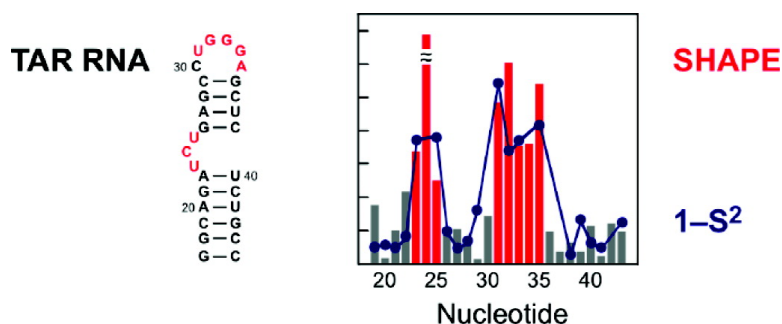## Communication

Strong Correlation between SHAPE Chemistry and the Generalized NMR Order Parameter (*S*) in RNA

Costin M. Gherghe, Zahra Shajani, Kevin A. Wilkinson, Gabriele Varani, and Kevin M. Weeks

More About This Article

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

- Supporting Information
- Links to the 1 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

Strong Correlation between SHAPE Chemistry and the Generalized NMR Order Parameter (S^2) in RNA

Costin M. Gherghe,[†] Zahra Shajani,[‡] Kevin A. Wilkinson,[†] Gabriele Varani,[‡] and Kevin M. Weeks*,[†]

Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290 and Department of Chemistry, University of Washington, Seattle, Washington 981950-1700

Received June 19, 2008; E-mail: weeks@unc.edu

RNA molecules perform important cellular functions that depend on the ability to form both rigid and dynamic structural elements and that often require large changes in conformation and motion. Critical examples include ribosomal protein synthesis, ribonucleoprotein assembly, and riboswitches.[1] The intrinsic RNA motions that underlie these processes occur over a wide range of time scales, from very fast ps motions to global conformational changes that require minutes.[2]

A number of NMR methods have been devised to measure RNA dynamics at single nucleotide resolution.[2a,3] One of the most useful is the model-free framework, originally developed for protein motions,[4] that interprets T1, T2, and NOE measurements in terms of two physically intuitive parameters: an effective correlation time for molecular tumbling and the generalized order parameter, S^2, which describes the spatial restriction of motion on a per-residue basis. S^2 can adopt values ranging from 0 (completely disordered) to 1 (fully ordered) (Figure 1A).[4]

Detailed motions at nucleotide resolution have been described for paradigmatic RNAs,[5,6] including the TAR element from HIV,[5a−c] a regulatory element that binds the U1A protein,[6] a stem loop derived from the U6 RNA,[5d] the lead-dependent ribozyme,[5e] and tetraloop-containing RNAs.[5f−i] These studies have yielded important information regarding RNA dynamics both alone and in complex with small molecule ligands[5a] and proteins.[6b] However, analysis of RNA dynamics by NMR is limited to small and spectroscopically well-behaved RNA motifs of ∼45 nucleotides or less. To analyze single-nucleotide resolution dynamics for large RNAs alone, as part of ribonucleoproteins, or *in vivo*, a different and more versatile approach is required.

Local motion in nearly any RNA can be easily measured at nucleotide resolution using the simple and rapid SHAPE (selective 2′-hydroxyl acylation analyzed by primer extension) experiment.[7] SHAPE chemistry is based on the discovery that conformationally flexible nucleotides react preferentially toward acylating agents such as 1-methyl-7-nitroisatoic anhydride (1M7) (Figure 1B) to form a 2′-O-adduct. In contrast, nucleotides that are constrained by base pairing or tertiary interactions are unreactive. Sites of 2′-O-adduct formation are then detected as stops to primer extension.[7]

SHAPE is proving to be a powerful approach for addressing a wide variety of structure−function relationships in RNA, from short oligonucleotides to RNAs thousands of nucleotides long.[7,8] To date, the correlation between SHAPE reactivity and local RNA motion, while clearly plausible, has not been rigorously established. In contrast, S^2 is derived from a well-understood quantitative framework[4] and is strongly correlated with local molecular motions and structure.[9]

We therefore sought to test whether SHAPE chemistry captures local nucleotide dynamics in a way that correlates with S^2, a parameter that possesses clear physical meaning. SHAPE and NMR analyses were conducted in parallel for three RNAs: (1) the TAR RNA from HIV-1 which activates transcriptional elongation in concert with the Tat protein; (2) the U1A protein binding site RNA that autoregulates mRNA processing; and (3) the *Tetrahymena* telomerase stem loop 4
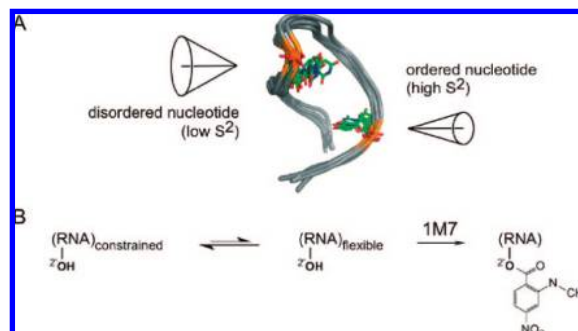

Figure 1. Schemes for interpreting (A) the generalized order parameter S^2 and (B) RNA SHAPE chemistry.
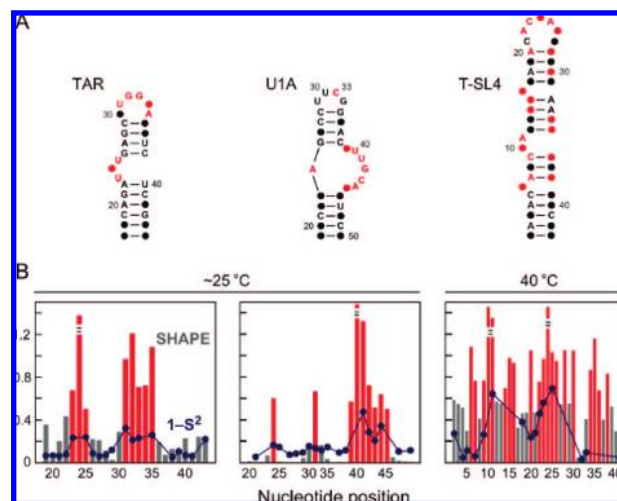

Figure 2. Local nucleotide structure in RNA analyzed by SHAPE and S^2. (A) Secondary structures for the TAR, U1A target, and T-SL4 RNAs. Nucleotides with SHAPE reactivities greater than 0.5 for TAR and U1A or greater than 0.6 for T-SL4 are red. Positions for which S^2 was not obtained are shown as red or black spheres. (B). Histograms of SHAPE reactivities as a function of nucleotide position (columns) compared to $1 − S^2$ measurements (blue spheres).

RNA (T-SL4) which promotes folding of a pseudoknot essential for enzyme activity (Figure 2A).[6,10]

S^2 values based on ^{13}C relaxation at the C1′ ribose position were calculated[4,6a] for all well-resolved nucleotides in the U1A and TAR RNAs and for C and A nucleotides in the T-SL4 RNA (results represented as black and red nucleotides in Figure 2A and blue spheres in Figure 2B). We also measured SHAPE reactivities for these RNAs under conditions similar to those used in the NMR experiments. SHAPE reactivities are normalized to a scale from 0 to ∼1.5, in which 1.0 is defined as the average intensity of highly reactive positions[8b] (bars, Figure 2B).

We then compared SHAPE reactivities with S^2, for all positions where both measurements could be made (compare blue spheres with

---

[†] University of North Carolina.
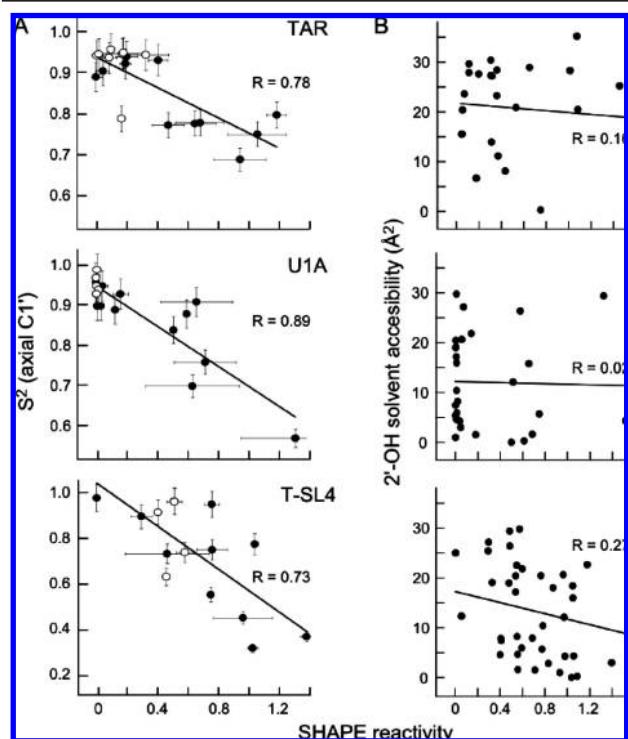[‡] University of Washington.

Figure 3. Quantitative correlation between SHAPE reactivity and (A) S^2, measured at C1′, or (B) solvent accessibility of the 2′-hydroxyl group. Pearson's linear R-values are shown.

columns, Figure 2B). Because high SHAPE reactivities and low S^2 values both correspond to a disordered site, we plot the generalized order parameter as $1 - S^2$. For the TAR RNA, the three-nucleotide UCU bulge and the apical loop are both reactive by SHAPE and are also disordered, as indicated by high $1 - S^2$ values. For the U1A RNA, the 39−45 loop is clearly identifiable by both SHAPE and S^2 values. Importantly, three of four nucleotides in the 29−34 loop are both unreactive by SHAPE and have S^2 values that reflect a high level of order, in agreement with previous observations that UUCG tetraloops are unusually stable, reflecting intramolecular base stacking and hydrogen bonding interactions.[11] SHAPE experiments with T-SL4 were performed at 40 °C, as required by the NMR analysis of this RNA. Nucleotides in the T-SL4 RNA are generally more reactive than those in the TAR and U1A target RNAs, consistent with a temperature-induced increase in RNA dynamics (compare panels, Figure 2B). For T-SL4, both SHAPE and S^2 detect increased nucleotide dynamics in the apical loop and at each of the three small bulges in this RNA.

SHAPE and $1 - S^2$ do not correlate well at three nucleotides: A24 and C33 of U1A and A20 of T-SL4. However, all three are single stranded or adjacent to unpaired positions. Therefore, the high observed SHAPE reactivity likely better represents the local structure in these regions than does the S^2 measurement.

We quantified the correlation between SHAPE reactivities and S^2 by plotting these values on a linear scale and calculating Pearson correlation coefficients, R (Figure 3A). We analyzed correlation coefficients in two ways. First, we determined R using all measured values and observed a strong correlation between SHAPE reactivities and S^2 in all cases: R values are 0.78, 0.89, and 0.73 for the TAR, U1A, and T-SL4 RNAs, respectively (all points, Figure 3A). We also calculated R values after excluding nucleotides that form canonical base pairs, which always have near-zero SHAPE reactivities and high S^2 values. Correlation coefficients from this more stringent calculation (filled spheres, Figure 3A) had similar correlations: R values are 0.79, 0.86, and 0.75 for the TAR, U1A, and T-SL4 RNAs, respectively.

S^2 is derived primarily from NMR measurements sensitive to ps−ns motions while SHAPE chemistry is likely to be additionally influenced by dynamics on slower time scales. Nucleotides that are disordered on these slower time scales, but whose motions are not detected in the NMR experiments, would fall above the line in our correlation plots (Figure 3A).

Finally, we assessed whether SHAPE chemistry is influenced by the solvent accessibility of the 2′-hydroxyl position. This is a critical control to establish whether SHAPE might also report a reactive spatial orientation of the 2′-OH group in addition to measuring RNA dynamics. There is essentially no correlation between SHAPE reactivity and solvent accessibility at the 2′-hydroxyl group for any of the three RNAs (Figure 3B).

We conclude that local disorder at individual RNA nucleotides as quantified by S^2 correlate strongly with SHAPE reactivities but not with solvent accessibility. SHAPE thus measures spatial disorder and structural dynamics at single nucleotide resolution in RNA. We anticipate that SHAPE chemistry will create many opportunities for understanding the roles of individual nucleotide dynamics in the structure of large RNAs, during ribonucleoprotein assembly, and upon RNA binding by proteins and other ligands and drugs, both *in vitro* and *in vivo*.

Supporting Information Available: Experimental procedures and additional information regarding the NMR experiments with T-SL4. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Korostelev, A.; Noller, H. F. *Trends Biochem. Sci.* **2007**, *32*, 434–441. (b) Buchmueller, K. L.; Webb, A. E.; Richardson, D. A.; Weeks, K. M. *Nat. Struct. Biol.* **2000**, *7*, 362–366. (c) Maity, T. S.; Weeks, K. M. *J. Mol. Biol.* **2007**, *369*, 512–524. (d) Bokinsky, G.; Nivón, L. G.; Liu, S.; Chai, G.; Hong, M.; Weeks, K. M.; Zhuang, X. *J. Mol. Biol.* **2006**, *361*, 771–784. (e) Kim, J. N.; Breaker, R. R. *Biol. Cell* **2008**, *100*, 1–11.

(2) (a) Shajani, Z.; Varani, G. *Biopolymers* **2007**, *86*, 348–359. (b) Buchmueller, K. L.; Hill, B. T.; Platz, M. S.; Weeks, K. M. *J. Am. Chem. Soc.* **2003**, *125*, 10850–10861. (c) Al-Hashimi, H. M.; Walter, N. G. *Curr. Opin. Struct. Biol.* **2008**, *18*, 321–329.

(3) Guéron, M.; Leroy, J. L. *Methods Enzymol.* **1995**, *261*, 383–413.

(4) (a) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546–4559. (b) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559–4570.

(5) (a) Dayie, K. T.; Brodsky, A. S.; Williamson, J. R. *J. Mol. Biol.* **2002**, *317*, 263–278. (b) Vallurupalli, P.; Scott, L.; Hennig, M.; Williamson, J. R.; Kay, L. E. *J. Am. Chem. Soc.* **2006**, *128*, 9346–9347. (c) Hansen, A. L.; Al-Hashimi, H. M. *J. Am. Chem. Soc.* **2007**, *129*, 16072–16082. (d) Blad, H.; Reiter, N. J.; Abildgaard, F.; Markley, J. L.; Butcher, S. E. *J. Mol. Biol.* **2005**, *353*, 540–555. (e) Hoogstraten, C. G.; Wank, J. R.; Pardi, A. *Biochemistry* **2000**, *39*, 9951–9958. (f) Hall, K. B.; Tang, C. *Biochemistry* **1998**, *37*, 9323–9332. (g) D'Souza, V.; Dey, A.; Habib, D.; Summers, M. F. *J. Mol. Biol.* **2004**, *337*, 427–442. (h) Duchardt, E.; Schwalbe, H. *J. Biomol. NMR* **2005**, *32*, 295–308. (i) Ferner, J.; Villa, A.; Duchardt, E.; Widjajakusuma, E.; Wöhnert, J.; Stock, G.; Schwalbe, H. *Nucleic Acids Res.* **2008**, *36*, 1928–1940.

(6) (a) Shajani, Z.; Varani, G. *J. Mol. Biol.* **2005**, *349*, 699–715. (b) Shajani, Z.; Drobny, G.; Varani, G. *Biochemistry* **2007**, *46*, 5875–5883.

(7) (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.

(8) (a) Mortimer, S. A.; Weeks, K. M. *J. Am. Chem. Soc.* **2007**, *129*, 4144–4145. (b) 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. (c) Wang, B.; Wilkinson, K. A.; Weeks, K. M. *Biochemistry* **2008**, *47*, 3454–3461. (d) Duncan, C. D. S.; Weeks, K. M. *Biochemistry* **2008**, *47*, 8504–8513. (e) Gherghe, C. M.; Mortimer, S. A.; Krahn, J. M.; Thompson, N. L.; Weeks, K. M. *J. Am. Chem. Soc.* **2008**, *130*, 8884–8885.

(9) (a) Palmer, A. G. *Chem. Rev.* **2004**, *104*, 3623–3640. (b) Jarymowycz, V. A.; Stone, M. J. *Chem. Rev.* **2006**, *106*, 1624–1671.

(10) (a) Sharp, P. A.; Marciniak, R. A. *Cell* **1989**, *59*, 229–230. (b) Hall, T. M. *Curr. Opin. Struct. Biol.* **2002**, *12*, 82–88. (c) Sperger, J. M.; Cech, T. R. *Biochemistry* **2001**, *40*, 7005–7016. (d) Chen, Y.; Fender, J.; Legassie, J. D.; Jarstfer, M. B.; Bryan, T. M.; Varani, G. *EMBO J.* **2006**, *25*, 3156–3166.

(11) Cheong, C.; Varani, G.; Tinoco, I. *Nature* **1990**, *346*, 680–682.

JA804541S