A Forced March across an RNA Folding Landscape

Scott K. Silverman[1,*]
[1]Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA
*Correspondence: scott@scs.uiuc.edu
DOI 10.1016/j.chembiol.2008.02.014

In a recent issue of *Science*, Greenleaf et al. (2008) report single-molecule force measurements to explore the sequential folding landscape of an adenine riboswitch aptamer domain. This study provides an exceptionally quantitative view of how an RNA molecule folds.

Single-molecule force measurements are increasingly used to investigate enzyme catalysis and protein folding landscapes (Cornish and Ha, 2007; Tinoco and Bustamante, 2002). The authors of a new report in *Science* (Greenleaf et al., 2008) previously used single-molecule force measurements to analyze the folding of individual DNA hairpin molecules (Woodside et al., 2006). In the present study, these investigators have probed a multistep RNA folding pathway in unprecedented quantitative detail. Their specific focus is the aptamer domain of the *pbuE* adenine riboswitch (Lemay et al., 2006; Wickiser et al., 2005), one of a growing number of riboswitches that rely on interactions between RNA and small-molecule metabolites—such as the purine nucleobases, adenine and guanine—to control gene expression (Mandal and Breaker, 2004). The new results provide detailed insight into the sequential folding pathway of this particular riboswitch, and they suggest a novel and likely general means of investigating cotranscriptional RNA folding.

RNA adopts three levels of structure: (1) primary sequence; (2) secondary structure, in which only Watson-Crick base pairs are present; and (3) tertiary structure, which depends on metal ions such as Mg^{2+}. For purine riboswitches, binding of the purine is also required for tertiary stabilization. Unlike for proteins, RNA secondary structure elements such as stem-loops (hairpins) are generally stable in isolation. As a consequence, the levels of RNA structure are ''hierarchical'' in that tertiary structure is usually formed from prefolded secondary structure components (Figure 1A; Tinoco and Bustamante, 1999). Despite this generally accepted hierarchical model, we still lack a deep understanding of the folding pathways by which RNAs attain functionally relevant tertiary structures.

The new report (Greenleaf et al., 2008) addresses this situation with single-molecule force measurements that explore folding of the adenine riboswitch aptamer domain, which has three double-helical (paired) regions of secondary structure denoted P1–P3. The adenine metabolite is engulfed entirely by the aptamer, as demonstrated by X-ray crystallography of several related riboswitches (Batey et al., 2004; Serganov et al., 2004). The P2 and P3 regions interact by tertiary contacts between the loops that terminate each stem. The new report uses force measurements to probe the RNA folding landscape traversed by individual riboswitch aptamer molecules as they interconvert from a fully unfolded state, which lacks any of P1–P3 and does not bind adenine, to the fully folded state, which has each of P1–P3 and also binds adenine (Figure 1B).

The aptamer was transcribed by a single *E. coli* RNA polymerase (RNAP) molecule held in an optical trap, with the leading nonriboswitch portion of the RNA held in a second optical trap by hybridization to a complementary DNA ''handle''. This setup allowed force to be applied to the nascent RNA transcript by moving apart the two traps with subnanometer spatial resolution. Two types of measurements were made: force-extension curves (FECs) measuring the relationship between the force applied to the molecule and its contour length, and extension records as a function of time at constant force.

The aptamer was first allowed to fold fully and was then unfolded by pulling. Hundreds of FECs from the same molecule were used to probe changes in contour length as various structural elements unfolded. The data suggested that an intermediate state was sometimes formed during unfolding. From the length change in nanometers, the intermediate was assigned as a state in which P2 is intact but P3 has been pulled apart; continued pulling then disrupted P2 as well. The work done to unfold the RNA allowed quantification of the equilibrium free energy of each folded state, thereby defining the location of the state on the y axis of a conventional energy diagram. Because the force measurements inherently involve a distance component, the data also allowed quantitative placement of the various observed states along the x axis of the same energy diagram. For example, from the distance to the first unfolding transition state as determined by fits to the FEC data (2.1 nm), the authors estimated that between 2 and 3 base pairs of the least stable P1 helix are disrupted to reach this transition state.

The aptamer was also allowed to refold when starting from the fully extended and hence unfolded state, monitoring the time-dependent molecular extension at constant force as the force was lowered in gradual steps. By carefully examining the four observed folding transitions and combining all of the unfolding and refolding data, the authors arrived at a landscape model that invokes five different states, in order from least to most folded (Figure 1B): the fully unfolded state; a state with P2 only; a state with both P2 and P3 folded separately but not interacting; a state with P2 and P3 interacting via tertiary contacts but P1 not yet formed; and finally the fully folded state with all of P1–P3 in place. The penultimate state on this pathway was identified as an ''adenine-competent'' state, because it is preorganized to bind adenine although it lacks P1. The final state on this pathway (which has each of P1–P3 intact) could be observed either with or without adenine bound, consistent with other work that has revealed preorganization of the aptamer for purine binding (Noeske et al., 2007; Ottink et al., 2007).

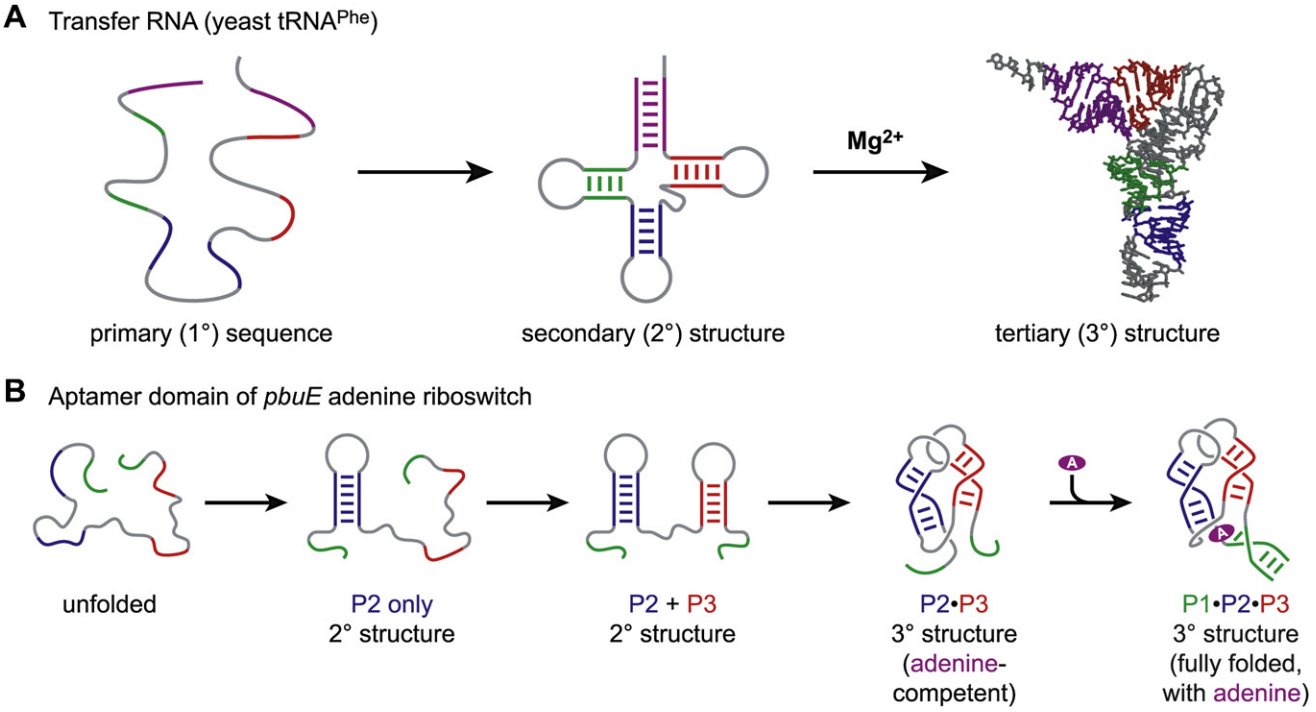

Figure 1. RNA Folding Hierarchy and Folding of the *pbuE* Adenine Riboswitch Aptamer Domain

(A) The standard RNA folding hierarchy, in which secondary structure forms first and tertiary structure then forms from the pre-assembled secondary structure elements. This is illustrated for transfer RNA [tRNA; PDB id 1EHZ (Shi and Moore, 2000)], which generally follows the folding hierarchy (Tinoco and Bustamante, 1999). As their length and structural complexity increase, RNA molecules may no longer strictly obey the hierarchy.

(B) Folding pathway of the *pbuE* adenine riboswitch aptamer domain, as revealed by new single-molecule force measurements (Greenleaf et al., 2008). The final step in the pathway involves formation of secondary structure element P1 although P2•P3 tertiary structure is already present, and therefore this RNA does not strictly follow the hierarchy of panel A. Nonetheless, folding of the aptamer is clearly sequential because a discrete set of folding intermediates is observed.

It is fair to say that these results revealed no qualitative surprises about the folding landscape of the adenine riboswitch aptamer. For example, although the P1 secondary structure element forms (in the folding direction) only after the P2•P3 tertiary contact is formed, thereby violating the simple "secondary structure before tertiary structure" hierarchy (Tinoco and Bustamante, 1999), this outcome is sensible for the riboswitch aptamer because these particular interactions have comparable thermodynamic stabilities. Perhaps in such cases we should refer to "sequential" rather than "hierarchical" RNA folding, because there is a well-defined folding sequence even though the strict hierarchy between RNA secondary and tertiary structure is not maintained. This viewpoint is reinforced by other well-established examples in which formation of RNA secondary and tertiary structure elements are not cleanly separable (Pan and Woodson, 1998; Russell et al., 2006; Wu and Tinoco, 1998).

The three most interesting aspects of the new study are (1) the integration of several recently developed equilibrium and nonequilibrium single-molecule force spectroscopy techniques; (2) the use of the combined data from these techniques to provide a comprehensive and highly precise view of both the thermodynamics and kinetics of sequential RNA folding; and (3) the likely future applications of these approaches to study cotranscriptional RNA folding pathways. Folding during synthesis is certainly important for "real" RNA folding in vivo but is challenging to study in vitro (Heilman-Miller and Woodson, 2003; Pan et al., 1999). Single-molecule force measurements allow monitoring of the close interplay between transcription and the dimensions of the RNA being transcribed. Therefore, it should be particularly interesting to see what analogous force experiments reveal for larger RNAs, in which misfolding via kinetic traps can dominate the folding landscape (Russell et al., 2006).

REFERENCES

Batey, R.T., Gilbert, S.D., and Montange, R.K. (2004). Nature *432*, 411–415.

Cornish, P.V., and Ha, T. (2007). ACS Chem. Biol. *2*, 53–61.

Greenleaf, W.J., Frieda, K.L., Foster, D.A., Woodside, M.T., and Block, S.M. (2008). Science *319*, 630–633.

Heilman-Miller, S.L., and Woodson, S.A. (2003). RNA *9*, 722–733.

Lemay, J.F., Penedo, J.C., Tremblay, R., Lilley, D.M.J., and Lafontaine, D.A. (2006). Chem. Biol. *13*, 857–868.

Mandal, M., and Breaker, R.R. (2004). Nat. Rev. Mol. Cell Biol. *5*, 451–463.

Noeske, J., Buck, J., Fürtig, B., Nasiri, H.R., Schwalbe, H., and Wöhnert, J. (2007). Nucleic Acids Res. *35*, 572–583.

Ottink, O.M., Rampersad, S.M., Tessari, M., Zaman, G.J., Heus, H.A., and Wijmenga, S.S. (2007). RNA *13*, 2202–2212.

Pan, J., and Woodson, S.A. (1998). J. Mol. Biol. *280*, 597–609.

Pan, T., Artsimovitch, I., Fang, X.W., Landick, R., and Sosnick, T.R. (1999). Proc. Natl. Acad. Sci. USA *96*, 9545–9550.

Russell, R., Das, R., Suh, H., Travers, K.J., Laederach, A., Engelhardt, M.A., and Herschlag, D. (2006). J. Mol. Biol. *363*, 531–544.

Serganov, A., Yuan, Y.R., Pikovskaya, O., Polonskaia, A., Malinina, L., Phan, A.T., Höbartner, C., Micura, R., Breaker, R.R., and Patel, D.J. (2004). Chem. Biol. *11*, 1729–1741.

Shi, H., and Moore, P.B. (2000). RNA *6*, 1091–1105.

Tinoco, I., Jr., and Bustamante, C. (1999). J. Mol. Biol. *293*, 271–281.

Tinoco, I., Jr., and Bustamante, C. (2002). Biophys. Chem. *101–102*, 513–533.

Wickiser, J.K., Cheah, M.T., Breaker, R.R., and Crothers, D.M. (2005). Biochemistry *44*, 13404–13414.

Woodside, M.T., Behnke-Parks, W.M., Larizadeh, K., Travers, K., Herschlag, D., and Block, S.M. (2006). Proc. Natl. Acad. Sci. USA *103*, 6190–6195.

Wu, M., and Tinoco, I., Jr. (1998). Proc. Natl. Acad. Sci. USA *95*, 11555–11560.