

Structure of the SAM-II riboswitch bound to S-adenosylmethionine

Sunny D Gilbert¹, Robert P Rambo², Daria Van Tyne¹ & Robert T Batey¹

In bacteria, numerous genes harbor regulatory elements in the 5' untranslated regions of their mRNA, termed riboswitches, which control gene expression by binding small-molecule metabolites. These sequences influence the secondary and tertiary structure of the RNA in a ligand-dependent manner, thereby directing its transcription or translation. The crystal structure of an S-adenosylmethionine-responsive riboswitch found predominantly in proteobacteria, SAM-II, has been solved to reveal a second means by which RNA interacts with this important cellular metabolite. Notably, this is the first structure of a complete riboswitch containing all sequences associated with both the ligand binding aptamer domain and the regulatory expression platform. Chemical probing of this RNA in the absence and presence of ligand shows how the structure changes in response to S-adenosylmethionine to sequester the ribosomal binding site and affect translational gene regulation.

Riboswitches generally act as genetic regulatory elements through the interplay of two distinct domains in the 5' untranslated region (5' UTR) of an mRNA: the aptamer domain, which directly binds a metabolite, and a downstream expression platform containing a secondary structural switch that determines whether the gene will be expressed¹. Among the known mRNA elements that specifically bind small molecules *in vivo*, there are now at least three characterized S-adenosylmethionine (SAM)-responsive riboswitches—SAM-I^{2,3}, SAM-II (SAM- α)^{4,5} and SAM-III (SAM_{MK})^{6,7}—underscoring the importance of riboregulation of sulfur metabolism by SAM. Each riboswitch has a distinct primary and secondary structure, and each seems to be mutually exclusive in a particular bacterial genome^{4,8}. To date, SAM-II riboswitches have been found only in proteobacteria, particularly α -proteobacteria, whereas SAM-I, the only other SAM riboswitch whose tertiary structure has been solved⁹, is generally found in Gram-positive bacterial genomes¹⁰.

The SAM-I and SAM-II riboswitches represent completely evolutionarily independent solutions for RNA-mediated SAM recognition coupled to gene regulation⁴. Although there are numerous examples of evolutionarily distinct proteins converging on the same function, we present the first structural analysis of this concept for biological RNAs. By determining the crystallographic structure of a SAM-bound SAM-II riboswitch that is found in the 5' UTR of the *metX* gene from an environmental sequence in the Sargasso Sea metagenome¹¹, we sought to show how two disparate RNA structures evolved to recognize the same ligand by comparison with the solved SAM-I structure⁹. Globally, the structures of the SAM-I and SAM-II riboswitches differ dramatically, but converge on similar chemical

strategies to bind ligand and to discriminate against noncognate ligands within their binding pockets. Chemical probing analysis of the SAM-II structure indicates that the tertiary architecture of the riboswitch undergoes a ligand-dependent structural transition that acts by tightening the global fold, sequestering putative regulatory sequences involved in translation. SAM-II also has a pseudoknot architecture that resembles other RNAs, expanding the known functions of this ubiquitous motif.

RESULTS

Crystallization of the SAM-II riboswitch

To further understand how mRNAs respond to SAM, we have determined the structure of the SAM-II riboswitch in complex with SAM. An RNA that yielded diffraction-quality crystals was found in a survey of 13 phylogenetic variants that differ in the lengths of two primary helices (P1 and P2a, **Fig. 1a**) in the predicted secondary structure (**Supplementary Table 1** online). Crystals that diffract X-rays to 2.8-Å resolution were obtained using a sequence found upstream of the *metX* gene (encoding homoserine acetyltransferase) in the Sargasso Sea metagenome¹¹ (**Supplementary Methods** online). The 3' end of this sequence is found 2 nucleotides (nt) upstream of the initiating AUG codon (**Supplementary Fig. 1a** online), suggesting that it regulates expression at the translational level by occluding the ribosome binding site (RBS), including the Shine-Dalgarno sequence. The crystallized SAM-II riboswitch (**Supplementary Fig. 1b**) fully encompasses the functional core of this regulatory element as it retains every sequence element that is greater than 80% conserved across a phylogenetic alignment of more than 70 representatives⁴ (**Supplementary Table 1**).

¹Department of Chemistry and Biochemistry, Campus Box 215, University of Colorado, Boulder, Colorado 80309, USA. ²Life Science Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Building 6 Room 2105, Berkeley, California 94720, USA. Correspondence should be addressed to R.T.B. (robert.batey@colorado.edu).

Received 10 August 2007; accepted 21 November 2007; published online 20 January 2008; doi:10.1038/nsmb.1371

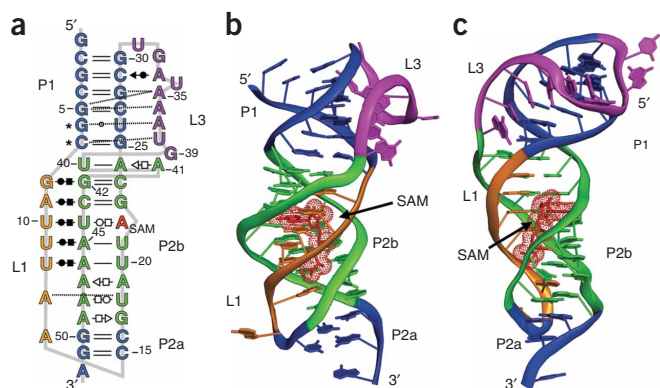


Figure 1 Global structure of SAM-II bound to *S*-adenosylmethionine (SAM). (a) Secondary structure of the Env12 *metX* SAM-II riboswitch showing base-pairing reflecting the tertiary structure of the SAM-bound RNA. Base interactions are shown using the notation of Leontis and Westhof⁴⁷. Circles indicate an interaction involving the Watson-Crick face, squares the Hoogsteen face and triangles the sugar edge; black symbols denote a parallel arrangement and open symbols denote an antiparallel arrangement. Dashed lines denote hydrogen bonds that cannot be described as one of the standard pairing interactions. Starred residues indicate sequences that deviate from the wild-type RNA. Colors of the bases reflect their position in the tertiary structure (blue, P1 and P2a; green, P2b; magenta, L3; orange, L1). (b) Cartoon of the global structure of the RNA; colors are consistent with the secondary structure. The red dots represent the van der Waals surface of SAM. (c) 90° rotation of the perspective shown in b.

Obtaining experimental phases required the alteration of two nonconserved Watson-Crick pairs in P1 (Fig. 1a, asterisks) for the purpose of creating a heavy-atom derivative binding site¹². This change has only a small effect on the affinity of the RNA for SAM (0.67 ± 0.06 and 1.35 ± 0.13 μM for the wild-type and altered RNAs, respectively; $\Delta\Delta G = 0.4$ kcal mol⁻¹) as measured by isothermal titration calorimetry (ITC) (Supplementary Fig. 2 online). Phasing was achieved by a combination of SAD with a cesium derivative data set to locate heavy atoms and single isomorphous replacement with anomalous scattering (SIRAS) using both the native and cesium derivative data sets to calculate experimental phases (Supplementary Fig. 3 online). All of the nucleotides in the RNA, as well as SAM in three slightly different RNA–ligand complexes in the asymmetric unit (Supplementary Fig. 4 online), were built and refined without noncrystallographic symmetry (NCS) constraints, to yield a final model with $R_{\text{xtal}} = 20.6\%$ and $R_{\text{free}} = 26.3\%$.

Structure of *S*-adenosylmethionine-bound SAM-II riboswitch

The global architecture of the SAM-II riboswitch conforms to a classic (H-type) pseudoknot¹³ (Fig. 1), different from the four-way-junction architecture of the SAM-I riboswitch⁹. The secondary structure consists of two Watson-Crick-paired helices (P1 and P2a) and two loop regions (L1 and L3). A third helical element, not predicted from sequence alignment, comprises highly conserved sequences in L2 and L3 to create P2b (Fig. 1a, Supplementary Fig. 1b). Loops L1 and L3 interact with the major and minor grooves of P2a/b and P1, respectively, to form an intricate tertiary structure with the SAM binding pocket located in the center of P2b (Fig. 1b,c). Each helical segment stacks

upon the next without any appreciable distortions to create a nearly straight structure.

The topology and sequence content of the L1 and L3 loops are similar to those observed in other classic pseudoknots, in which L1 is often uracil rich and crosses the major groove of P2 (refs. 14,15), whereas L3 typically is adenine rich and crosses the minor groove of P1 (refs. 15–18). In the case of SAM-II, a stack of four adenosines in L3 (A33, A35–37) rotate clockwise (as viewed from A33) along the minor groove face of P1, such that A33 interacts through its Hoogsteen face, whereas A37 uses its sugar face in a similar way to a type-I A minor groove triple¹⁹ (Fig. 2a). Bases in L3 do not form planar triples with the minor groove of P1, but rather are skewed at an $\sim 70^\circ$ angle with respect to P1, forming hydrogen bonds with two successive base pairs (Fig. 2a). As a result, hydrogen bonds are formed extensively between residues of L3 and the minor groove side of every base pair of P1, except the C2–G30 base pair at the 5' end of the helix (Fig. 2a). The *glmS* ribozyme contains a similar tertiary interaction between an adenosine-rich internal loop in the P4–P4.1 helix and the minor groove of the adjacent P2.1 helix^{20,21} (Supplementary Fig. 4 online). Three adenosines rotate clockwise from the Watson-Crick face to the sugar face, completing a 120° turn compared to the 180° turn

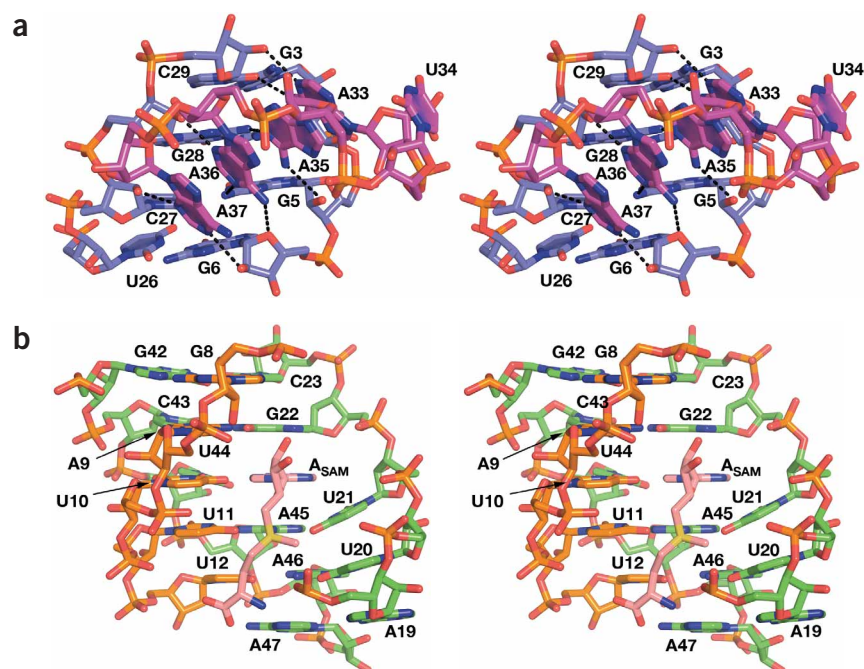


Figure 2 Detailed structural view of the A minor twist motif and the ligand binding pocket, key tertiary interactions in the ligand-bound structure. (a) Stereoview of the interactions between L3 (magenta) and the P1 helix (blue), emphasizing the role of four stacked adenosine residues in cementing the loop to the minor groove. (b) Stereoview of the binding pocket of *S*-adenosylmethionine (SAM; salmon) with the P2b helix. The adenine base of SAM (A_{SAM}) is accommodated by an opening in the 5' strand of P2b between U21 and G22, whereas the ribose sugar and methionine group reside in the narrow major groove of the triplex.

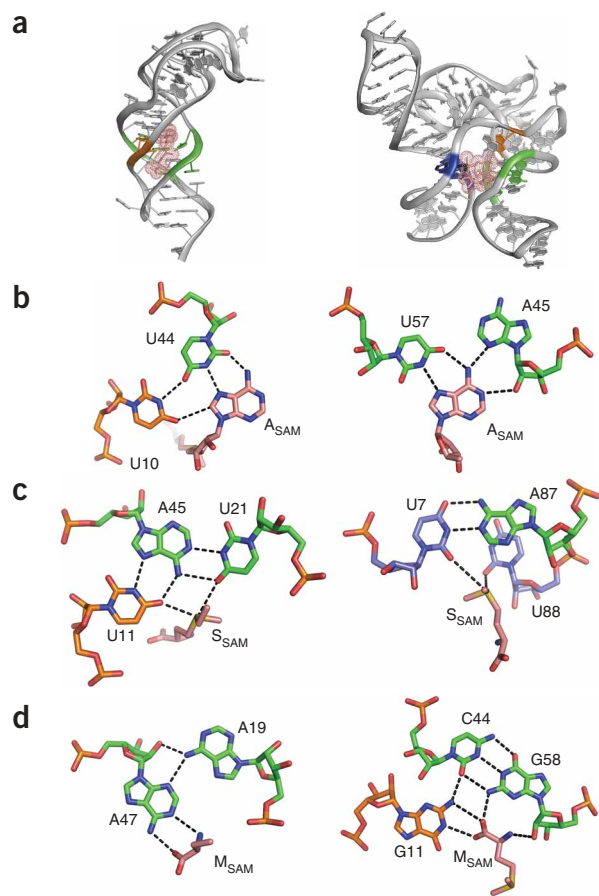


Figure 3 Comparison of the SAM-II (left) and SAM-I (right) binding pocket interactions with *S*-adenosylmethionine (SAM). **(a)** Global structures of SAM-II and SAM-I riboswitches. Specific residues that interact directly with SAM are colored to indicate their position around the binding pocket. **(b)** Hydrogen-bonding interactions involving the adenine moiety of SAM (A_{SAM}). **(c)** Hydrogen-bonding and electrostatic interactions involving the positively charged sulfur moiety and the methyl group of SAM. **(d)** Interactions between the RNA and the main chain atoms of the methionine residue of SAM.

observed in SAM-II. This A minor twist motif, distinct from the more common A minor triple motif^{22,23}, is likely to be another fundamental means by which an adenine-rich sequence forms a long-range tertiary interaction with the minor groove of an RNA helix.

The SAM binding pocket is created by formation of an extended triplex between L1 and the major groove of P2b. One end of the P2b-L1 interaction is defined by a single minor groove triple between A24–U40•A41 (Fig. 1a), followed by a series of base triples formed by nucleotides G8–U12 in L1 with the major groove face of P2b. The Watson-Crick face of each nucleotide in L1 interacts with the Hoogsteen face of nucleotides in the 3' strand of the P2b helix (nucleotides 42–46) (Fig. 2b). A sheared A19•A47 pair terminates this near-perfect triplex. Below this pair, the P2b triplex transitions into the P2a Watson-Crick-paired helix, defining the lower boundary of the SAM binding pocket. An isolated A13•(U18–A48) triple and a sheared G17•A49 pair comprise this transition. The >80% purine character of residues at positions 17 and 18 represents the only phylogenetic conservation within these pairs (Supplementary Table 1). The final nucleotide of L1, which is not conserved in phylogeny, is flipped out and makes no contacts with the 2-base-pair (bp) helix of P2a, which corresponds to the second helix in a classic pseudoknot fold.

SAM binds in an extended configuration along the major groove face of the P2b-L1 triplex, forming direct contacts with five successive base pairs and triples (Fig. 2b). The adenine moiety of SAM (A_{SAM}) participates in a base triple between U10 and U44, using its Hoogsteen face to pair with U44, similarly to what is observed for the SAM-I structure⁹ (Fig. 3a,b). Notably, this site seems to be created by the deletion of a single residue between U21 and G22; the gap in helix P2b thus requires adenine to be positioned in the helix such that its

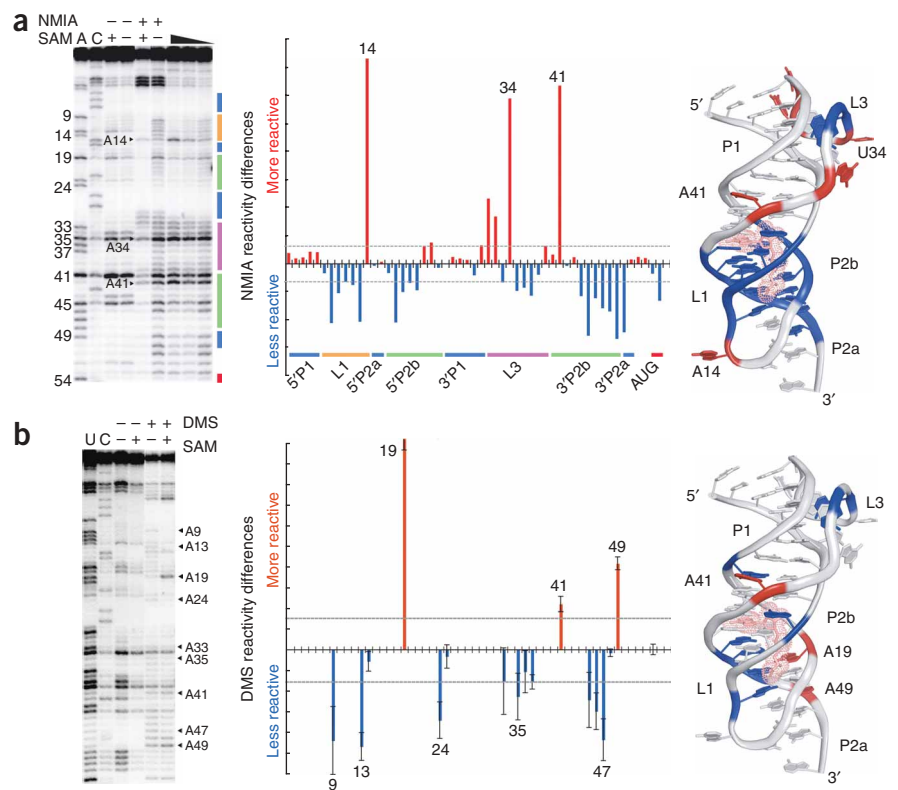
Watson-Crick face is solvent exposed in the minor groove of the triplex (Fig. 3b). The positively charged sulfur moiety and the activated methyl group are recognized by the carbonyl oxygen groups of U11 and U21 (Fig. 3c, left), explaining how SAM-II discriminates between SAM and *S*-adenosylhomocysteine (SAH)²⁴. Again, this is markedly similar to discrimination by SAM-I, which uses the minor groove carbonyl moieties of two universally conserved A-U pairs to interact electrostatically with the positive charge (Fig. 3c, right). The main chain atoms of methionine (carboxylate and amino groups) are positioned in the major groove adjacent to the sheared A19•A47 pair (Fig. 3d, left). However, these groups have different configurations in the three protomers in the asymmetric unit (Supplementary Fig. 5 online). In molecules A and C the amino and carboxylate groups lie along the Watson-Crick face of A47 (Fig. 3d, left), whereas in molecule B the amino group does not hydrogen-bond with the RNA. Recognition of the main chain atoms of methionine by the Watson-Crick face of an adenine base is similar to that observed in SAM-I, in which the Watson-Crick face of a guanine interacts with the carboxylate moiety (Fig. 3d, right). Only the methylene groups of the methionine side chain on the ligand gave poor definition, as little electron density was observed in this region in all three protomers. This probably results from weak van der Waals interactions between the hydrophobic groups of methionine and the adjacent carbonyl residues of U11 and U21 in the major groove of the RNA. Thus, all of the available functional groups in SAM seem to be directly or indirectly recognized by the mRNA, consistent with the results of a binding affinity analysis of SAM analogs to the SAM-II riboswitch²⁴. Despite the extensive recognition of SAM, the ligand is less extensively buried within the RNA (64% solvent inaccessible) than in other riboswitch–small molecule complexes²⁵.

SAM-II tertiary structure formation upon SAM binding

To assess potential conformational changes that occur in this riboswitch upon SAM binding, we chemically probed the complex with *N*-methylisatoic acid (NMIA), which surveys the dynamics of the ribose sugars²⁶, and dimethyl sulfate (DMS), which reacts with the N1 of unpaired adenosines and, to a lesser extent, the N3 of unpaired cytidines²⁷. The RNA used here contained the wild-type SAM-II sequence plus 21 residues located downstream of the aptamer domain, including the start codon for the *metX* gene and a potential switching sequence that is complementary to a consensus sequence in the aptamer domain⁴ (Supplementary Fig. 1b).

Chemical probing revealed a clear ligand-dependent stabilization of the pseudoknot, centered about the P2b helix. NMIA modification of residues in the P1 helix was not observed in the absence of SAM (Supplementary Fig. 6a online), and these residues remain unmodified upon addition of SAM (Fig. 4a), consistent with previous studies done by inline probing⁴. L3, which packs against the minor groove of P1, also shows moderate decreases in NMIA reactivity upon SAM binding, indicating that this region of the RNA's structure is well established in the absence of ligand. In contrast, the 2'-hydroxyl

Figure 4 Chemical probing of the structural difference between unliganded and liganded S-adenosylmethionine II (SAM-II). All of the reactions used to calculate reactivity differences incurred upon SAM addition contained saturating amounts of ligand. **(a)** Reactivity of RNA with N-methylisatoic acid (NMIA), a reagent that probes backbone flexibility by specifically modifying dynamic 2'-hydroxyl groups. Positions of bars along the x-axis indicate nucleotide residues left to right and 5' to 3'. The y-axis represents the change in reactivity percentage at each position from the unliganded to liganded states relative to the whole RNA. Reactivity differences are considered significant (>0.5 for increased reactivity; <-0.5 for decreased reactivity; gray dashed lines) if they are >1 s.d. from the mean of the percentage change in NMIA reactivity of the 15% of residues (all in P1 and P2a). At right, residues that are more reactive in the SAM-bound structure are colored red on the crystal structure; residues that are less reactive are colored blue. The color scheme is the same in the histogram, with prominent residues indicated. **(b)** Dimethyl sulfate (DMS) reactivity difference, specifically modifying the solvent accessibility of the Watson-Crick face of adenosines. The x-axis corresponds to that of **a**, and the y-axis indicates the magnitude of the change (in 1×10^3 counts) in the amount of radioactivity and the corresponding change in the level of modification by DMS observed upon ligand binding. Gray dashed lines, outside of which changes in reactivity were considered significant, indicate 1 s.d. above the average change in reactivity observed for all nonadenosine residues (1,600) that we would expect to remain constant in the presence and absence of ligand. Most adenosines, with the exception of A19, A49 and to a lesser extent A41 because their Watson-Crick faces become exposed upon ligand binding, show a decrease in DMS reactivity in the presence of SAM.



groups of residues forming the SAM binding pocket in P2b and L1 become strongly protected from NMIA reactivity by SAM. These two elements of the pseudoknot require interactions with SAM to create a stable major groove triplex. Further supporting this conclusion, bases that are flipped out of the helical structure (A14, U34 and A41) had a strongly increased level of NMIA reactivity relative to the unliganded RNA (**Fig. 4a**). The increased reactivity of A14 correlates with the increased reactivity of a guanine residue at the same position in the *metA* SAM-II riboswitch identified by inline probing⁴. Finally, sequences downstream of the pseudoknot have little effect on its NMIA reactivity (as compared to an RNA lacking this additional sequence; S.D.G. and R.T.B., data not shown) in the presence or absence of SAM. This is consistent with the translationally repressive activity of this RNA, in which SAM binding serves to occlude the Shine-Dalgarno sequence contained within the pseudoknot, rather than a mechanism involving a secondary switch in downstream elements, as seen in transcriptional regulators.

Patterns of DMS reactivity of the N1 of adenosines in the SAM-bound RNA correlate well with the crystal structure (**Supplementary Fig. 6b**). Adenosines whose Watson-Crick faces are exposed because of their involvement in sheared base-pairing interactions (A19, A41 and A49) are all significantly modified in the presence of SAM. Watson-Crick base-paired adenosines in the P1 and P2b helices show low reactivity to DMS (A24 and A25) (**Supplementary Fig. 6b**). In particular, the protection of A25, which is part of the U38•A25-U7 base triple in the wild-type RNA, is consistent with the crystal structure, in which this residue was altered to form a U38•G25-C7 base triple. Adenosines elsewhere in the structure whose Watson-Crick

faces are involved in tertiary structure formation, such as the interaction of A33 and A35–A37 of L3 with the minor groove of P1, are also only weakly modified (**Fig. 4b**). In the crystallized RNA, A37 interacts with the minor groove of the engineered G6•U26 base pair in P1, probably reflecting a nearly identical interaction with a Watson-Crick pair at this position in wild-type RNA.

Direct contacts between adenosines and the ligand protect certain residues from DMS modification. A45, which along with U21 and U11 is involved in a hydrogen-bonding network that recognizes the charged sulfur group (**Fig. 3c**), is only weakly modified in the presence of SAM (**Supplementary Fig. 6b** online). The Watson-Crick face of A47 is also protected from DMS modification. This residue is involved in a sheared A•A pair and is conserved in $>95\%$ of SAM-II RNAs. It forms two hydrogen bonds through its N1 and N6 to the carbonyl and amino groups of the amino acid side chain, respectively, in two of the three protomers in the unit cell (**Fig. 3d**). These data support the idea that ligand-dependent formation of the sheared A•A pair is required to specifically recognize this moiety of SAM.

Differences found in chemical probing of the unliganded and liganded structures indicate that binding of SAM to the riboswitch clearly alters the local conformation of several adenosines. Consistent with the NMIA probing results, A25, the only adenosine residue in the P1 helix, shows only a small difference in modification level in the unliganded and liganded RNA (**Fig. 4b**). This is mirrored by the adenosine in the start codon (A54), further demonstrating that SAM binding does not influence the RNA's structure beyond the pseudoknot. Residues in the ligand binding pocket (A19, A41 and A49) that form noncanonical base-pairing interactions in the

ligand-bound structure in which their Watson-Crick faces are solvent exposed (**Supplementary Fig. 6c,d**) show an increased reactivity with DMS (**Fig. 4b**). Adenosine residues in P2b involved in base-pairing interactions and contacts with SAM (A45–A47) show a decrease in DMS reactivity upon ligand binding. SAM binding clearly induces tertiary architecture formation in the ligand binding pocket between L1 and P2b. Our data support a hypothesis in which P2b is composed of loosely stacked bases that achieve specific pairing patterns only in the presence of SAM, evidenced by the substantial rearrangements throughout the core of the RNA between the unliganded and liganded RNA.

DISCUSSION

With the increasing number of structurally characterized riboswitches, it is evident that these RNAs use tertiary architectures in a way similar to that of other biologically important RNAs. In a fashion reminiscent of the resemblance between the global organization of the guanine riboswitch²⁸ and the hammerhead ribozyme²⁹, and the architectural similarity of the SAM-I riboswitch⁹ and the core of the group I intron³⁰, SAM-II uses a motif that is highly recurrent in nature. The pseudoknot is ubiquitous throughout biological RNAs, with a number of diverse functions, including roles in translational regulation of viral genomes by inducing mRNA frameshifting³¹ and in telomerase RNA as a key element in telomerase activity^{15,32,33}. The architecture of SAM-II is consistent with other H-type pseudoknots, most notably the human telomerase RNA (hTR) pseudoknot core, which is involved in telomerase reverse transcriptase (TERT) repeat addition processivity³² (**Supplementary Fig. 7** online). Both the hTR pseudoknot and SAM-II contain triplexes at the junction between P1 and P2a/b that have a strong preference for A•U base pairs in this region¹⁵. The hTR core is a 5-bp triplex broken only by a noncanonical A•U Hoogsteen base pair in the center; the SAM-II core is also a 5-bp triplex, containing the A_{SAM} motif in the center forming a Hoogsteen base pair with U44. The riboswitch has created a ligand-dependent structure out of the pseudoknot through the deletion of a crucial residue in the middle of the L1–P2b triplex between U21 and G22. Interaction of the adenine moiety of SAM with U10 and U44 is then required to reestablish the central base triple, thereby stabilizing the pseudoknot. SAM-II adds to the functional diversity of the H-type pseudoknot, one of many recurrent structural motifs in RNA.

Riboswitches affect gene regulation by sequestering sequences within their ligand-bound tertiary structures that would otherwise be involved in secondary structure formation, or by directly shielding regulatory sequences involved in transcription or translation^{34–36}. In SAM-II, the RBS, including the purine-rich Shine-Dalgarno element (the AUG start codon is 2 nt downstream of the crystallized RNA), comprises the 3' side of the P2a/b helices and is intimately involved in SAM recognition⁴ (**Supplementary Fig. 1**). Chemical probing data indicate that SAM-dependent stabilization of the L1–P2b tertiary interaction is likely to be responsible for translational repression via an occlusion mechanism rather than by secondary structural switching. This is in contrast to purine³⁷ and thiamine pyrophosphate^{38,39} riboswitches, which repress translation by sequestering sequences involved in mutually exclusive secondary structures in a downstream expression platform. However, there are other instances in RNA biology involving the direct sequestering of translation regulators in tertiary structures. For example, leader sequences in the 5' UTR of mRNAs encoding ribosomal proteins often contain regulatory elements⁴⁰ that bind protein that is not incorporated into ribosomes to directly block access by the 30S subunit^{40,41}. Recently identified orphan riboswitches *sucA* and COG4708 also seem to operate by the

same mode of gene regulation⁴². Thus, as revealed by a combination of X-ray crystallographic and chemical probing approaches, the SAM-II riboswitch creates a complete genetic regulatory element from a compact structural motif.

METHODS

RNA library synthesis and purification. A series of RNAs corresponding to secondary structure and sequence variations of the SAM-II RNA observed across phylogeny was constructed according to the length of the P1 and P2 helices of the minimal riboswitch⁴ (**Supplementary Fig. 1**). These helices vary in length by 5–8 bp and 2–6 bp, respectively. RNAs for ITC, X-ray crystallography and chemical probing were transcribed *in vitro* from PCR-amplified DNA fragments, as described⁴³. RNA transcription and purification details can be found in **Supplementary Methods**.

Crystallization and structure determination. Crystals of the 5' leader sequence of the *metX* gene from the Env12 genome representing the SAM-II aptamer domain in complex with SAM were obtained by hanging drop vapor diffusion and flash frozen in mother liquor (8 mM cobalt hexamine chloride, 640 mM ammonium acetate, 10% PEG 1K, 10 mM barium chloride, 50 mM sodium cacodylate, pH 6.0) plus 8% (v/v) (2*R*,3*R*)-butanediol. Data were collected on a home X-ray source using CuK α radiation. Crystallization, structure determination and refinement details are supplied in the **Supplementary Methods**. All figures were generated in PyMOL (DeLano Scientific)⁴⁴.

Chemical probing and band intensity analysis. Methods for NMIA and DMS probing of the RNA were similar to those previously described^{26,27,45}. RNA was heat refolded before each reaction. NMIA and DMS modification reactions were carried out at 25 °C and 30 °C, respectively. Band intensity was assessed using the SAFA program available from MatLab⁴⁶. Data was normalized and averaged in Microsoft Excel. Procedures for RNA preparation, RNA

Table 1 Data collection, phasing and refinement statistics (SIRAS)

	Native (lithium)	Derivative (cesium)
Data collection		
Space group	C2	C2
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	115.26, 48.10, 109.62	115.64, 48.29, 109.51
α , β , γ (°)	90, 108.26, 90	90, 108.20, 90
Resolution (Å)	20–2.6 (2.69–2.60)	20–2.8 (2.90–2.80) ¹
<i>R</i> _{merge}	0.053 (0.368)	0.061 (0.349)
<i>I</i> / σ <i>I</i>	28.5 (2.5)	31.9 (5.0)
Completeness (%)	89.6 (54.4)	97.9 (94.7)
Redundancy	3.6 (2.8)	7.8 (7.7)
Refinement		
Resolution (Å)		20–2.8 (2.97–2.80)
No. reflections		26,866 (4,033)
<i>R</i> _{work} / <i>R</i> _{free}		20.6/26.3 (40.7/40.8)
No. atoms		
RNA		3,375
Ligand/ion		84
Water		154
<i>B</i> -factors		
RNA		64.90
Ligand/ion		64.87
Water		45.50
r.m.s. deviations		
Bond lengths (Å)		0.0059
Bond angles (°)		1.017

¹Reflections obtained from a single crystal. Values in parentheses are for the highest-resolution shell.

modification and reverse-transcription reactions, denaturing PAGE for sequencing and the specifics of data handling are included in the **Supplementary Methods**.

Isothermal titration calorimetry. Binding of SAM to both the wild-type and crystallization RNAs was measured at 30 °C in 50 mM potassium HEPES, pH 7.5, 100 mM KCl and 10 mM MgCl₂ on a MicroCal isothermal titration calorimeter⁴³. Details regarding sample preparation and experimental parameters are available in the **Supplementary Methods**.

Accession codes. Protein Data Bank: Coordinates and structure factors have been deposited with accession code 2QWY.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank S. Edwards for maintaining the home X-ray source, A. Edwards for critical comments on the manuscript and J. Kieft for useful discussions and input on this project. This work was supported by grants from the US National Institutes of Health (GM 073850) and the American Heart Association (Scientist Development Grant) to R.T.B.

AUTHOR CONTRIBUTIONS

S.D.G. and R.T.B. conceived and carried out the experiments, solved the structure and wrote the paper. R.P.R. provided critical crystallographic expertise throughout the data collection and refinement process. D.V.T. assisted in screening the initial RNA library to find a crystallizable RNA variant.

Published online at <http://www.nature.com/nsmb/>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions>

- Winkler, W.C. & Breaker, R.R. Regulation of bacterial gene expression by riboswitches. *Annu. Rev. Microbiol.* **59**, 487–517 (2005).
- Grundy, F.J. & Henkin, T.M. The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in Gram-positive bacteria. *Mol. Microbiol.* **30**, 737–749 (1998).
- Winkler, W.C., Nahvi, A., Sudarsan, N., Barrick, J.E. & Breaker, R.R. An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat. Struct. Biol.* **10**, 701–707 (2003).
- Corbino, K.A. *et al.* Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in α -proteobacteria. *Genome Biol.* **6**, R70 (2005).
- Lim, J., Winkler, W.C., Nakamura, S., Scott, V. & Breaker, R.R. Molecular-recognition characteristics of SAM-binding riboswitches. *Angew. Chem. Int. Edn. Engl.* **45**, 964–968 (2006).
- Fuchs, R.T., Grundy, F.J. & Henkin, T.M. The S_{MK} box is a new SAM-binding RNA for translational regulation of SAM synthetase. *Nat. Struct. Mol. Biol.* **13**, 226–233 (2006).
- Fuchs, R.T., Grundy, F.J. & Henkin, T.M. S-adenosylmethionine directly inhibits binding of 30S ribosomal subunits to the S_{MK} box translational riboswitch RNA. *Proc. Natl. Acad. Sci. USA* **104**, 4876–4880 (2007).
- Barrick, J.E. & Breaker, R.R. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol.* **8**, R239 (2007).
- Montange, R.K. & Batey, R.T. Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature* **441**, 1172–1175 (2006).
- Griffiths-Jones, S. *et al.* Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* **33**, D121–D124 (2005).
- Venter, J.C. *et al.* Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**, 66–74 (2004).
- Keel, A.Y., Rambo, R.P., Batey, R.T. & Kieft, J.S. A general strategy to solve the phase problem in RNA crystallography. *Structure* **15**, 761–772 (2007).
- Hilbers, C.W., Michiels, P.J.A. & Heus, H.A. New developments in structure determination of pseudoknots. *Biopolymers* **48**, 137–158 (1998).
- Nonin-Lecomte, S., Felten, B. & Dardel, F. NMR structure of the *Aquifex aeolicus* tmRNA pseudoknot PK1: new insights into the recoding event of the ribosomal translation. *Nucl. Acids Res.* **34**, 1847–1853 (2006).
- Theimer, C.A., Blois, C.A. & Feigon, J. Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol. Cell* **17**, 671–682 (2005).
- Chen, X. *et al.* Structural and functional studies of retroviral RNA pseudoknots involved in ribosomal frameshifting: nucleotides at the junction of the two stems are important for efficient ribosomal frameshifting. *EMBO J.* **14**, 842–852 (1995).
- Pallan, P.S. *et al.* Crystal structure of a luteoviral RNA pseudoknot and model for a minimal ribosomal frameshifting motif. *Biochemistry* **44**, 11315–11322 (2005).
- Su, L., Chen, L., Egli, M., Berger, J.M. & Rich, A. Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. *Nat. Struct. Biol.* **6**, 285–292 (1999).
- Doherty, E.A., Batey, R.T., Masquida, B. & Doudna, J.A. A universal mode of helix packing in RNA. *Nat. Struct. Biol.* **8**, 339–343 (2001).
- Cochrane, J.C., Lipchick, S.V. & Strobel, S.A. Structural investigation of the *glmS* ribozyme bound to its catalytic cofactor. *Chem. Biol.* **14**, 97–105 (2007).
- Klein, D.J. & Ferre-D'Amare, A.R. Structural basis of *glmS* ribozyme activation by glucosamine-6-phosphate. *Science* **313**, 1752–1756 (2006).
- Battle, D.J. & Doudna, J.A. Specificity of RNA-RNA helix recognition. *Proc. Natl. Acad. Sci. USA* **99**, 11676–11681 (2002).
- Nissen, P., Ippolito, J.A., Ban, N., Moore, P.B. & Steitz, T.A. RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc. Natl. Acad. Sci. USA* **98**, 4899–4903 (2001).
- Lim, J., Winkler, W.C., Nakamura, S., Scott, V. & Breaker, R.R. Molecular-recognition characteristics of SAM-binding riboswitches. *Angew. Chem. Int. Edn. Engl.* **45**, 964–968 (2006).
- Edwards, T.E., Klein, D.J. & Ferre-D'Amare, A.R. Riboswitches: small-molecule recognition by gene regulatory RNAs. *Curr. Opin. Struct. Biol.* **17**, 273–279 (2007).
- Wilkinson, K.A., Merino, E.J. & Weeks, K.M. Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.* **1**, 1610–1616 (2006).
- Moazed, D. & Noller, H.F. Transfer RNA shields specific nucleotides in 16S ribosomal RNA from attack by chemical probes. *Cell* **47**, 985–994 (1986).
- Batey, R.T., Gilbert, S.D. & Montange, R.K. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature* **432**, 411–415 (2004).
- Martick, M. & Scott, W.G. Tertiary contacts distant from the active site prime a ribozyme for catalysis. *Cell* **126**, 309–320 (2006).
- Adams, P.L., Stahley, M.R., Kosek, A.B., Wang, J. & Strobel, S.A. Crystal structure of a self-splicing group I intron with both exons. *Nature* **430**, 45–50 (2004).
- Brierley, I., Pennell, S. & Gilbert, R.J. Viral RNA pseudoknots: versatile motifs in gene expression and replication. *Nat. Rev. Microbiol.* **5**, 598–610 (2007).
- Moriarty, T.J., Marie-Egyptienne, D.T. & Autexier, C. Functional organization of repeat addition processivity and DNA synthesis determinants in the human telomerase multimer. *Mol. Cell. Biol.* **24**, 3720–3733 (2004).
- Tzfati, Y., Knight, Z., Roy, J. & Blackburn, E.H. A novel pseudoknot element is essential for the action of a yeast telomerase. *Genes Dev.* **17**, 1779–1788 (2003).
- Winkler, W.C. & Breaker, R.R. Genetic control by metabolite-binding riboswitches. *ChemBioChem* **4**, 1024–1032 (2003).
- Gilbert, S.D. & Batey, R.T. Riboswitches: fold and function. *Chem. Biol.* **13**, 805–807 (2006).
- Nudler, E. & Mironov, A.S. The riboswitch control of bacterial metabolism. *Trends Biochem. Sci.* **29**, 11–17 (2004).
- Serganov, A. *et al.* Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem. Biol.* **11**, 1729–1741 (2004).
- Mironov, A.S. *et al.* Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* **111**, 747–756 (2002).
- Serganov, A., Polonskaia, A., Phan, A.T., Breaker, R.R. & Patel, D.J. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* **441**, 1167–1171 (2006).
- Zengel, J.M. & Lindahl, L. Diverse mechanisms for regulating ribosomal protein synthesis in *Escherichia coli*. *Prog. Nucleic Acid Res. Mol. Biol.* **47**, 331–370 (1994).
- Merianos, H.J., Wang, J. & Moore, P.B. The structure of a ribosomal protein S8/spc operon mRNA complex. *RNA* **10**, 954–964 (2004).
- Weinberg, Z. *et al.* Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Res.* **35**, 4809–4819 (2007).
- Gilbert, S.D., Stoddard, C.D., Wise, S.J. & Batey, R.T. Thermodynamic and kinetic characterization of ligand binding to the purine riboswitch aptamer domain. *J. Mol. Biol.* **359**, 754–768 (2006).
- DeLano, W.L. The PyMOL Molecular Graphics System. (2002).
- Kjems, J., Egebjerg, J. & Christiansen, J. *Analysis of RNA-Protein Complexes in vitro* (Elsevier, Amsterdam, 1998).
- Das, R., Laederach, A., Pearlman, S.M., Herschlag, D. & Altman, R.B. SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *RNA* **11**, 344–354 (2005).
- Leontis, N.B. & Westhof, E. Geometric nomenclature and classification of RNA base pairs. *RNA* **7**, 499–512 (2001).