE reactivity profile was very similar to that for NMIA (Figures 2A and S1B). In particular, all of the base-paired nucleotides were unreactive and many single-stranded nucleotides exhibited similar reactivities toward the two reagents. Critically, a few nucleotides exhibited strongly enhanced reactivity toward one of the two reagents (Figure 2A, asterisks). SHAPE chemistry is quantitative; therefore, reagent-specific reactivities can be identified simply by subtracting one profile from another. After excluding nucleotides that participate in crystal contacts or have poorly defined electron densities in the previously determined crystal structure (Figure 2B, gray columns), we identified six nucleotides exhibiting statistically significant differential reactivities to NMIA versus 1M6 (Figure 2B).

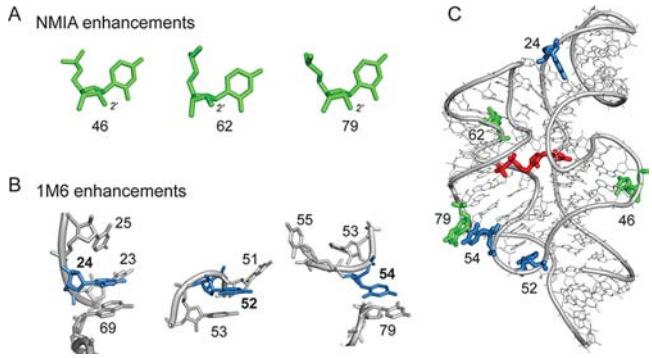


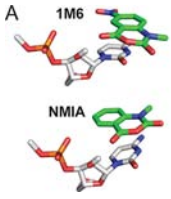
Figure 3. Conformations and structural contexts for nucleotides exhibiting differential reactivities in the ligand-bound state of the TPP riboswitch. (A) Sites of enhanced reactivity toward NMIA correspond to nucleotides in the C2′-endo ribose conformation. (B) Sites of 1M6 enhancement reflect one-sided stacking conformations. (C) NMIA (green) and 1M6 (blue) enhancements superimposed on a model for the TPP riboswitch aptamer domain with the bound ligand shown in red (PDB entry 2gdi).

The three nucleotides that react preferentially with NMIA (Figure 3A) each adopt the relatively rare C2′-endo conformation, consistent with previous studies.[8] The mechanism by which nucleotides might react preferentially with 1M6 has not been previously explored. These three nucleotides share the characteristic that one *face* of the nucleobase is available for $\pi-\pi$ stacking interactions with a small molecule such as 1M6 (Figure 3B). This conformation is unusual because in both A-form helices and most highly folded RNAs, base–base stacking is nearly fully saturated.[2,3] Only a few nucleotides in special structural contexts, especially at bulges, turns, and the termini of some helices, form "one-sided" stacking interactions.

We hypothesized that the nitro substituent polarizes the two-ring system to stabilize 1M6–nucleobase stacking interactions. We evaluated this model experimentally by varying the electron-withdrawing ability of the ring functional group of the reagent from slightly electron-donating (methyl) to moderately electron-withdrawing (bromo) to strongly electron-withdrawing (nitro) (Figures 1B and S2A). SHAPE reactivities of the one-sided stacking nucleotide C24 increased monotonically with increasing electron-withdrawing ability of the reagent substituent as reflected by its Hammett coefficient[19] (Pearson's linear r = 0.97; Figure S2B). In contrast, this trend was not observed for A45, which is also reactive toward SHAPE reagents but forms stacking interactions on both sides of the adenine base. These reactivity patterns are consistent with the formation of favorable reagent–nucleobase stacking interactions,[20] which are possible at C24 but not A45.

Since the effect of electron-withdrawing groups on the stacking interaction and resulting SHAPE reactivity is quantitative, we estimated representative electronic contributions for this interaction through first-principles calculations (Figure 4). Complexes formed between 1M6 and the four RNA nucleotides were −2 to −5 kcal/mol more stable than those formed with NMIA (Figure 4B). These values are comparable to the net stabilization energy of 2–3 kcal/mol for a two-base-pair stack.[21] Favorable stacking appears to enhance 1M6 reactivity by increasing the local reagent concentration at nucleotides where one face is available for a one-sided stacking interaction.

In sum, the three nucleotides with enhanced reactivity toward NMIA have unusual ribose geometries that either



B

		Energies (kcal/mol)		
	Sugar pucker	1M6	NMIA	Stabilization
Cytidine	C3′-endo	-23.0	-18.3	-4.7
Guanosine	C3′-endo	-28.6	-24.0	-4.6
Adenosine	C2′-endo	-21.5	-17.8	-3.7
Uridine	C2′-endo	-20.2	-18.1	-2.1

Figure 4. Electronic structure calculations of the stabilities of 1M6– and NMIA–nucleotide complexes. (A) The most stable stacking conformations for the cytidine–1M6 and –NMIA complexes. (B) Representative stacking complex energies and net stabilization energies as functions of nucleotide, ribose conformation, and reagent.

stabilize the RNA fold (nucleotide 79; Figure 3C) or are adjacent to the ligand binding pocket (nucleotides 46 and 62). Each nucleotide with enhanced reactivity toward 1M6 forms either a long-range stacking interaction (nucleotides 24 and 54; Figure 3C) or lies at a turn in the backbone (nucleotide 52). These differential reactivities comprise a noncanonical and tertiary RNA structure "fingerprint" highlighting nucleotides that form structurally critical interactions.

To examine the generality of differential SHAPE reactivity analysis, we probed three additional RNAs with well-characterized structures: the adenine and lysine riboswitches and RNase P specificity domain[8,22] (Figure S3). Differential reactivities in the adenine riboswitch were found in the ligand binding pocket and in a loop–loop interaction (nucleotides 47 and 63, respectively; Figure S3A). In the lysine riboswitch, the three differentially reactive nucleotides are located in the P2a-L2 turn motif, which stabilizes the folded riboswitch structure[23] (nucleotides 38, 40, and 52; Figure S3B). The RNase P specificity domain has many nucleotides with significant differential reactivities (nucleotides 130 and 194 are involved in long-range interactions, 142 is in a sarcin–ricin motif, and 186 is adjacent to a T-loop motif;[24] Figure S3C).

In the four RNAs analyzed in this study, 18 nucleotides displayed statistically significant reagent-specific differential reactivities (Figures 3 and S3). Ten nucleotides displayed preferential reactivity with NMIA and presumably experience slow local dynamics[8] (Figures 3 and S3, emphasized in green). Of these 10 nucleotides, nine are in the C2′-endo ribose conformation (nucleotide 180 in the RNase P RNA is modeled as C3′-endo). Eight nucleotides reacted preferentially with 1M6 (Figures 3 and S3, emphasized in blue), and seven represent clear examples in which one face of the nucleobase is accessible for a stacking interaction; the remaining nucleotide (position 142 in the RNase P RNA) forms a partial one-sided stacking conformation. These 18 nucleotides were detected out of ~470 RNA nucleotides probed, and strong observed correlation between reactivity and location in structurally distinctive regions highlights the ability of differential SHAPE reactivity analysis to identify unique nucleotide dynamics and conformations on a large scale.

We next examined the TPP riboswitch aptamer domain in the less structured ligand-free state (Figure 5A). These experiments were motivated in part by studies suggesting that the J3-2 region of the TPP binding pocket may exist in a preformed state[14,15,25] and that the ligand-free states of many riboswitches have significant structure.[25,26] Two features were immediately evident. First, the magnitudes of the reagent-specific differential reactivities in the ligand-free state were smaller than in the ligand-bound state, consistent with a model in which local tertiary structure motifs in the ligand-free state

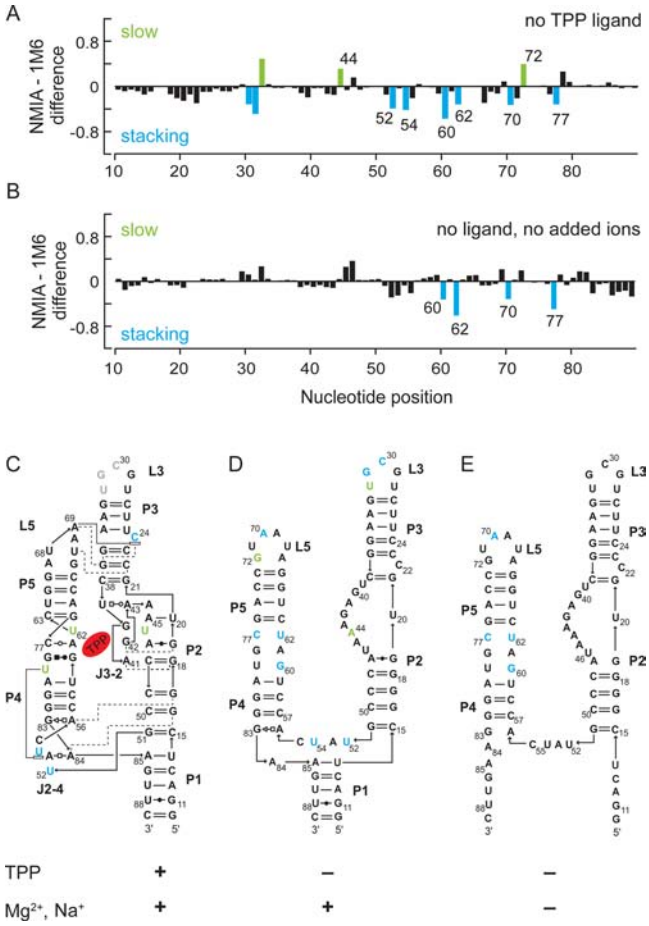


Figure 5. Differential SHAPE reactivities as a function of the extent of stable RNA structure. (A, B) Differential SHAPE reactivity analyses of the (A) ligand-free and (B) ligand-free with no-added-ion states of the TPP riboswitch. Nucleotides that showed significant reagent-specific enhancements (as defined in Figure 2) are highlighted. (C–E) Reagent enhancements superimposed on secondary structure models for the (C) ligand-bound, (D) ligand-free, and (E) no-added-ion states of the TPP riboswitch.

form transiently or less stably than in the ligand-bound state. Second, there were *more* nucleotides with differential reactivities in the ligand-free state, especially 1M6 enhancements, in comparison with the ligand-bound RNA state (Figures 2B and 5A, numbered positions). This observation suggests that the TPP riboswitch samples diverse structured microstates in the absence of ligand.

We superimposed the differential reactivities onto secondary structure models for the ligand-bound (Figure 5C) and ligand-free states (Figure 5D). The differential reactivities in the ligand-free state occurred predominantly at nucleotides in two structural contexts: (*i*) at or adjacent to the TPP ligand binding pocket (nucleotides 44, 60, 62, and 77; Figure 5D) and (*ii*) at nucleotides that form key tertiary structures in the fully folded ligand-bound structure (nucleotides 52, 54, 70, and 72). The differential reactivity data thus support a model in which a population of ligand-free RNA molecules adopts a TPP binding pocket structure similar to that of the ligand-bound RNA (compare J3-2 and nucleotides 60–62 and 77–79 in Figure 5C,D). A preformed binding pocket likely facilitates rapid binding by the ligand, consistent with experiments showing fast TPP binding and ligand-induced folding.[15]

dx.doi.org/10.1021/ja304027m | J. Am. Chem. Soc. 2012, 134, 13160–13163

Finally, we probed the ligand-free TPP riboswitch in the absence of added monovalent (Na^+) or divalent (Mg^{2+}) ions and generated an experimentally supported secondary structure model[7] for this RNA state. The major structural difference occurred in the P1 helix. In the absence of ions, nucleotides in the P1 helix were reactive toward SHAPE reagents, suggesting that this helix does not form stably (Figure 5E). We observed four statistically significant differential reactivities in the ion-free TPP RNA (nucleotides 60, 62, 70, and 77; Figure 5B,E). All four nucleotides that exhibited enhanced reactivity toward 1M6 were also observed in the ligand-free +Mg^{2+} state of the TPP riboswitch. These observations support a model in which the P4−P5 side of the TPP ligand binding pocket is partially prestructured independent of the ion environment.

Differential SHAPE reactivity probing of the ligand-bound, ligand-free, and low-ion states of the TPP riboswitch demonstrated that, as the extent of higher-order structure was systematically decreased, the magnitude and number of reagent-specific differential reactivities also decreased. Nonetheless, even in apparently loosely structured states, the TPP riboswitch RNA samples conformations that allow for favorable, potentially transient, tertiary structure interactions involved in ligand binding.

Differential SHAPE analysis uses a dual-reagent detection scheme that takes advantage of two fundamental features of local RNA structure: (1) some nucleotides are constrained in special environments, causing them to undergo structural rearrangements on long time scales, and (2) although the vast majority of nucleotides in folded RNAs have fully "saturated" base-stacking interactions, a few nucleotides form conformations that leave one face of the nucleobase accessible for a binding interaction. To a first-order approximation, these two features are selectively detectable by enhanced reactivity toward NMIA and 1M6, respectively. The longer lifetime of NMIA in solution prior to degradation by hydrolysis provides a longer window for nucleotides experiencing slow nucleotide dynamics to achieve a SHAPE-reactive conformation[8] (Figure 1). Enhanced reactivity toward 1M6 reflects preferential reagent stacking at accessible nucleobase sites (Figures 3B, 4, S2).

In sum, differential SHAPE analysis enables rapid identification of nucleotides involved in key noncanonical and tertiary interactions in RNA essentially independent of the RNA size, complexity, and dynamics. This approach detects the local dynamic and structural differences that are characteristic of important tertiary structure motifs and yields information complementary to hydroxyl radical footprinting,[27] which detects tertiary interactions by virtue of their ability to shield nucleotides from reaction with a solvent-based probe.

We anticipate that differential SHAPE reactivity analysis and the resulting RNA tertiary structure fingerprints will prove especially useful in ongoing efforts to improve secondary and tertiary structure prediction, for identification of ligand and protein binding sites, and in large-scale searches for tertiary interactions in complex RNAs.

■ ASSOCIATED CONTENT

Supporting Information

Experimental methods, additional discussion, and three figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

weeks@unc.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by grants from the NSF (MCB-0919666) and NIH (GM064803) to K.M.W. We are grateful to Alexander Serganov for providing electron density maps for the riboswitch RNAs and to Shubin Liu for assistance during early phases of the electronic structure calculations.

■ REFERENCES

(1) (a) Sharp, P. A. *Cell* **2009**, *136*, 577. (b) Darnell, E. J. *RNA: Life's Indispensable Molecule*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2011.

(2) Leontis, N. B.; et al. *Curr. Opin. Struct. Biol.* **2006**, *16*, 279.

(3) Butcher, S. E.; Pyle, A. M. *Acc. Chem. Res.* **2011**, *44*, 1302.

(4) Weeks, K. M. *Curr. Opin. Struct. Biol.* **2010**, *20*, 295.

(5) Weeks, K. M.; Mauger, D. M. *Acc. Chem. Res.* **2011**, *44*, 1280.

(6) Gherghe, C. M.; Shajani, Z.; Wilkinson, K. A.; Varani, G.; Weeks, K. M. *J. Am. Chem. Soc.* **2008**, *130*, 12244.

(7) Deigan, K. E.; Li, T. W.; Mathews, D. H.; Weeks, K. M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 97.

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

(9) Mortimer, S. A.; Weeks, K. M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15622.

(10) Duncan, C. D.; Weeks, K. M. *Biochemistry* **2010**, *49*, 5418.

(11) (a) Holbrook, S. R. *Annu. Rev. Biophys.* **2008**, *37*, 445. (b) Stombaugh, J.; Zirbel, C. L.; Westhof, E.; Leontis, N. B. *Nucleic Acids Res.* **2009**, *37*, 2294.

(12) Sokoloski, J. E.; Godfrey, S. A.; Dombrowski, S. E.; Bevilacqua, P. C. *RNA* **2011**, *17*, 1775.

(13) (a) Serganov, A.; Polonskaia, A.; Phan, A. T.; Breaker, R. R.; Patel, D. J. *Nature* **2006**, *441*, 1167. (b) Thore, S.; Leibundgut, M.; Ban, N. *Science* **2006**, *312*, 1208. (c) Kulshina, N.; Edwards, T. E.; Ferre-D'Amare, A. R. *RNA* **2010**, *16*, 186.

(14) Edwards, T. E.; Ferre-D'Amare, A. R. *Structure* **2006**, *14*, 1459.

(15) Lang, K.; et al. *Nucleic Acids Res.* **2007**, *35*, 5370.

(16) (a) Baird, N. J.; Ferre-D'Amare, A. R. *RNA* **2010**, *16*, 598. (b) Ali, M.; et al. *J. Mol. Biol.* **2010**, *396*, 153.

(17) Steen, K. A.; et al. *J. Am. Chem. Soc.* **2010**, *132*, 9940.

(18) Merino, E. J.; Wilkinson, K. A.; Coughlan, J. L.; Weeks, K. M. *J. Am. Chem. Soc.* **2005**, *127*, 4223.

(19) Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; Wiley-Interscience: New York, 1979.

(20) (a) Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. *J. Chem. Soc., Perkin Trans. 2* **2001**, 651. (b) Mignon, P.; Loverix, S.; Steyaert, J.; Geerlings, P. *Nucleic Acids Res.* **2005**, *33*, 1779.

(21) Petersheim, M.; Turner, D. H. *Biochemistry* **1983**, *22*, 256.

(22) (a) Serganov, A.; Yuan, Y. R.; Pikovskaya, O.; Polonskaia, A.; Malinina, L.; Phan, A. T.; Hobartner, C.; Micura, R.; Breaker, R. R.; Patel, D. J. *Chem. Biol.* **2004**, *11*, 1729. (b) Serganov, A.; Huang, L.; Patel, D. J. *Nature* **2008**, *455*, 1263. (c) Krasilnikov, A. S.; Yang, X.; Pan, T.; Mondragon, A. *Nature* **2003**, *421*, 760.

(23) Blouin, S.; Lafontaine, D. A. *RNA* **2007**, *13*, 1256.

(24) Krasilnikov, A. S.; Mondragon, A. *RNA* **2003**, *9*, 640.

(25) Montange, R. K.; Batey, R. T. *Annu. Rev. Biophys.* **2008**, *37*, 117.

(26) Haller, A.; et al. *Acc. Chem. Res.* **2011**, *44*, 1339.

(27) (a) Brenowitz, M.; Chance, M. R.; Dhavan, G.; Takamoto, K. *Curr. Opin. Struct. Biol.* **2002**, *12*, 648. (b) Tullius, T. D.; Greenbaum, J. A. *Curr. Opin. Chem. Biol.* **2005**, *9*, 127.