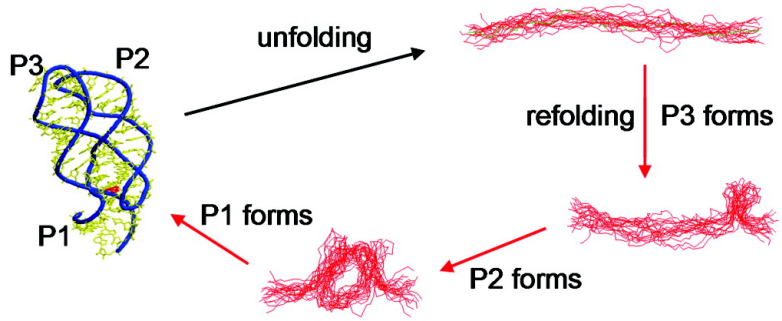
Communication

Relative Stability of Helices Determines the Folding Landscape of Adenine Riboswitch Aptamers

Jong-Chin Lin, and D. Thirumalai

J. Am. Chem. Soc., **2008**, 130 (43), 14080-14081 • DOI: 10.1021/ja8063638 • Publication Date (Web): 02 October 2008



More About This Article

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

- Supporting Information
- 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 COMMUNICATIONS

Published on Web 10/02/2008

Relative Stability of Helices Determines the Folding Landscape of Adenine Riboswitch Aptamers

Jong-Chin Lin and D. Thirumalai*

Department of Chemistry and Biochemistry and Biophysics Program, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742

Received August 11, 2008; E-mail: thirum@umd.edu

Riboswitches, found in the untranslated region (UTR) of mRNAs of both prokaryotes and eukaryotes,[1] are RNA elements that regulate gene expression by sensing and binding target cellular metabolites. They contain a conserved metabolite-binding aptamer domain and a downstream expression platform. In bacteria, ligand binding to the aptamer domain usually results in a conformational change,[2] which alters the folding pattern of the expression platform and controls transcription termination or translation initiation.[3] Among the simplest riboswitches, the purine (guanine and adenine) riboswitches[4] display remarkable ligand selectivity and carry out markedly different functions despite the structural similarity of their aptamers.[5,6] For the *pbuE* adenine (A) riboswitch, the ligand binding activates the gene expression when an antiterminator is formed.[5] In the absence of adenine, part of the aptamer region is involved in the formation of a terminator stem with the expression platform, which results in transcription termination. The *add* A-riboswitch, on the other hand, activates the gene expression by forming a translational activator upon ligand binding, while, in the absence of adenine, the riboswitch adopts the structure with a translational repressor stem in the downstream region.[3] Thus, it is important to quantitatively map the folding landscape of aptamers to understand the differences in the function of structurally similar riboswitches.

Purine riboswitch aptamers[3] form a three-way junction formed from helix P1 and hairpins P2 and P3, that are stabilized by tertiary interactions in the folded state (Figure 1 and the contact map in Figure S1 in the Supporting Information, SI). A recent single molecule optical tweezer experiment has been used to directly observe the hierarchical folding in the *pbuE* A-riboswitch aptamer.[7] Here, we describe force(f)-triggered unfolding and refolding of the A-riboswitch aptamer theoretically using the self-organized polymer model[8] with the Langevin dynamics in the overdamped limit (see SI for details). The native structure, taken from the crystal structure of the aptamer domain of the *Vibrio vulnificus add* A-riboswitch,[3] has 63 nucleotides (from U17 to A79, PDB code: 1Y26; see Figure 1).

In the absence of adenine the force−extension curve, at a constant loading rate of 960 pN/s, shows that unfolding occurs in three distinct steps (Figure 1c). When $f \approx 11$ pN the aptamer contour length increases by ∼9 nm. This step is associated with the unfolding of helix P1 and the triple-helix junction (see Figure S2). In the second step, at $f \approx 12$ pN, the contour length increases from 10 to 16 nm (0.35 nm/nt) and is associated with the unfolding of the 21-nt P2. At $f \approx 13$ pN, the contour length increases by ∼5 nm, which is associated with the unfolding of the 19-nt P3. By investigating the loss of contacts during the unfolding process, we find that when hairpins P2 and P3 are both folded, the tertiary loop−loop interaction stabilizes P2 and P3 hairpin loops (see Figure S2). However, the two loops unbind transiently due to thermal fluctuations, which is consistent with the observation that P2−P3 tertiary interactions require adenine.[5,9]

The presence of adenine in the binding pocket in the triple-helix junction increases the unfolding force (Figure 1d). Surprisingly, the

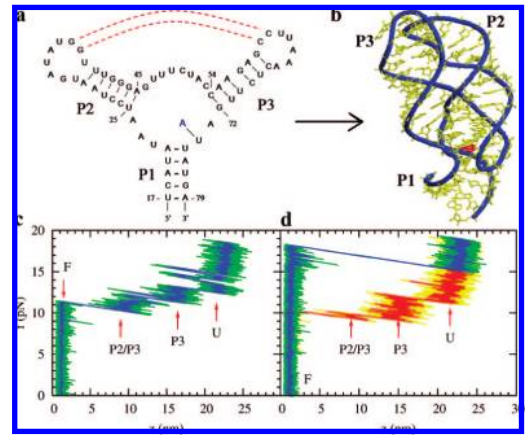


Figure 1. (a) Secondary and (b) tertiary structures of the *add* A-riboswitch aptamer. Dashed lines in red indicate loop−loop interaction in the folded aptamer. Typical force−extension curve for unfolding the A-riboswitch aptamer (c) without and (d) with adenine at constant loading rate. The curve for refolding (red) in the presence of adenine is shown in (d). The unfolding sequence (F → P2/P3 (P2 and P3 are folded) → P3 (folded) → U) is the reverse of refolding reflecting the hierarchical assembly of the riboswitch. The last step in the folding is the metabolite-induced formation of P1.

binding of adenine results in the absence of intermediate unfolding steps; the aptamer stays folded until f is increased to 18 pN, and then it unfolds completely (Figure 1d). The unfolding force is comparable to that found in experiments[7] for the pbuE adenine riboswitch aptamer, which suggests that metabolite-bound stabilities of the purine riboswitches are similar.

After the aptamer is globally unfolded, we reduce the force, with the same loading rate, to initiate the refolding of the riboswitch. An identical refolding pathway is found by quenching the force from an initial high force to constant low forces. The forces at which P3, P2, and P1 form are larger than the experimental values for the *pbuE* aptamer[7] by only ∼40% (Figures S3 and S4). In both methods of force reduction, several intermediate states are resolved during the hierarchical refolding process; P3 folds first followed by P2, and finally the triple-helix junction and the helix P1 form (Figure 1d). The finding that P3 folds first implies that it is more stable than P2, which is consistent with the smaller free energy predicted for the secondary structure of P3 (smaller than P2 by 1 kcal/mol) in the *add* riboswitch aptamer using the Vienna RNA package.[10] Remarkably, despite the structural similarity between *pbuE* and *add* A-riboswitch aptamers, experiments[7] show that P2 in *pbuE* unfolds last and presumably is the first structural element to refold. Our results would imply that, in the *pbuE* riboswitch aptamer, P2 ought to be more stable than P3. Indeed, the relative stability of P2/P3 is different in the *pbuE* A-riboswitch aptamer where the predicted free energy for the secondary structure of P2 is smaller than that of P3 by 2 kcal/mol. The stability difference explains the reversed order of the folding of P2 and P3 in

10.1021/ja8063638 CCC: $40.75

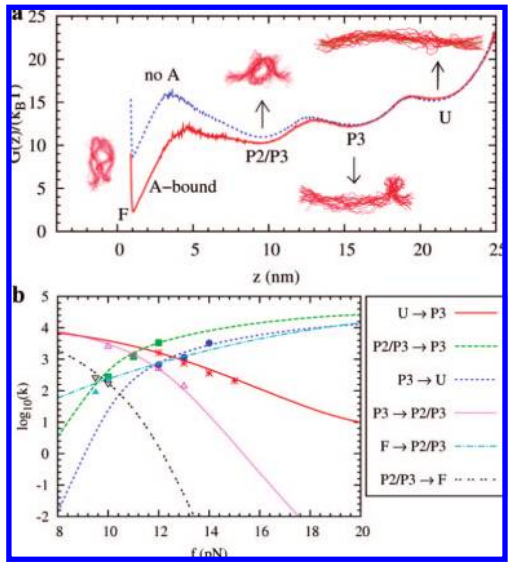


Figure 2. (a) Folding landscape for A-riboswitch aptamer as a function of z at $f = 10$ pN without (dashed line) and with adenine (solid line). The ensemble of structures in the intermediate basins are shown. (b) The logarithm of transition rates, $\log(k = 1/\bar{\tau})$, between the distinct states calculated using eq 1 (shown in lines), and directly from the time traces of the extension of the aptamer in simulations (shown in points). The variances of $\log(\tau)$ from the time traces are within ± 0.75. The excellent agreement between theory and simulations shows that z is an excellent reaction coordinate under tension even if the folding landscape has multiple basins.

the *pbuE* A-riboswitch aptamer.[7] Comparison of the responses to f between the purine riboswitch aptamers shows that the folding landscape is determined by the local stability of the structural elements. The similarity in the distributions of the number of contacts between P2 and P3 hairpin loops over time with and without bound adenine (data not shown) shows that adenine binding does not affect the interactions between P2 and P3 hairpin loops. Rather, the adenine binding stabilizes the triple-helix junction and makes it more difficult to unfold P1. From these findings, we surmise that the rate limiting step in the fully folded aptamer is the formation of P1.

The multiple steps in the hierarchical folding, characteristic of RNA assembly, is also reflected in the free energy profiles (Figures 2a and S4), which are obtained from the distribution of the extension (z) of the aptamer using $G(z) = -k_BT \ln P(z)$, where $P(z)$ is the probability of z. The distance from the folded state ($z = 1$ nm) to the first barrier in the absence of adenine is ~2.5 nm (Figure 2a), which indicates unzipping of 3 base pairs. The location of the transition state (TS) is in excellent agreement with experiments.[6] In the presence of adenine, we predict the position of the first barrier shifts to ~4 nm, which means that 5 base pairs of P1 next to the nucleotide that has direct interaction with adenine are ruptured at the TS. Thus adenine binding becomes the key barrier in the first unfolding step. We also find that, at $f = 10$ pN, binding of adenine stabilizes the folded state by $\sim 6k_BT$ and increases the energy barrier for leaving the folded state by $2k_BT$. These results support the hypothesis that the ligand binding stabilizes P1, a feature that is common to most riboswitches.[2]

The utility of the free energy profiles, $G(z)$, is that they can be used to obtain force-dependent transition rates.[11] We calculated the transition rates between the four distinct states (F, P2/P3, P3, and U) in the complex folding landscape of the riboswitch aptamer (Figure 3a). The rate of transition from an initial point x_A in one state to the final point x_B in another is given by

$$k^{-1}(f) = \bar{\tau}(f) \approx \int_{X_A}^{X_B} \mathrm{d}y\, e^{\beta G(f,y)} \frac{1}{D_{A\leftrightarrow B}} \int_{X_0}^{y} \mathrm{d}z\, e^{-\beta G(f,z)} \qquad (1)$$

where $D_{A\leftrightarrow B}$ is the diffusion coefficient along the extension coordinate for transition between states A and B, $\beta = 1/(k_BT)$, $x_0 = -\infty$ if $x_A < x_B$, and $x_0 = \infty$ if $x_A > x_B$. Here, the diffusion coefficient is obtained by equating the transition rate calculated using eq 1 to that obtained from time traces (Figure S3) at $f = 12$ pN for U→P3 and P3→P2/P3 transitions and at $f = 10$ pN for the F→P2/P3 transition (see SI for details).

The simulated transition rates without adenine are in excellent agreement with predictions based upon eq 1 (Figure 2b). When $f = 11-13$ pN, the aptamer switches between the unfolded, P3 folded, and P2/P3 folded states with rates on the order of 10^3 s^{-1} with the energy barriers $\sim(2-3)k_BT$. Transitions between the folded state and P2/P3 folded (or P1 unfolded) state have to cross higher barriers $(5-7k_BT)$ with rates ~100 s^{-1} at $f = 10$ pN. The calculated transition rate for the F→P2/P3 transition, at zero force, using $G(0,z) = G(f,z) + fz$ in eq 1, is ~0.07 s^{-1}, which compares favorably with single molecule experiments.[7]

With the binding of adenine, the unfolding rate decreases by 2 orders of magnitude due to the deeper well of the folded state and larger energy barrier for transition from the folded state. The slow unfolding rate is in agreement with that observed for FMN riboswitches,[12] suggesting that functions of riboswitches are kinetically, rather than thermodynamically, controlled. The result is crucial for transcription of the complete riboswitch. In vivo, without metabolite binding, the riboswitch favors the formation of the downstream terminator hairpin, which disrupts the aptamer structure. Ligand binding thus stabilizes the aptamer structure during transcription and prevents the formation of the terminator stem before transcription is completed (*pbuE* riboswitches) or the formation of translation repressor stem before translation is initiated (*add* riboswitches). Further study of cotranscriptional folding of the complete riboswitch including the downstream expression platform is necessary to fully understand the mechanism of gene regulation by riboswitches.

Acknowledgment. We thank R. T. Batey and N. Toan for critical comments on the manuscript. This work is supported by grants from the National Science Foundation (CHE-05-14056) and AFOSR (FA9550-07-1-0098).

Supporting Information Available: Theoretical methods and figures of number of contacts, time traces of extension, free energy profiles for the aptamer at constant forces. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Winkler, W. C.; Breaker, R. R. *Annue. Rev. Microbiol.* **2005**, *59*, 487–517. (b) Cheah, M. T.; Wachter, A.; Sudarsan, N.; Breaker, R. R. *Nature* **2007**, *447*, 497–500. (c) Breaker, R. R. *Science* **2008**, *319*, 1795–1797.
(2) (a) Montange, R. K.; Batey, R. T. *Ann. Rev. Biophys.* **2008**, *37*, 117–133. (b) Batey, R. T.; Gilbert, S. D.; Montange, R. K. *Nature* **2004**, *432*, 411–415. (c) Gilbert, S. D.; Rambo, R. P.; Van Tyne, D.; Batey, R. T. *Nat. Struc. Mol. Biol.* **2008**, *15*, 177–182. (d) Montange, R. K.; Batey, R. T. *Nature* **2006**, *441*, 1172–1175.
(3) 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*, 1729–1741.
(4) Edwards, T. E.; Klein, D. J.; Ferre-D'Amare, A. R. *Curr. Opin. Struct. Biol.* **2007**, *17*, 273–279.
(5) Mandal, M.; Boese, B.; Barrick, J. E.; Winkler, W. C.; Breaker, R. R. *Cell* **2003**, *113*, 577–586.
(6) Mandal, M.; Breaker, R. R. *Nat. Struct. Mol. Biol.* **2003**, *11*, 29–35.
(7) Greenleaf, W. J.; Frieda, K. L.; Foster, D. A.; Woodside, M. T.; Block, S. M. *Science* **2008**, *319*, 630–633.
(8) (a) Hyeon, C.; Thirumalai, D. *Biophys. J.* **2007**, *192*, 731–743. (b) Hyeon, C; Dima, R. I.; Thirumalai, D. *Structure* **2006**, *14*, 1633–1645.
(9) Lemay, J. F.; Penedo, J. C.; Tremblay, R.; Lilley, D. M. J.; Lafontaine, D. A. *Chem. Biol.* **2006**, *13*, 857–868.
(10) Hofacker, I. L. *Nucleic Acids Res.* **2003**, *31*, 3429–3431.
(11) Hyeon, C.; Morrison, G.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9604–9609.
(12) Wickiser, J. K.; Winkler, W. C.; Breaker, R. R.; Crothers, D. M. *Mol. Cell* **2005**, *18*, 49–60.

JA8063638