Published in final edited form as:

J Mol Biol. 2015 November 6; 427(22): 3469–3472. doi:10.1016/j.jmb.2015.08.001.

Gene Regulation Gets In Tune: how riboswitch tertiary-structure networks adapt to meet the needs of their transcription units

Debapratim Dutta and Joseph E. Wedekind

Department of Biochemistry & Biophysics, and Center for RNA Biology, University of Rochester School of Medicine and Dentistry, Rochester NY 14642, USA

Regulatory RNA elements called riboswitches are derived typically from the 5'-leader sequences of bacterial messenger RNAs where they control expression of the associated, downstream gene by direct binding to a small molecule effector, such as S-adenosyl-L-methionine (SAM) [1]. Riboswitches are notable not only for their lack of an accessory protein requirement, but also because they are targets of antibacterial compounds [2-4] and hold promise as drug targets [5, 6]. In terms of organization, riboswitches comprise a conserved aptamer (receptor) domain that binds the cognate effector (ligand), and a downstream expression platform harboring transcription- or translation-related sequences. Ligand binding stabilizes a distinct three-dimensional conformation of the riboswitch, masking or unfurling expression platform signals that result in gene "on" or "off" structural states. In this manner, riboswitches serve as economical sensors that provide rapid feedback for the constantly changing cellular environment. To date, >35 riboswitch classes have been characterized [7].

Remarkably, at least six riboswitch classes sense SAM (i.e., the SAM-I, II, III, IV, I/IV, and SAM/SAH riboswitches) [8]. SAM clan members I, I/IV, and IV recognize this methyl-donor coenzyme by an almost invariant set of bases organized around a four-way-helical junction core flanked by coaxially stacked helices [9-13]. In *Bacillus subtilis*, regulatory RNAs control >4.1% of genes [14], including 11 SAM-I riboswitches that govern discrete transcriptional units involved in methionine metabolism or transport [15]. These 'S-boxes' promote read-through of a transcriptional terminator under methionine starvation [16]. Despite conservation of SAM-binding sequences, S-boxes exhibit a range of ligand-dependent transcription termination efficiencies *in vitro*, ranging from 6% (no SAM) to 92% (3 μ M SAM) with half-maximal termination between 0.35 to 40 μ M SAM [16]. Importantly, the latter values do not correlate strictly with equilibrium $K_{D,app}$ values, which range from 14 nM to 3.5 μ M [16]. Intriguingly, differences in S-box regulatory prowess appear adapted to cellular requirements for a particular transcriptional unit. S-boxes that control costly methionine biosynthesis (e.g., *metE* and *yitJ*) show lowest expression of their associated genes during cell growth in methionine; conversely, methionine starvation produces the

Correspondence to: Joseph E. Wedekind; joseph.wedekind@rochester.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

highest expression, in accordance with the minute concentrations of SAM required to elicit half-maximal S-box-mediated transcription-termination. By contrast, the *yusC* S-box controls methionine transport, and allows higher expression of its associated genes in the presence of methionine; appropriately, the *yusC* dynamic range is relatively narrow, consistent with the observation that import is a more frugal way to obtain methionine than *de novo* synthesis [16].

So, *how do* riboswitches with a common ligand-binding core become tuned to the regulatory requirements of individual transcription units? In this issue, Batey and co-workers address this challenge by investigating the *B. subtilis metE* S-box riboswitch [17]. Their experimental design considers that RNA folding is dictated by kinetics rather than thermodynamics. To illustrate, high SAM concentrations promote integration of *metE* anti-terminator sequences into the newly transcribed aptamer, predisposing downstream sequences to fold into a terminator stem preceding a patch of uridines that pause the polymerase, leading to DNA template release [18, 19]. In low SAM, an alternative mRNA conformation arises early at the expense of aptamer folding. The resulting anti-terminator suppresses formation of the downstream terminator, resulting in a full-length transcript for protein translation [1].

To capture co-transcriptional folding of *metE* at various SAM concentrations, the authors added heparin to the preformed transcription-elongation complex, thereby trapping the polymerase upon dissociation from the template after the initial round of transcription [20]. The key experimental parameters obtained were: (i) T_{50} , the SAM concentration required to elicit half-maximal termination; and (ii) dynamic range, the difference in response level derived from transcriptional “on” and “off” states based on a reporter transcript. The approach was complemented with $K_{D,app}$ measurements from calorimetry, and structure-guided interpretation. Although similar single-turnover transcription analysis has been used [21, 22], the large number of S-box riboswitches and functional diversity of their associated transcriptional units enabled the authors to conduct a broader analysis, elevating the impact of their findings.

Borrowing a page from the “ribozyme folding” playbook, Batey and co-workers sought evidence for tertiary networks, with the hypothesis that such interactions tune individual S-box regulation. Prior work using Mg(II)-induced folding established that complex RNA transcripts can exhibit stable, prolonged intermediates devoid of activity, as illustrated by the relatively slow timescale of *in vitro* splicing by the group I intron [23, 24]. To quell nonnative structures, this ribozyme uses coupled, cooperative long-range interactions early in folding that help navigate the jagged folding-energy landscape on the path to a native conformation [25]. A telltale indication of such networks is that only the most harmful changes to tertiary structures compromise folding, and hence activity [17, 25, 26].

Using structure-guided analysis and prior knowledge, the authors examined four tertiary-structure modules. First, they hypothesized that the internal loop (IL)3a-to-IL3b interaction tunes riboswitch function. A prior structure of the *yitJ* riboswitch reveals A74 of IL3b stacked on A47 of IL3a, which flanks the base triple that recognizes the adenosyl moiety of SAM (**Fig. 1a**). Accordingly, *metE* transcription termination was impaired by IL3b deletion,

or by conversion to an A•C-wobble by the 153 mutation. By contrast, reduction of the extended sequence flanking IL3b to a GNRA tetraloop had modest changes on T_{50} and dynamic range. Retention of a single bulged base in IL3b was sufficient to maintain T_{50} and dynamic range, in addition to $K_{D,app}$, consistent with structural and sequence observations supporting the requirement of a cross-strand IL3a-to-IL3b interaction for SAM binding.

The k-turn was scrutinized next, which resides distant from the SAM-binding pocket in subdomain PK1 (pseudoknot1). K-turns produce sharp bends in the backbone arising from three asymmetrically bulged nucleotides flanked by tandem G•A mismatches (**Fig. 1b**) [27]. The benchmark for SAM-I k-turn impairment is A20C in the context of the *yitJ* riboswitch, which decreased $K_{D,app}$ for SAM ~9-fold [28]. The comparable A57C mutation in *metE* eliminated switching, reflected in both T_{50} and the dynamic range. By contrast, the capacity to form alternative tertiary interactions could explain why the nearby A73C variant led to only modest changes in dynamic range and T_{50} . Similarly, change of the closing G58•U68 pair to a stronger G-C pair had no effect, nor did changes in the composition of the bulge, consistent with its role as a structural spacer (**Fig. 1b**).

K-turn integrity is important because it positions nearby pseudoknot PK-1, whose L2 loop caps stem P2, and forms a long-range contact with J3/4 (**Fig. 1c**) in the four-helix junction. This tertiary interaction accounts for why some mutations that disrupt isolated k-turns are suppressed in the context of the SAM-I structure [26]. Notably, disruption of the tandem Watson-Crick pairs central to the *yitJ* pseudoknot blunts *in vitro* SAM binding, and *in vivo* switching [28, 29]. Batey and co-workers demonstrated an analogous double mutant, G64C/G65C, is similarly disruptive to *metE*. However, because PK-1 of *metE* consists of three Watson-Crick pairs and a closing U62•U168 pair, the authors decided to replace the wobble interaction with A-U and G-C pairs, but saw no effect. Instead, disrupting individual Watson-Crick pairs revealed that PK-1 must comprise at least two canonical pairs inclusive of G65-C165, whose disruption abolished switching. Flanking the L2 loop, two G-C pairs of P2b form minor-groove triplexes with conserved J3/4 adenines that strengthen PK-1 (**Fig. 1c**). Individual A-to-C mutations raised T_{50} , with little effect on dynamic range, suggesting an important role in stabilizing the SAM-bound conformation.

The authors next hypothesized that P2a length (i.e., the helical foundation of L2) might be the cornerstone of PK-1 formation, and hence S-box tuning. Comparative sequence analysis revealed the nine base pair P2a stem of *metE* as atypical, and a preferred length of eight [17]. Deletion of one *metE* P2a base pair had little effect, whereas stems of six, seven or ten were intolerable, suggesting only a narrow window of helix lengths is permissible.

In a final set of experiments, the authors explored why IL3b deletion (e.g. 153) in *metE* compromised SAM binding and switching, despite its poor conservation and prior results suggesting IL3b is dispensable for activity in the context of the *Tte metF* riboswitch [10, 30, 31]. Using their established experimental conditions, the authors validated that a deletion within *metF*, comparable to *metE* 153, retains significant SAM binding [17]. The authors correctly hypothesized that IL3b ablation by 153 can be ameliorated by introduction of compensatory interactions in the distant PK1 subdomain. Accordingly, augmenting Watson-Crick pairing in PK-1 or shortening P2a each restored equilibrium SAM binding, and

improved dynamic range, thereby providing specific proof-of-principle examples of tertiary interactions that contribute to riboswitch tuning.

Conclusions and Future Directions

The mechanism by which individual riboswitches are tuned to meet the regulatory needs of associated transcriptional units has remained unclear. Batey and co-workers established that each of four tertiary-structure modules in the *metE* SAM-I riboswitch tolerates numerous individual variations but exhibits at least one nucleotide or base pair critical for switching, which supports the ‘coupled tertiary-interaction network’ hypothesis. By analyzing their findings in the context of the regulatory properties and sequences of other S-boxes, they conclude that IL3b and P2a are the principal tuning modules of SAM-I riboswitches. Overall the results help codify our understanding of regulatory RNA structure and function.

Future analysis would benefit from a broader analysis of S-box (and other) riboswitch tertiary-interaction networks and mutations thereof – not only by use of methods described here, but also by *in vivo* analysis where cooperative folding interactions are enhanced [32]. To develop a molecular level understanding, it would be informative to determine crystal structures of specific mutants, especially those that function synergistically, restoring activity lost by individual variations. Such approaches hold the key to unlocking our understanding of the gateways that control entry into specific biochemical pathways.

Acknowledgment

This work was supported in part by National Institutes of Health grant GM063162 to J.E.W.

References

- [1]. Breaker RR. Riboswitches and the RNA world. Cold Spring Harb Perspect Biol. 2012; 4:a003566. [PubMed: 21106649]
- [2]. Blount KF, Wang JX, Lim J, Sudarsan N, Breaker RR. Antibacterial lysine analogs that target lysine riboswitches. Nat Chem Biol. 2007; 3:44–9. [PubMed: 17143270]
- [3]. Lee ER, Blount KF, Breaker RR. Roseoflavin is a natural antibacterial compound that binds to FMN riboswitches and regulates gene expression. RNA Biol. 2009; 6:187–94. [PubMed: 19246992]
- [4]. Sudarsan N, Cohen-Chalamish S, Nakamura S, Emilsson GM, Breaker RR. Thiamine pyrophosphate riboswitches are targets for the antimicrobial compound pyrithiamine. Chem Biol. 2005; 12:1325–35. [PubMed: 16356850]
- [5]. Blount KF, Megyola C, Plummer M, Osterman D, O’Connell T, Aristoff P, et al. A Novel Riboswitch-Binding Flavin Analog that Protects Mice Against Clostridium difficile Infection without Inhibiting Cecal Flora. Antimicrob Agents Chemother. 2015
- [6]. Blount KF, Breaker RR. Riboswitches as antibacterial drug targets. Nat Biotechnol. 2006; 24:1558–64. [PubMed: 17160062]
- [7]. Breaker RR. Prospects for riboswitch discovery and analysis. Mol Cell. 2011; 43:867–79. [PubMed: 21925376]
- [8]. Batey RT. Recognition of S-adenosylmethionine by riboswitches. Wiley Interdiscip Rev RNA. 2011; 2:299–311. [PubMed: 21957011]
- [9]. Lu C, Ding F, Chowdhury A, Pradhan V, Tomsic J, Holmes WM, et al. SAM recognition and conformational switching mechanism in the Bacillus subtilis yitJ S box/SAM-I riboswitch. J Mol Biol. 2010; 404:803–18. [PubMed: 20951706]

- [10]. Montange RK, Batey RT. Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature*. 2006; 441:1172–5. [PubMed: 16810258]
- [11]. Trausch JJ, Xu Z, Edwards AL, Reyes FE, Ross PE, Knight R, et al. Structural basis for diversity in the SAM clan of riboswitches. *Proc Natl Acad Sci U S A*. 2014; 111:6624–9. [PubMed: 24753586]
- [12]. Weinberg Z, Regulski EE, Hammond MC, Barrick JE, Yao Z, Ruzzo WL, et al. The aptamer core of SAM-IV riboswitches mimics the ligand-binding site of SAM-I riboswitches. *RNA*. 2008; 14:822–8. [PubMed: 18369181]
- [13]. Weinberg Z, Wang JX, Bogue J, Yang J, Corbino K, Moy RH, et al. Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol*. 2010; 11:R31. [PubMed: 20230605]
- [14]. Irnov, Kertsburg A, Winkler WC. Genetic control by cis-acting regulatory RNAs in *Bacillus subtilis*: general principles and prospects for discovery. *Cold Spring Harb Symp Quant Biol*. 2006; 71:239–49. [PubMed: 17381303]
- [15]. Winkler WC, Nahvi A, Sudarsan N, Barrick JE, Breaker RR. An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat Struct Biol*. 2003; 10:701–7. [PubMed: 12910260]
- [16]. Tomsic J, McDaniel BA, Grundy FJ, Henkin TM. Natural variability in S-adenosylmethionine (SAM)-dependent riboswitches: S-box elements in *Bacillus subtilis* exhibit differential sensitivity to SAM in vivo and in vitro. *J Bacteriol*. 2008; 190:823–33. [PubMed: 18039762]
- [17]. Wostenberg C, Ceres P, Polaski JT, Batey RT. A highly coupled network of tertiary interactions in the SAM-I riboswitch and their role in regulatory tuning. *J. Molecular Biology*. 2015; 427:3473–90.
- [18]. Gusarov I, Nudler E. The mechanism of intrinsic transcription termination. *Mol Cell*. 1999; 3:495–504. [PubMed: 10230402]
- [19]. Yarnell WS, Roberts JW. Mechanism of intrinsic transcription termination and antitermination. *Science*. 1999; 284:611–5. [PubMed: 10213678]
- [20]. Artsimovitch I, Henkin TM. In vitro approaches to analysis of transcription termination. *Methods*. 2009; 47:37–43. [PubMed: 18948199]
- [21]. Dann CE 3rd, Wakeman CA, Sieling CL, Baker SC, Irnov I, Winkler WC. Structure and mechanism of a metal-sensing regulatory RNA. *Cell*. 2007; 130:878–92. [PubMed: 17803910]
- [22]. Trausch JJ, Ceres P, Reyes FE, Batey RT. The structure of a tetrahydrofolate-sensing riboswitch reveals two ligand binding sites in a single aptamer. *Structure*. 2011; 19:1413–23. [PubMed: 21906956]
- [23]. Treiber DK, Williamson JR. Beyond kinetic traps in RNA folding. *Curr Opin Struct Biol*. 2001; 11:309–14. [PubMed: 11406379]
- [24]. Woodson SA. Folding mechanisms of group I ribozymes: role of stability and contact order. *Biochem Soc Trans*. 2002; 30:1166–9. [PubMed: 12440997]
- [25]. Behrouzi R, Roh JH, Kilburn D, Briber RM, Woodson SA. Cooperative tertiary interaction network guides RNA folding. *Cell*. 2012; 149:348–57. [PubMed: 22500801]
- [26]. Schroeder KT, Daldrop P, Lilley DM. RNA tertiary interactions in a riboswitch stabilize the structure of a kink turn. *Structure*. 2011; 19:1233–40. [PubMed: 21893284]
- [27]. Lilley DM. The K-turn motif in riboswitches and other RNA species. *Biochim Biophys Acta*. 2014; 1839:995–1004. [PubMed: 24798078]
- [28]. Heppell B, Lafontaine DA. Folding of the SAM aptamer is determined by the formation of a K-turn-dependent pseudoknot. *Biochemistry*. 2008; 47:1490–9. [PubMed: 18205390]
- [29]. McDaniel BA, Grundy FJ, Henkin TM. A tertiary structural element in S box leader RNAs is required for S-adenosylmethionine-directed transcription termination. *Mol Microbiol*. 2005; 57:1008–21. [PubMed: 16091040]
- [30]. Montange RK, Mondragon E, van Tyne D, Garst AD, Ceres P, Batey RT. Discrimination between closely related cellular metabolites by the SAM-I riboswitch. *J Mol Biol*. 2010; 396:761–72. [PubMed: 20006621]

- [31]. Stoddard CD, Montange RK, Hennelly SP, Rambo RP, Sanbonmatsu KY, Batey RT. Free state conformational sampling of the SAM-I riboswitch aptamer domain. *Structure*. 2010; 18:787–97. [PubMed: 20637415]
- [32]. Strulson CA, Boyer JA, Whitman EE, Bevilacqua PC. Molecular crowders and cosolutes promote folding cooperativity of RNA under physiological ionic conditions. *RNA*. 2014; 20:331–47. [PubMed: 24442612]

Figure 1.

SAM-I riboswitch tertiary-interaction networks from crystal structures. (a) IL3a-IL3b cross-strand base stacking in the *B. subtilis yitJ* riboswitch (PDB code 4KQY) [9] exhibits bulged A47 in IL3a, whereas IL3b comprises two adenosines in one strand but no bulged bases in the opposing strand; by contrast, IL3b of *metE* possesses an adenosine and cytidine opposite a single bulged adenosine. Primary labels and the color scheme are derived from *metE* [17]; parenthetical labels are from crystal structures. Δ indicates a deletion that precludes alignment of the *metE* sequence to the *yitJ* structure. (b) The k-turn of the *Thermoanaerobacter tengcongensis (Tte) metF* riboswitch (PDB code 2GIS) [10] comprises bulged nucleotides 5'-GUA-3' flanked by sheared G•A pairs as in *metE*. A third, unconserved G•A interaction closes the 3'-end, whereas *metE* uses C55•C74. (c). The PK-1 interaction from the *Tte metF* riboswitch [10] has three canonical pairs and one U•U mismatch like *metE*, although the latter wobble forms the closing pair in *metE*. The *metF* P2b-J3/4 structure illustrates how a comparable minor-groove triplex likely forms in *metE*.