Nonlocal Helix Formation Is Key to Understanding S-Adenosylmethionine-1 Riboswitch Function

Paul C. Whitford,[†] Alexander Schug,[†] John Saunders,[†] Scott P. Hennelly,[‡] José N. Onuchic,[†] and Kevin Y. Sanbonmatsu[‡*]

[†]Center for Theoretical Biological Physics and Department of Physics, University of California at San Diego, La Jolla, California; and [‡]Theoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico

ABSTRACT Riboswitches are noncoding RNAs that regulate gene expression in response to changing concentrations of specific metabolites. Switching activity is affected by the interplay between the aptamer domain and expression platform of the riboswitch. The aptamer domain binds the metabolite, locking the riboswitch in a ligand-bound conformation. In absence of the metabolite, the expression platform forms an alternative secondary structure by sequestering the 3′ end of a nonlocal helix called P1. We use all-atom structure-based simulations to characterize the folding, unfolding, and metabolite binding of the aptamer domain of the S-adenosylmethionine-1 (SAM-1) riboswitch. Our results suggest that folding of the nonlocal helix (P1) is rate-limiting in aptamer domain formation. Interestingly, SAM assists folding of the P1 helix by reducing the associated free energy barrier. Because the 3′ end of the P1 helix is sequestered by an alternative helix in the absence of metabolites, this observed ligand-control of P1 formation provides a mechanistic explanation of expression platform regulation.

Received for publication 30 August 2008 and in final form 8 October 2008.

*Correspondence: kys@lanl.gov

Structure formation in mRNA often regulates genetic expression. Multiple compact conformations may be accessed while kinetic and thermodynamic competition of these structures determines the functional state of the mRNA (1). In these systems the folding dynamics can play a critical role in biological function. Riboswitches are one class of functional mRNA units that are often found in specific 5′-untranslated regions of mRNA (2). They regulate transcription and translation in response to changing concentrations of metabolites via communication between an aptamer (metabolite binding) domain and the expression platform (Fig. 1 *a*). Conformational changes in the aptamer domain are essential for this functional response. Little is known about riboswitch function from a theoretical perspective. Significant computational efforts have focused on RNA tetraloops (3). One question of interest is: How does ligand binding influence the formation of secondary and tertiary structure? Recent single molecule force spectroscopy experiments (4) have suggested the helix formed by the 3′ and 5′ ends of a *pbuE* adenine riboswitch is the least thermodynamically stable helix and is the helix most sensitive to metabolite concentrations. In contrast, fluorescence experiments suggest native 5′-3′ helix formation occurs before metabolite binding in a *thiM* riboswitch (5).

In this letter we describe the role of the 5′-3′ helix (P1) folding and S-Adenosylmethionine (SAM) binding in the activity of the SAM-I riboswitch (6) (Fig. 1). We adopt the energy landscape theory of protein folding (7) and apply it to RNA via an all-atom structure-based model (our ideas are based on the model presented in (8); structure-based models have been used for RNA folding in (9–11). See also the Supporting Material). We compare aptamer domain folding with and without its associated metabolite, SAM. The functional state of the riboswitch is regulated by the balance of aptamer domain folding and formation of an alternate conformation involving a terminator sequence binding the 3′ tail of the riboswitch (12). It has been suggested that breaking of the 3′ tail (in the nonlocal helix) is needed to regulate the expression platform. Although the terminator sequence has been identified, the structure of the full riboswitch has not been solved and the precise details of the decision process need to be determined. However, the folding of both conformations must occur on the same energy landscape. Thus, rate-limiting steps in aptamer formation may provide opportunities for the alternate structure to form and the functional decision to be made. We perform simulations using the recently solved x-ray structure of the SAM-1 riboswitch aptamer domain (6), allowing us to isolate the role of P1 formation in aptamer folding. Our results suggest the rate-limiting step in aptamer domain folding is the initiation of P1 helix formation. SAM reduces the associated free-energy barrier by binding to the preformed P3 helix and then attracting the unstructured strands of the P1 helix.

Energy landscape theory states that nature has selected for protein sequences that maximize the energetic bias for the native state and minimize trapping of nonnative structures. Namely, they have been selected to be minimally frustrated. The principle of minimal frustration has been validated through comparison of structure-based models and experimental results, which has led to the funnel paradigm of protein folding (7). For structured RNA, one can envision a frustrated landscape where there is a marginal bias to reach the native state.

Editor: Kathleen B. Hall.

doi: 10.1016/j.bpj.2008.10.033

The RNA would then randomly search all possible basepairs and the folded state would only be reached by chance. This would result in a ''Levinthal's paradox'', where searching takes the age of the universe, whereas, in reality, folding of functional RNAs can be fast (approximately milliseconds). Therefore, evolutionary pressure to reduce frustration must exist. Although RNA is likely frustrated to some degree, by understanding energetically unfrustrated models one can partition the structural and energetic effects in folding and function.

The principle of minimal frustration is applied via structure-based simulations in which all heavy atoms are explicitly represented. The model is energetically unfrustrated since only native interactions are attractive and all other interactions are repulsive. Kinetic (temperature jump) and thermodynamic (constant temperature) simulations of the aptamer domain were performed, both with and without SAM present. Thermodynamic simulations ranged in temperature such that the full folding/binding landscape could be characterized (Fig. 1, *c–f*, and Fig. 2). For SAM-present simulations, one copy of the aptamer domain and 100 copies of the SAM molecule are placed in a box with periodic boundary conditions. SAM molecules are free to associate and only native SAM-aptamer interactions are attractive. Since SAM-SAM interactions are strictly repulsive, metabolite aggregation and nonspecific binding are not possible. To our knowledge, this is the first simulation in which a bath of ligands (with atomic resolution) is able to freely associate and dissociate with a RNA molecule during folding.

In thermodynamic simulations of the apo aptamer domain, the largest free-energy barrier is associated with initial formation of the P1 helix (Fig. 1, *c* and *e*, *black arrow*). In the presence of SAM, the initiation of P1 helix formation and the free-energy barrier are encountered earlier in the folding process (Fig. 1, *d* and *f*, *black arrows*) and the free-energy barrier is reduced. P1 forms after all other secondary structure (and some tertiary structure) is formed and SAM primarily affects P1 folding in both thermodynamic and kinetic simulations (see Fig. S3 in the Supporting Material). In the SAM riboswitch, the SAM molecule stabilizes the rate-limiting step (largest free energy barrier; see the Supporting Material) in folding, which leads to a kinetically accessible and thermodynamically more stable folded aptamer domain.

Since the P3 domain is formed before SAM binding (Fig. 1 *c*, *green curve*), P3 can serve as a platform for SAM binding. Fig. 2 shows that upon binding to P3, SAM stabilizes the P1 domain by predominantly interacting with the 3′ strand and then the 5′ strand of P1 (see Movie S1 in the Supporting Material).

Another notable feature in Fig. 1, *c* and *d*, is the apparent interplay between P1 and the pseudo-knot (PK, *starred*). In kinetic simulations (see the Supporting Material) this partial unfolding of the PK is more pronounced, suggesting that a dynamic balance between PK and P1 formation exists.

The current picture of RNA folding is hierarchical (13). In this view, it is important to distinguish between local helices (formed by simple stem-loops) and nonlocal helices (formed by two strands distant in sequence) (14). Relative to a stem loop, a nonlocal helix has a larger loss of entropy associated with its formation. This unfavorable driving force is often accounted for in secondary structure prediction algorithms, where scoring penalties are imposed on large loops (15). Thus, it may not be surprising to find a nonlocal helix (P1) that is less stable than the local helices. As we have shown, the entropic barrier due to bringing together distant (in sequence) bases also gives rise to the rate-limiting step: initiation of P1 folding.

Since P1 folding is rate-limiting, it is an ideal stage for SAM to bind and the on/off decision to be made. Our results provide a detailed mechanism for both this switching decision and SAM binding. Our results also suggest the structural mechanism of control is the same, regardless of whether the process is thermodynamically, or kinetically, regulated (16).

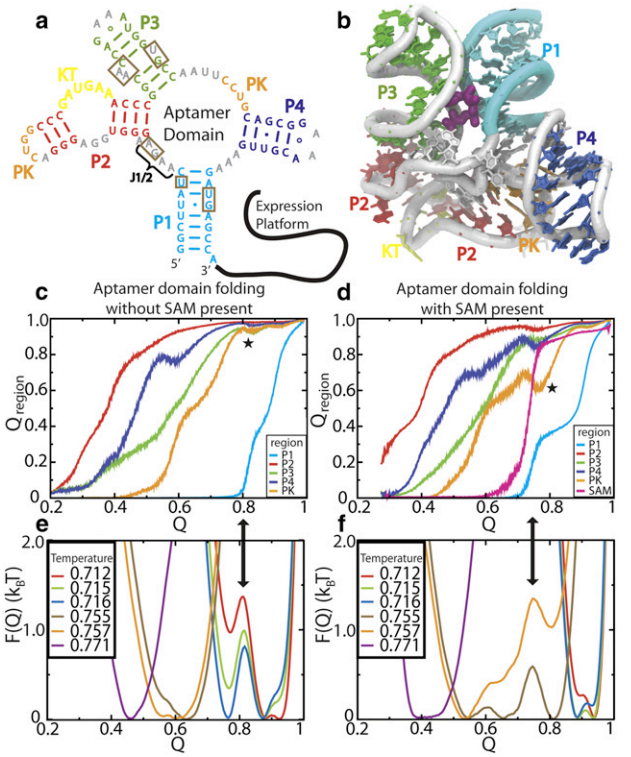

FIGURE 1 (***a***) **Secondary and (*b*) tertiary structure (PDB entry: 2GIS) of the SAM-I riboswitch. Average secondary structure formation as a function of the fraction of native contacts formed (*Q*; see the Supporting Material) for the (*c*) SAM-free and (*d*) SAM-present simulations. (*a–d*) Color scheme: P1, cyan; P2, red; P3, green; P4, blue; PK, orange; and SAM, purple in panels *b* and *d*. In panel *a*, SAM-contacting residues are highlighted by brown boxes. The most notable difference in folding mechanism is earlier initial folding of P1 (*black arrows*) at the expense of the PK (*starred*) when SAM is present. The folding free-energy profiles for the (*e*) SAM-free and (*f*) SAM-present simulations are shown for several temperatures (with temperature indicated by *color*). The most significant free-energy barrier in both systems is associated with initial P1 folding. When SAM is present, the free-energy barrier is reduced and encountered earlier in the folding process.**

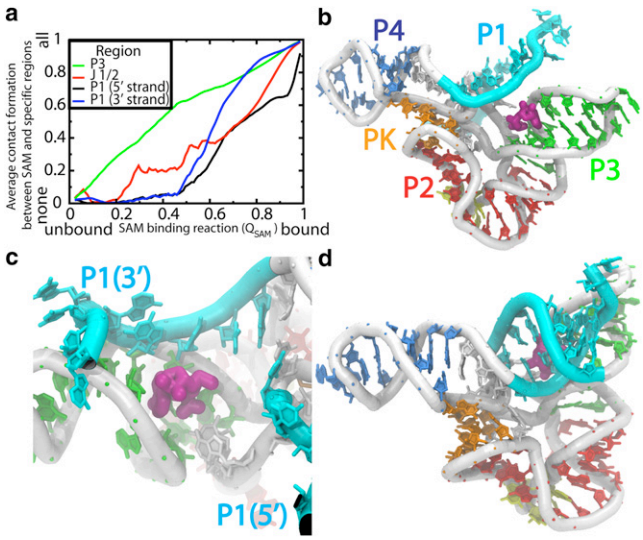

FIGURE 2 **(*a*) Average percent of SAM-aptamer domain interactions formed by region as a function of the fraction of native SAM-aptamer domain contacts formed Q_{SAM}. Simulation images illustrating the SAM binding mechanism: (*b*) SAM binds a preformed P3 helix; (*c*) SAM recruits 3′ strand of P1; (*d*) SAM binds 5′ strand of P1; and P1 helix formation proceeds.**

The less stable, and slower forming, P1 helix results in a compact state where some tertiary structure (the PK) can be formed. In this partially structured state SAM may bind to a preformed P3 helix. After SAM binds to P3, it localizes the 3′ and 5′ strands that compose the P1 helix. SAM binding to P1 initiates P1 helix formation (Fig. 1 *d*), after which P1 continues to form without any significant free-energy barriers.

Since the P1 helix is a fragile structure (relative to P2, P3, and P4), it is likely more sensitive to the cellular environment. Force spectroscopy experiments have shown a coupling between nonlocal helix formation and ligand binding in an adenine riboswitch (4). In *Azoarcus* ribozyme (17), a near-native, compact state with partial tertiary structure has been experimentally observed. This is also consistent with nonlocal helix formation being the final folding step. Although nonlocal helix formation is important in some RNA-ligand systems, loop ordering (18,19) and tertiary structure formation (5) may also be important in the decision processes of other riboswitches.

Several recent results have shown that molecular recognition, control, and signaling do not necessarily occur by surface matching between biomolecules. Rather, a more interesting process occurs where folding of the biomolecular parts is signaled through binding. Our results suggest that initial P1 formation is a central step for further recognition and function in the SAM aptamer.

SUPPORTING MATERIAL

Figure, movie and caption, and data files are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(08)00075-1.

ACKNOWLEDGMENTS

P.C.W. and A.S. thank Ulrich Müller at the University of California at San Diego for useful discussions. We also thank the Tera-Grid Computing Facilities at Virginia Tech and Los Alamos National Laboratory Institutional Computing.

This work was supported in part by grants No. PHY-0216576 and 0225630 from the National Science Foundation-sponsored Center for Theoretical Biological Physics, National Science Foundation grant No. PHY-0822283, the Los Alamos National Laboratory LDRD program, and National Institutes of Health grant No. R01-GM072686.

REFERENCES and FOOTNOTES

1. Breaker, R. 2008. Complex riboswitches. *Science*. 319:1795–1797.
2. Montange, R. K., and R. T. Batey. 2008. Riboswitches: emerging themes in RNA structure and function. *Annu. Rev. Biophys.* 37:117–133.
3. Garcia, A. E., and D. Paschek. 2008. Simulation of the pressure and temperature folding/unfolding equilibrium of a small RNA hairpin. *J. Am. Chem. Soc.* 130:815–817.
4. Greenleaf, W. J., K. L. Frieda, D. A. N. Foster, M. T. Woodside, and S. M. Block. 2008. Direct observation of hierarchical folding in single riboswitch aptamers. *Science*. 319:630–633.
5. Lang, K., R. Rieder, and R. Micura. 2007. Ligand-induced folding of the thiM TPP riboswitch investigated by a structure-based fluorescence spectroscopic approach. *Nucleic Acids Res.* 35:5370–5378.
6. Montange, R. K., and R. T. Batey. 2006. Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature*. 441:1172–1175.
7. Onuchic, J. N., Z. Luthey-Schulten, and P. G. Wolynes. 1997. Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* 48:545–600.
8. Whitford, P. C., J. K. Noel, S. Gosavi, A. Schug, K. Y. Sanbonmatsu, et al. 2008. An all-atom structure-based potential for proteins: bridging minimal models with all-atom empirical force fields. *Prot. Struct. Funct. Bioinform. 10.1002/port.22253*.
9. Sorin, E. J., B. J. Nakatani, Y. M. Rhee, G. Jayachandran, V. Vishal, et al. 2004. Does native state topology determine the RNA folding mechanism? *J. Mol. Biol.* 337:789–797.
10. Hyeon, C., and D. Thirumalai. 2007. Mechanical unfolding of RNA: from hairpins to structures with internal multiloops. *Biophys. J.* 92:731–743.
11. Hyeon, C., R. I. Dima, and D. Thirumalai. 2006. Pathways and kinetic barriers in mechanical unfolding and refolding of RNA and proteins. *Structure*. 14:1633–1645.
12. Winkler, W. C., A. Nahvi, N. Sudarsan, J. E. Barrick, and R. R. Breaker. 2003. An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat. Struct. Biol.* 10:701–707.
13. Brion, P., and E. Westhof. 1997. Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* 26:113–117.
14. Woodson, S. A. 2002. Folding mechanisms of group I ribozymes: role of stability and contact order. *Biochem. Soc. Trans.* 30:1166–1169.
15. Chen, S. 2008. RNA folding: conformational statistics, folding kinetics, and ion electrostatics. *Annu. Rev. Biophys.* 37:197–214.
16. Coppins, R. L., K. B. Hall, and E. A. Groisman. 2007. The intricate world of riboswitches. *Curr. Opin. Struct. Biol.* 10:176–181.
17. Chauhan, S., and S. A. Woodson. 2008. Tertiary interactions determine the accuracy of RNA folding. *J. Am. Chem. Soc.* 130:1296–1303.
18. Gilbert, S. D., C. D. Stoddard, S. J. Wise, and R. T. Batey. 2006. Thermodynamic and kinetic characterization of ligand binding to the purine riboswitch aptamer domain. *J. Mol. Biol.* 359:754–768.
19. Ottink, M. O., S. M. Rampersad, M. Tessari, G. J. R. Zaman, H. A. Heus, et al.. 2007. Ligand-induced folding of the guanine-sensing riboswitch is controlled by a combined predetermined-induced fit mechanism. *RNA*. 13:2202–2212.