

The structure of the SAM/SAH-binding riboswitch

A. Katharina Weickhmann¹, Heiko Keller¹, Jan P. Wurm^{1,2}, Elisabeth Strebitzer³, Michael A. Juen³, Johannes Kremser³, Zasha Weinberg⁴, Christoph Kreutz³, Elke Duchardt-Ferner¹ and Jens Wöhnert^{1,*}

¹Institute for Molecular Biosciences and Center for Biomolecular Magnetic Resonance (BMRZ), Goethe-University Frankfurt, Max-von-Laue-Strasse 9, 60438 Frankfurt/M., Germany, ²Institute of Biophysics and Physical Biochemistry, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Bavaria, Germany, ³Institute of Organic Chemistry, Centre for Molecular Biosciences (CMBI), University of Innsbruck, Innrain 80/82, 6020 Innsbruck, Austria and ⁴Bioinformatics Group, Department of Computer Science and Interdisciplinary Centre for Bioinformatics, Institute of Informatics, University of Leipzig, Härtelstrasse 16-18, 04107 Leipzig, Germany

Received October 09, 2018; Revised December 12, 2018; Editorial Decision December 13, 2018; Accepted December 26, 2018

ABSTRACT

S-adenosylmethionine (SAM) is a central metabolite since it is used as a methyl group donor in many different biochemical reactions. Many bacteria control intracellular SAM concentrations using riboswitch-based mechanisms. A number of structurally different riboswitch families specifically bind to SAM and mainly regulate the transcription or the translation of SAM-biosynthetic enzymes. In addition, a highly specific riboswitch class recognizes S-adenosylhomocysteine (SAH)—the product of SAM-dependent methyl group transfer reactions—and regulates enzymes responsible for SAH hydrolysis. High-resolution structures are available for many of these riboswitch classes and illustrate how they discriminate between the two structurally similar ligands SAM and SAH. The so-called SAM/SAH riboswitch class binds both ligands with similar affinities and is structurally not yet characterized. Here, we present a high-resolution nuclear magnetic resonance structure of a member of the SAM/SAH-riboswitch class in complex with SAH. Ligand binding induces pseudoknot formation and sequestration of the ribosome binding site. Thus, the SAM/SAH-riboswitches are translational ‘OFF’-switches. Our results establish a structural basis for the unusual bispecificity of this riboswitch class. In conjunction with genomic data our structure suggests that the SAM/SAH-riboswitches might be an evolutionary late invention and not a remnant of a primordial RNA-world as suggested for other riboswitches.

INTRODUCTION

S-adenosylmethionine (SAM, Figure 1A, top) is a central metabolite and a widely used co-substrate in many different enzymatic reactions. SAM is produced enzymatically from adenosine triphosphate (ATP) and methionine by SAM-synthetase. In the reaction catalyzed by this enzyme, the sulfur atom of methionine attacks the C5' of ATP and cleaves off its entire triphosphate chain while forming a new carbon-sulfur bond. SAM contains a positively charged sulfonium center substituted by a methyl, an aminocarboxypropyl and a 5'-deoxyadenosyl group. Each of these three functional groups can now be transferred from SAM to other substrates in S_N2 reactions (1). Most common is the transfer of the methyl group to substrates ranging from small molecules to protein sidechains, RNAs and DNAs catalyzed by a myriad of specialized methyltransferase enzymes (2). In these reactions, SAM is converted into the neutral compound S-adenosylhomocysteine (SAH, Figure 1A, bottom). SAH is toxic for the cell and rapidly broken down by SAH-nucleosidases into adenine and S-ribosylhomocysteine (3,4) or by SAH-hydrolases into homocysteine and adenosine (5,6). Less common is the use of SAM as a cosubstrate in reactions transferring the aminocarboxypropyl group to another substrate. Aminocarboxypropyl transfer reactions occur for instance during post-transcriptional modifications of rRNAs and tRNAs leaving 5'-deoxy-5'-methylthioadenosine as the second reaction product (7,8). Furthermore, in some metabolic pathways SAM is modified prior to any transfer reactions by SAM-decarboxylase (9,10). The aminopropyl group of the decarboxy-SAM produced by this enzyme is then used in the biosynthesis of spermine, spermidine and other polyamines (11). The transfer of the 5'-deoxyadenosyl group from SAM to a substrate is for instance observed in enzymatic fluorination reactions (12,13).

*To whom correspondence should be addressed. Tel: +49 69 79829785; Fax: +49 69 798 29527; Email: woehnert@bio.uni-frankfurt.de

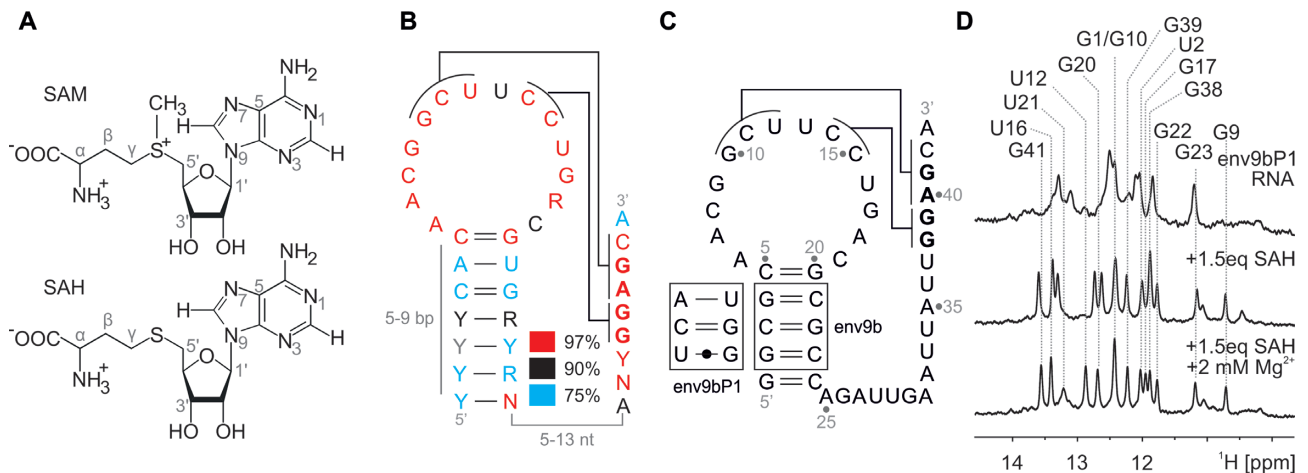


Figure 1. The SAM/SAH riboswitch and its ligands. (A) Chemical structures of SAM (top) and SAH (bottom). (B) Consensus sequence and secondary structure of the SAM/SAH-binding riboswitches. The ribosomal binding site is highlighted in bold letters. (C) Secondary structure of the SAM/SAH-binding riboswitch variant env9b and the P1 stem mutant env9bP1. The ribosomal binding site is highlighted in bold letters. The pseudoknot interactions indicated by connecting lines in (B) and (C) conform to the structure determined in this work and differ slightly from the original prediction. (D) Imino proton spectra of the env9bP1 RNA in the absence (top) and the presence of an 1.5-fold excess of SAH (middle) and 2 mM magnesium acetate (bottom). Resonance assignments in the SAH and Mg^{2+} bound state are indicated.

Due to the central importance of SAM-dependent reactions in metabolism, intracellular SAM concentrations and SAM biosynthesis are tightly regulated in all organisms. Transcription regulation mechanisms using a SAM-binding repressor protein (MetJ) are known for instance in *Escherichia coli* (14). However, many other bacteria use riboswitch-based transcriptional or translational control mechanisms to regulate SAM-biosynthesis (15,16). Riboswitches are structured domains in the 5'-UTRs of bacterial mRNAs that bind specifically to small molecules and change their structure upon ligand binding. These structural changes usually modulate either the stability of transcription termination signals or the accessibility of ribosome binding sites in these mRNAs. Depending on the structural properties of each riboswitch ligand binding can either activate ('ON'-switch) or inhibit ('OFF'-switch) transcription or translation (17).

SAM-sensing riboswitches are not only widespread in bacteria but also structurally highly diverse. Based on difference in sequence, secondary and tertiary structure they can be divided into three large superfamilies covering seven subclasses called SAM-I to SAM-VI and SAM-I/IV (15,16,18,19). Many SAM-riboswitches act as 'OFF'-switches for the expression of SAM-synthetase. While they differ in their tertiary structures and the architecture of their SAM-binding pockets, a common feature is the extensive amount of interactions between the RNA and the functional groups of the ligand SAM. The adenine moiety of SAM is involved in intermolecular base pairing or base-triplet interactions. The positively charged sulfonium group interacts electrostatically with polar carbonyl or hydroxyl groups of the RNA. Furthermore, in the SAM-I and SAM-II riboswitches, the carboxy and amino groups of the methionine side chain of SAM form hydrogen bonds to RNA nucleobases and/or the RNA backbone. Due to this extensive set of interactions between the ligand and the RNA, all SAM-riboswitches are highly specific for SAM.

In particular, these riboswitches discriminate strongly between SAM and SAH and bind SAM with a 100- to 1000-fold higher affinity compared to SAH. Interestingly, there is a specialized SAH-sensing riboswitch with the opposite pattern of specificity (20). It binds SAH with an ~1000-fold higher affinity compared to SAM and acts as an 'ON'-switch for the expression of SAH-hydrolase—an enzyme responsible for the breakdown of the toxic SAH. The high specificity of these riboswitches and their ability to discriminate strongly between SAM and SAH prevents unwanted regulatory cross talk between those two structurally very similar but functionally very different metabolites. This is an important feature for the regulatory logic in the associated metabolic pathways.

Defying these considerations, members of a novel class of riboswitches discovered in 2010 in the order Rhodobacterales of the α -proteobacteria bind SAM and SAH with similar affinities (21). Accordingly, this class of riboswitches was named the SAM/SAH-family of riboswitches. However, virtually all of its members were found in the 5'-UTRs of SAM-synthetase (MetK) mRNAs and were proposed to act as translational 'OFF'-switches. While it is immediately obvious, that translational repression of the SAM-synthetase upon SAM-binding to this riboswitch is a regulatory useful response, it is less obvious why these riboswitches should be able to bind to the breakdown product of SAM—SAH—and thereby shut off the expression of SAM synthetase. Sequence comparisons suggested a pseudoknot secondary structure for this riboswitch (Figure 1B). A 5'-hairpin loop structure contains a P1 stem with 5–9 bp and a formally single stranded loop region with 14 nt. A total of 12 of the 14 nt in the loop region are strictly conserved. The hairpin loop element is followed by a putatively single-stranded linker element of 5–13 nt of variable sequence and 8 conserved nucleotides that include the ribosome binding site. Some of these conserved nucleotides are complementary to nucleotides in the loop region of the 5'-hairpin al-

lowing the formation of the P2 stem of a putative H-type pseudoknot structure (Figure 1B). The three-dimensional structure of this riboswitch and its ligand recognition mode have not yet been determined.

Here, we present the high-resolution nuclear magnetic resonance (NMR) solution structure of the SAM/SAH-riboswitch bound to SAH. The structure rationalizes why this riboswitch in contrast to the classical SAM riboswitches binds SAM and SAH with similar affinity. We also show that ligand binding promotes the formation of an H-type pseudoknot structure that sequesters the ribosome-binding site and thereby induces ligand-dependent translational repression. Ligand binding also supports additional base pairing interactions in the loop region and between the hairpin loop and conserved nucleotides in the 3'-single-stranded region. SAM and SAH are bound in an intermolecular *trans* Hoogsteen/Watson-Crick ('reversed Hoogsteen') base pairing interaction between the adenosyl moiety of the cofactor and a strictly conserved U nucleotide of the riboswitch. Contacts between the sulfur atom of the cofactor and the RNA are limited to only a single polar interaction and therefore less extensive than in the other SAM-binding riboswitch families. Only a single electrostatic interaction is found between the amino group of the aminocarboxypropyl side chain of the ligand and the RNA. Overall, our structure and additional extensive thermodynamic data obtained for mutant RNAs and ligand analogs suggest that the SAM/SAH-riboswitch represents a structurally minimalistic solution for SAM recognition. Taken together with the poor ability to discriminate against SAH, its gene associations and its restricted occurrence in the Rhodobacterales suggest that the SAM/SAH riboswitch evolved in organisms that already contained an SAH-degrading enzyme or other mechanisms that keep SAH concentrations low enough to prevent interference with SAM-dependent riboswitch signaling.

MATERIALS AND METHODS

RNA preparation

RNAs were synthesized by *in vitro* transcription with T7 RNA polymerase and a linearized plasmid template DNA. The *in vitro* transcription mix contained either unlabeled (Sigma-Aldrich) or appropriately isotope-labeled nucleoside triphosphates (NTPs) (Silantes GmbH and Cambridge Isotope Laboratories, Inc.). To obtain RNA with a homogeneous 3'-end, the primary RNA-transcript contained a hepatitis delta virus ribozyme fused to the 3'-end of the target RNA. The processed RNA was purified as described (22). The purified RNAs were folded by heating to 80°C for 10 min followed by a rapid 5-fold dilution with ice cold water. The samples were exchanged into 25 mM potassium phosphate buffer (pH 6.2) containing 50 mM potassium chloride in multiple cycles of concentration and dilution in buffer using centrifugation devices (VivaSpin 2, MWC 3 kDa). RNA concentrations in the NMR samples ranged between 0.04 and 1.2 mM containing either 5–10% or 100% (v/v) D₂O, 2 mM magnesium acetate and up to four equivalents of ligand.

For assignment and structure calculation experiments samples containing unlabeled, uniformly ¹⁵N-, uniformly

¹³C,¹⁵N-, uniformly 5-D₁/ribose-3',4',5',5''-D₄-labeled and three different nucleotide-type selectively labeled samples (¹³C,¹⁵N-A/C, ¹³C,¹⁵N-G or ¹³C,¹⁵N-U) of the env9bP1 riboswitch (Figure 1C) in complex with either unlabeled or ¹³C,¹⁵N-labeled SAM or SAH were used. Furthermore, site and atom-specifically labeled samples (env9b: ¹⁵N₂-U12, ¹⁵N₂-U16, ¹⁵N₂-U37, ¹³C-C8-A7/¹³C-C6-U12, ¹³C-C8-G10/¹³C-C6-U16, ¹³C-C8-A6/¹³C-C6-C15, ¹³C-C6-C11/¹³C-C8-G17, ¹³C-C8-G9/¹³C-C6-C14; env9bP1: ¹³C-C6-U13/¹³C-C6-C19, ¹³C-C6-C8/¹³C-C8-A18, ¹³C-C8-A35/¹³C-C6-U36, ¹³C-C2,C8-A35/¹³C-C6-U37, ¹³C-C6-U34/¹³C-C2,C8-A35) were chemically synthesized (23–25). Pde-1-1 RNA was purchased commercially (Thermo Scientific) and deprotected according to the manufacturer's instructions.

Unlabeled SAM, SAH, adenosine, S-(5'-adenosyl)-3-thiopropylamine (decarboxy-SAH, Sigma-Aldrich), 5'-deoxy-5'-methylthioadenosine (Sigma-Aldrich) and S-(5'-adenosyl)-3-methylthiopropylamine (decarboxy-SAM, TRC, Canada) were purchased commercially and dissolved in water or NMR buffer, respectively.

Synthesis and purification of ¹³C,¹⁵N-labeled SAM and SAH

¹³C,¹⁵N-labeled SAM was enzymatically synthesized as described earlier (26) starting from ¹³C,¹⁵N-labeled methionine and ¹³C,¹⁵N-labeled ATP and purified using cation exchange chromatography.

¹³C,¹⁵N-labeled SAM was then incubated with the methyltransferase PaMTH1 from *Pseudospira anserina* and its substrate myricetin yielding ¹³C,¹⁵N-labeled SAH (27) which was further purified using RP-HPLC. Product purity was confirmed by NMR and thin-layer chromatography using ethanol/H₂O/acetic acid (65/34/1) as running buffer. The synthesis pathway for SAM and SAH is shown in Supplementary Figure S1.

NMR spectroscopy

All NMR experiments were carried out using Bruker AVANCE 600, 700, 800, 900 and 950 MHz spectrometers equipped with cryogenic triple resonance HCN- or HCP-probes and z-axis gradients. NMR experiments for resonance assignments have been described in detail before (28). NMR spectra were recorded at 5°C, 10°C and 20°C in 5–10% (v/v) D₂O for the detection of exchangeable protons and at 20°C in 100% D₂O for the non-exchangeable protons.

The data were processed using TOPSPIN 3.2 or 3.5 (Bruker BioSpin, Germany) and analyzed using CARA (29). Proton chemical shifts were referenced directly to 4,4-dimethyl-4-silapentane-1-sulfonic acid. Carbon, nitrogen and phosphorous chemical shifts were referenced indirectly (30).

Input restraints for structure calculations

All intra-, as well as intermolecular ¹H,¹H distance restraints were obtained from NOE intensities of 2D ¹H,¹H-NOESY, 2D ¹⁵N-CPMG-NOESY (31,32) and 2D or 3D ¹H,¹³C-NOESY-HSQC (optimized for either aromatic or

aliphatic CH moieties) experiments. Experiments optimized for aromatic CH moieties were carried out in 5–10% (v/v) D₂O, while those optimized for aliphatic CH moieties were recorded in 100% (v/v) D₂O. All NOESY experiments were performed with a mixing time of 120 ms. Hydrogen bonds between imino or amino hydrogen groups as donors and nitrogen acceptors were identified experimentally using HNN-COSY, amino HNN-COSY, long-range HNN-COSY and sellr-HNN-COSY experiments using appropriately labeled samples (33–37).

For the determination of the ribose sugar pucker, a TOCSY experiment with a mixing time of 20 ms was recorded on a uniformly 5-D₁/ribose-3',4',5',5''-D₄-labeled sample in 100% (v/v) D₂O. Nucleotides that showed a significant H1'-H2' TOCSY cross peaks were restraint to a C2'-endo conformation.

Structure calculation

All extracted NOE cross-peak intensities were referenced to the averaged intranucleotide H5/H6 NOEs in pyrimidine residues (2.4 Å). For spectra without H5/H6 NOE cross peaks, referencing was performed by matching the peak intensities to the corresponding intensities of identical NOEs in reference spectra. 1.9 Å were added to all upper limit distance restraints. They were classified in three distinct groups corresponding to weak, intermediate and strong NOEs. The geometry of canonical base pairs identified in the HNN-COSY experiments was constrained by introducing four upper and four lower distance restraints for A:U and G:U base pairs as well as for the U16:SAH intermolecular base pair and six upper and six lower distance restraints for G-C base pairs. All additional experimentally detected hydrogen bonds were introduced using two upper and two lower distance restraints. For residues in canonical base pairs (nucleotides 1–5 and 20–24 in P1 as well as 10–14 and 38–42 in the pseudoknot) the backbone torsion angles α , β , γ , δ , ϵ and ζ were set to standard A-form helical values ($\pm 20^\circ$). The ribose conformation of nucleotides 2–4 and 21–23 in P1 and 11–12 and 40–41 in P2 was constrained to C3'-endo. Residues G9, A18 and U12 were constrained to a C2'-endo ribose conformation. The ribose conformations of the flexible linker residues as well as U13 and U37 were left unconstrained. The glycosidic torsion angle χ for the bases in the A-form helical parts was set to restrict the nucleobases in an *anti*-conformation as was the torsion angle χ for A6 due to the characteristic NOE pattern in this region.

The structures were calculated using CYANA version 3.97. PyMol (DeLano Scientific LLC, Schrödinger Inc.) or MOLMOL (38) was used for structure visualization.

Isothermal titration calorimetry (ITC)

All isothermal titration calorimetry (ITC) measurements were performed using a MicroCal iTC₂₀₀ instrument (Malvern Pananalytical, UK) at 25°C. Samples of unlabeled RNAs and all ligands were prepared in 25 mM potassium phosphate buffer (pH 6.2) containing 50 mM potassium chloride and 2 mM magnesium acetate. The ligand (400–4000 μ M) was injected into a solution of 40–100 μ M RNA. After an initial delay of 120 s, the first injection of 0.2 μ l

was followed by 19 serial injections of 2 μ l, separated by intervals of 180 s using a reference power of 11 μ cal⁻¹ and a stirring speed of 750 rpm with high feedback mode. The titration of the ligand into the buffer solution resulted in negligible evolution of heat. Three independent titrations for env9b-derived RNAs and two independent titrations for Pde-1-1 or SK209-52 RNAs were performed for each ligand and the reported K_D values are the average from these titrations. The thermograms were processed using Origin 7.0 (OriginLab) assuming a one-site binding model. The first injection was excluded from the analysis. An overview of all binding parameters derived from the ITC measurements is given in Supplementary Table S1.

Bioinformatics

Additional SAM/SAH riboswitches were found using Infernal version 1.1 (39) to search the RefSeq nucleotide database version 72 (40) and environmental sequences essentially as previously described (39). We analyzed the occurrence of enzymes in different species using KEGG (41), using nine completely sequenced organisms with SAM/SAH riboswitches that are also annotated by KEGG (see Supplementary Table S2). We did not consider organisms with incomplete genomes (i.e. whose RefSeq accessions did not start with 'NC'), since apparently absent genes could be located in the missing parts of the genome.

RESULTS AND DISCUSSION

NMR-spectroscopy of the SAM/SAH-riboswitch

In order to find SAM/SAH riboswitch candidates suitable for high-resolution structure determination we screened sequences from different organisms in NMR-based titration experiments with SAM and/or SAH at different temperatures and Mg²⁺ concentrations. Changes in imino proton 1D ¹H-NMR spectra upon ligand addition as well as signal dispersion and signal line widths served as selection criteria. The spectra of two 43-nt RNA variants (env9b, env9bP1, Figure 1C) derived from a metagenomic sequence in an environmental sample called env9 (21) showed large spectral changes as well as well-dispersed imino proton signals with narrow line widths upon SAH addition (Figure 1D and Supplementary Figure S2). Spectral quality increased even further upon addition of Mg²⁺ (Figure 1D). Thus, this RNA is capable of binding SAH already in the absence of Mg²⁺. The imino proton spectrum of the free RNA showed only a small number of broadened signals suggesting that in the absence of the ligand the RNA is only partially structured and interconverting between different conformations (Figure 1D). This is in agreement with the previous proposal (21) that the ligand-free state of this riboswitch corresponds to the 'ON'-state where the ribosome-binding site is accessible to the ribosome for ligand binding. Addition of SAM lead to spectral changes in the RNA imino proton 1D-¹H spectra very similar to those observed upon SAH binding. Thus, this riboswitch binds indeed to both ligands. Furthermore, SAM and SAH induce the folding of this RNA into very similar structures. This was verified by the comparison of the ¹⁵N-HSQC-spectra of both complexes (Figure 2A). ITC showed that the env9bP1 ri-

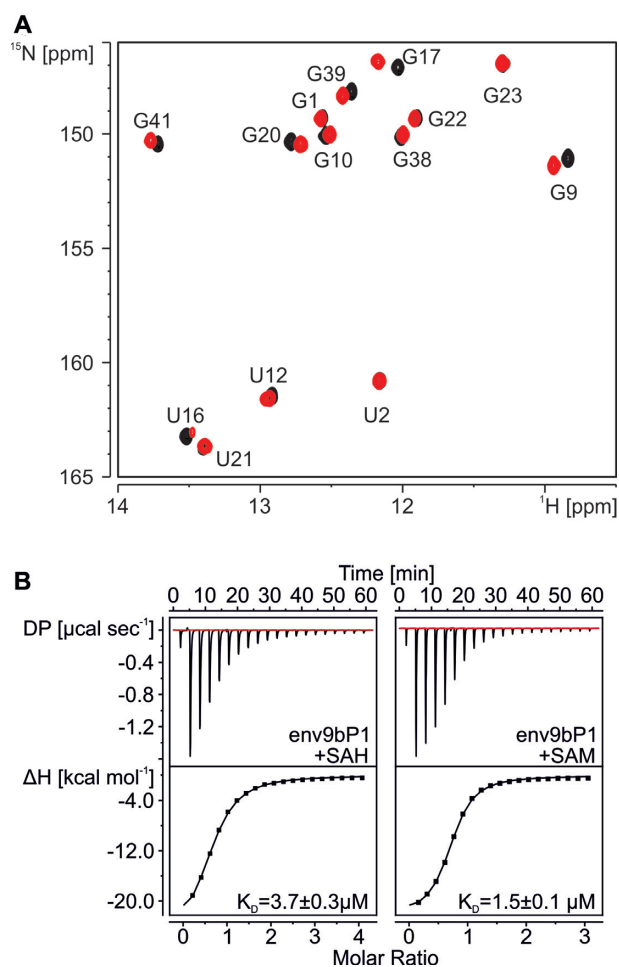


Figure 2. The env9 SAM/SAH-riboswitch binds SAM and SAH with a similar binding mode and with similar affinities. (A) Overlay of the ^{15}N -HSQC spectra of the imino region of uniformly ^{15}N -labeled env9bP1 RNA in complex with SAH (black) and SAM (red) recorded at 20°C in the presence of 2 mM magnesium acetate. NMR signal assignments are indicated. (B) Representative ITC thermograms for the env9bP1 riboswitch titrated with SAH (left) or SAM (right) in the presence of 2 mM magnesium acetate. The resulting K_D values are indicated.

riboswitch bound SAM ($K_D = 1.5 \mu\text{M}$) with a slightly higher affinity than SAH ($K_D = 3.7 \mu\text{M}$) in the presence of 2 mM MgCl_2 (Figure 2B). In the absence of Mg^{2+} the affinity of the RNA for both ligands is significantly diminished (Supplementary Figure S2). Despite its slightly lower affinity, we chose to determine the NMR-solution structure of the SAH-RNA complex in the presence of 2 mM magnesium acetate since NMR-experiments for the collection of structural constraints can require multiple days of measurement time and SAM is unstable at room temperature on this time scale. For NMR-signal assignment purposes and to minimize NMR-signal overlap we used both the env9b and the env9bP1 variant of the riboswitch. Since the slight sequence difference in the P1 stem region (Figure 1C) had no consequences for ligand binding and RNA tertiary structure (Figure 2 and Supplementary Figure S2), spectra recorded for both constructs could be used to derive structural constraints (Supplementary Table S1).

Table 1. Structural statistics for the SAM/SAH-binding riboswitch in complex with SAH (PDB entry: 6HAG)

NMR restraints	
Total distance restraints	743
RNA	
Intra-residue	234
Sequential	184
Long-range	166
Hydrogen-bond	76
RNA-SAH	
NOE	79
Hydrogen-bond	4
Total dihedral restraints	168
Ribose pucker	57
Backbone	111
Structural statistics	
Heavy atom RMSD to mean structure (Å)	0.58 ± 0.15
(U2-U12, C14-G23, A35-U36, G38-C42, SAH)	

Importantly, the quality of the fingerprint ^{15}N -HSQC NMR-spectrum for the SAH-RNA complex was similar to that recorded for the SAM-RNA complex (Figure 2A). Due to significant structural deviations from A-form double-helical RNA, NMR signal assignments had to rely heavily on through-bond magnetization transfer experiments for the identification of base spin systems (42–46) and for connecting base and sugar spin systems (47,48) as well as on the use of a multitude of chemically synthesized site-specifically labeled RNAs (23,25) as described in detail previously (28). In order to collect a large number of NOE-based distance restraints as input for NMR-structure calculations we used base-type selectively labeled samples as well as a sample prepared from partially deuterated nucleotides (5- D_1 /ribose-3',4',5',5''- D_4 -labeled (49)). Furthermore, we employed variants of the HNN-COSY experiments optimized for different types of hydrogen bond donor and acceptor groups on both uniformly and base-type selectively ^{15}N -labeled samples in order to delineate the hydrogen bond network in the RNA-SAH complex (36,37). The final NMR-structure was calculated from 663 NOE-based distance restraints, 168 torsion angle restraints and 80 hydrogen bond restraints. The overall heavy atom RMSD for the 10 structures with the lowest CYANA-target function is $0.58 \pm 0.15 \text{ \AA}$ when including all structurally well-defined nucleotides (Table 1).

Overall structure of the SAM/SAH-riboswitch in the ligand bound state

The structure of the SAM/SAH-riboswitch bound to SAH is shown in Figure 3A. As predicted from sequence comparisons, the RNA in the complex forms an H-type pseudoknot with the P2 stem consisting of base pairs between nucleotides of the apical loop of the 5' hairpin and nucleotides stemming from the conserved region in the 3'-single stranded region (Figures 1C, 3A and B). However, the base pairing interactions in P2 differ from the earlier prediction (21). The previously proposed secondary structure model predicted a continuous P2 helix consisting of nucleotides C11 to C14 (env9b numbering) in the 5'-hairpin loop and nucleotides 38 and 41 from the 3'-single stranded region. In our structure, we observe base pairing between

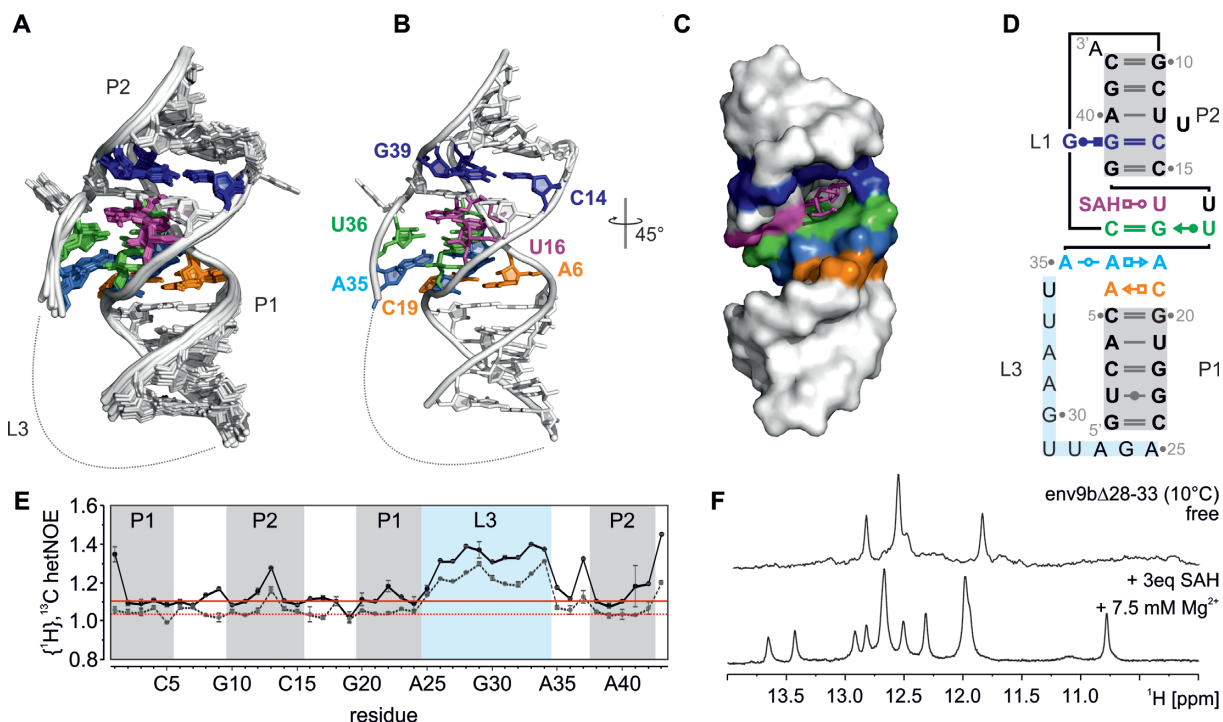


Figure 3. Overall structure of the SAM/SAH-riboswitch in complex with SAH. (A) Superimposition of the ten riboswitch/SAH complex structures with the lowest CYANA target function. The base triplet between the bases G9, C14 and G39 is shown in dark blue, U16 and SAH in magenta, the base triplet C8:G17:U36 in green, the base triplet of A6:A18:A35 in light blue and A6 and C19 in orange. (B) Average structure of the riboswitch/SAH complex with the same coloring as used in (A). (C) Surface representation of the riboswitch RNA and SAH in stick representation with the same coloring as in (A). (D) Structure diagram of the complex structure using the Leontis-Westhof notation. (E) $\{^1\text{H}\}, ^{13}\text{C}$ hetNOE values for the aliphatic H1'/C1' (solid black lines) and aromatic H8/C8- and H6/C6- (dashed gray lines) moieties along the sequence of the riboswitch in the SAH-bound state. The average hetNOE-values of the helical residues for aliphatic H1'/C1' aromatic H6/C6 and H8/C8 groups are marked by a solid and a dashed red line, respectively. The different structural regions of the riboswitch are indicated by different background colors. (F) Imino proton spectra of the mutant with a truncated linker env9b Δ 28-33 in the ligand-free state (top) and in the presence of 3-fold excess of SAH and 7.5 mM magnesium acetate (bottom) at 10°C.

nucleotides G10, C11, U12, C14 and C15 from the hairpin loop and the G38-G39-A40-G41-C42 sequence of the 3'-single stranded region (Figures 1C and 3D). All nucleotides involved in pseudoknot-helix base pairing are strongly conserved (Figure 1B). Nucleotide U13 is looped-out from the pseudoknot helix and is not conserved (Figure 1B). It is conformationally flexible as seen from elevated $\{^1\text{H}\}, ^{13}\text{C}$ -hetNOEs for its H6/C6, H8/C8 and H1'/C1' moieties (Figure 3E) and induces a slight kink in the P2 helix (Figure 3A and C). Replacement of this nucleotide with A or C preserved the ligand binding capabilities of the RNA while its deletion lead to a complete loss of ligand binding (Supplementary Figure S3). Compared to the previously predicted secondary structure the P1 helix of the RNA is extended by two non-Watson-Crick base pairs—A6:C19 and A7:A18—and a Watson-Crick base pair between C8 and G17 (Figure 3). G9 formally corresponds to loop 1 (L1) of the pseudoknot and spans the major groove of the entire pseudoknot P2 helix (Figure 3). U16 on the other side of the hairpin loop represents loop 2 (L2) of the pseudoknot. It is not base paired with an RNA-nucleotide but instead binds the adenosyl moiety of SAH (see below). Formally, loop 3 (L3) of the pseudoknot consists of nucleotides 25–37 and crosses the minor groove of the P1 stem (Figure 3). $\{^1\text{H}\}, ^{13}\text{C}$ -hetNOE-measurements (Figure 3E) show that nucleotides 26–34 of L3 are highly flexible and not involved

in stabilizing interactions with the remainder of the RNA-structure. An RNA with a significantly shortened linker (env9b Δ 28-33) where nucleotides 28–33 were deleted is still able to bind the ligand (Supplementary Figure S4) and folds into a wild-type (WT)-like structure demonstrating that these nucleotides are not important for the overall structure (Figure 3F). Nucleotide 35 of the linker is strictly conserved as an A and is conformationally rigid (Figure 3E). It stabilizes the non-canonical P1 base pair A7:A18 by forming a minor groove base triple (see below). Nucleotide U36 stabilizes the closing base pair of the P1 stem C8:G17 by forming another minor groove base triple (see below). Nucleotide U37 is flexible according to the $\{^1\text{H}\}, ^{13}\text{C}$ -HetNOE-data (Figure 3E) and does not interact with other parts of the RNA. It serves as a spacer for creating the SAM-binding pocket. Overall, the structure of the SAM/SAH riboswitch in complex with its ligand is significantly less compact compared to the structures of other riboswitches adopting similar H-type pseudoknot structures upon ligand binding. In the SAM-II riboswitch, the tertiary interactions between linker nucleotides in L1 (7 nt) and L3 (11 nt) and the remainder of the RNA are much more pronounced (50). Nucleotides in L1 are an integral part of the ligand binding site whereas almost all nucleotides in L3 form stabilizing tertiary interactions with the major groove of the P1 stem. Only one nucleotide in L1 and three nucleotides in L3 are

flexible in this structure. Similarly, in the H-type pseudoknot structure adopted by the preQ1 riboswitch, all nucleotides of L1 (2 nt) and L3 (6nt) are involved in tertiary interactions (51).

Ligand recognition and the ligand binding site

The ligand binding site of the SAM/SAH riboswitch is sandwiched between the C8:G17 closing base pair of the P1 stem and the C15:G38 closing base pair of the pseudoknot P2 stem (Figure 3). The adenosyl moiety of the ligand forms an intermolecular base pair with the strictly conserved U16 nucleotide of the riboswitch. An HNN-COSY spectrum recorded for a selectively ^{15}N -U-labeled RNA bound to $^{13}\text{C}^{15}\text{N}$ -labeled SAH showed directly that the U16 imino group is hydrogen bonded to the N7 of the ligand (Figure 4A) and together with intermolecular NOEs established a *trans* Hoogsteen/Watson–Crick (‘reversed Hoogsteen’) geometry for the SAH:U16 intermolecular base pair (Figure 4B). A U16A mutant of the riboswitch is not capable of binding SAH anymore (Supplementary Figure S5). Unambiguous NMR resonance assignments of the U16 spin system were further supported by the use of a chemically synthesized U16 N3- ^{15}N and a U16 C6- ^{13}C selectively labeled sample (Supplementary Figure S5). A *trans* Hoogsteen/Watson–Crick base pairing interaction between the adenosyl moiety of the ligand and an U residue of the riboswitch structurally similar to the one observed here is also found in the SAM-II riboswitch (50). However, in this case the U residue binding the ligand base moiety is further stabilized by additional hydrogen bonding interactions with another U residue of the riboswitch. In contrast, in the SAM-I riboswitch the adenine base of SAM forms a *cis* Hoogsteen/Watson–Crick base pair with an U of the riboswitch (16).

The formation of an intermolecular base pair between U16 and SAH establishes continuous stacking between the P1 and the P2 stems of the pseudoknot (see Figure 3). The adenosyl moiety of SAH stacks with the bases of G38 from the C15:G38 base pair closing P2 on one site and of G17 from the C8:G17 closing base pair of P1 on the other side (Figure 3). The sulfur atom from SAH is in an appropriate distance ($3.11 \pm 0.03 \text{ \AA}$) and orientation to form a chalcogen bond (52) with the O6 carbonyl oxygen of G38 (Figure 4C). Notably, in a SAM-bound complex the sulfur atom would carry an additional positive charge strengthening this interaction due to the negative partial charge of G38 O6. The ribose hydroxyl groups of SAH also interact with the RNA. The 2'-OH group of the ligand is in hydrogen bonding distance to the 5'-phosphate group of G38 (average O–O distance for 10 best structures $3.1 \pm 0.4 \text{ \AA}$) (Figure 4C). The 3'-OH group of SAH is in hydrogen bonding distance to the O4' of the ribose ring in G9 as well as to the 2'-OH-group of C8 in P1 ($3.0 \pm 0.2 \text{ \AA}$ and $3.2 \pm 0.1 \text{ \AA}$).

The C8:G17 base pair at the lower base of the ligand binding site is further stabilized through the formation of a minor groove base triple with U36 from L1 (Figure 4D). The U36 C2 carbonyl group is a hydrogen bond acceptor for the G17 amino group (Figure 4D). The 2'-OH group of U36 acts as a hydrogen bond acceptor for the amino group of SAH (Figure 4E). A U36C mutant RNA that keeps these

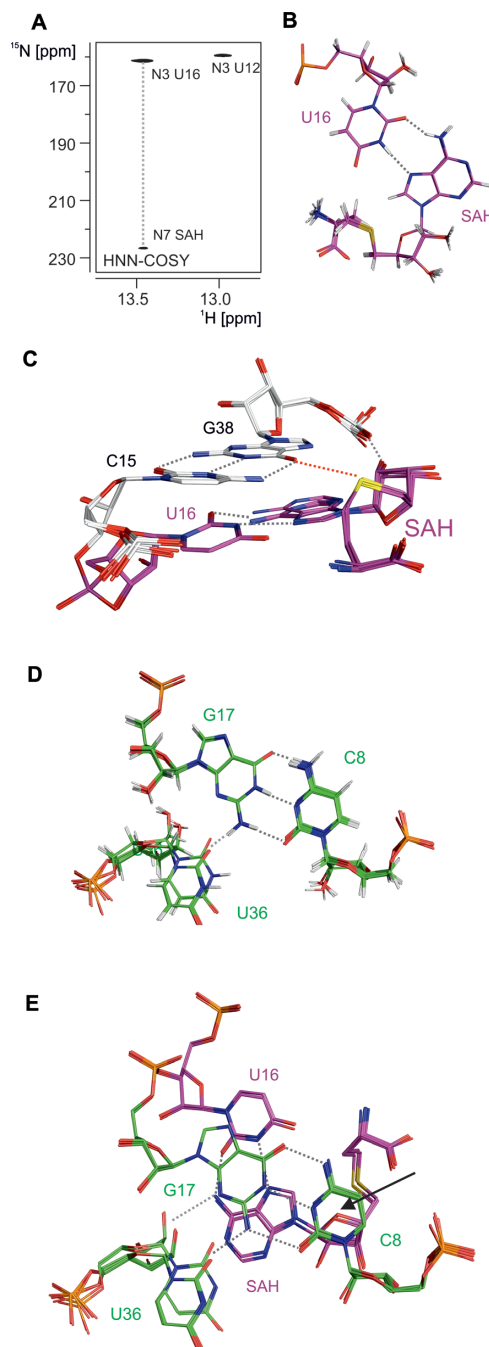


Figure 4. The SAH binding site. (A) HNN-COSY experiment on a ^{15}N -U-labeled RNA in complex with $^{13}\text{C}^{15}\text{N}$ -labeled SAH showing a correlation between the U16 imino group and the N7 nitrogen of SAH. (B) SAH recognition in a ‘reversed Hoogsteen’ (*trans* Hoogsteen/Watson–Crick) base pair between SAH and U16 in a bundle representation. Hydrogen bonds are indicated by dashed lines. (C) Interactions between the sulfur atom and the ribose moiety of SAH with the C15:G38 base pair forming the ‘roof’ of the SAH-binding site. Dashed gray lines indicate hydrogen bonds and a dashed red line indicates the putative chalcogen bond between U16 and SAH. (D) The C8:G17:U36 minor groove base triplet (carbon atoms in green) forming the floor of the SAH binding pocket in a bundle representation. (E) Stacking interactions between the U16:SAH intermolecular base pair (carbon atoms in magenta) and the C8:G17:U36 (carbon atoms in green) minor groove base triplet. Hydrogen bonds are indicated by dashed lines. A black arrow points toward the O4' oxygen of the SAH ribose group located atop the base of C8 in an orientation suitable for a lone pair– π interaction.

hydrogen bonding interactions intact binds SAM and SAH with the same affinity as the WT (Supplementary Figure S6 and Table S1). Furthermore, the base of C8 is packed against the ribose moiety of SAH (Figure 4E). The SAH ribose O4' oxygen is in an orientation suitable for a lone pair- π interaction with the C8 base (3.34 ± 0.08 Å) (53,54). At last, the positively charged amino group of the aminocarboxypyl group of SAH is in a distance range (4.2 ± 0.1 Å) for an electrostatic interaction with O4 of U16 (Figure 4E).

A major groove base triple between L1 and helix P2 of the pseudoknot

In the immediate vicinity of the binding pocket, the structure of the SAH:RNA complex is stabilized by additional tertiary interactions (Figure 5A). Directly above the 'roof' of the SAH binding pocket formed by C15:G38, the closing base pair of P2, there is a major groove base triple consisting of the C14:G39 Watson-Crick base pair and G9, which formally represents L1 of the pseudoknot (Figure 5B). The amino group of G9 is hydrogen bonded to the N7 of G39 and the G9 imino group is hydrogen bonded to the O6 of G39 establishing a *cis* Watson-Crick/Hoogsteen base pairing interaction between G9 and G39. In addition, the G9 O6 serves as a hydrogen bond acceptor for the C14 amino group (Figure 5B). The geometry of this base triple was established based on a large number of NOE contacts involving the G9 imino group (Figure 5C). Its proton chemical shift of 10.8 ppm is in agreement with an oxygen containing group as a hydrogen bond acceptor. An RNA with a G39C/C14G mutation that should be able to form a stable P2 Watson-Crick base pair but no hydrogen bonds with G9 is not capable of ligand binding (Figure 5D).

Stabilizing tertiary interactions between residues of L3 and the extended P1 helix

Below the 'floor' of the SAH-binding pocket, which consists of the minor groove base triple C8:G17:U36 (see above), lies another base triple (Figure 6A). It involves base pairing between A7 and A18 as well as between A7 and A35. A7 and A18 are part of the extended stem P1 whereas A35 stems from the L3 loop. A7 and A18 interact with each other in a *trans* Hoogsteen/Sugar-edge base pairing interaction stabilized via a single hydrogen bond between the A7 amino group and the N3 nitrogen of A18 (Figure 6B). Two hydrogen bonds are present between the amino groups of A7 and A35 and their respective N1 nitrogen atoms establishing a *trans* Watson-Crick/Watson-Crick base pair (Figure 6B). The geometry of this base triplet is supported by a dense network of NOE-constraints (Figure 6C). The *trans* Hoogsteen/Sugar-edge base pair A7:A18 is isosteric to the classical 'sheared' G:A base pair (a *trans* Sugar-edge/Hoogsteen G:A base pair in the Westhof-Leontis nomenclature, Figure 6D). Interestingly, the consensus sequence of the SAM/SAH-riboswitch shows that position 18 can be either A or G whereas in position 7 A is strictly conserved (Figure 1B). A 'sheared' A7:G18 in this position would be stabilized by two hydrogen bonds—one between the A N6 amino group and the G N3 nitrogen and the other

between the G N2 amino group and the A N7 nitrogen (Figure 6D). In line with this idea, an RNA containing an A18G mutation not only bound SAM and SAH with a structure very similar to the WT (Figure 6D) but its affinity for both SAM and SAH was increased compared to the WT (Figure 6E). The ^{15}N -HSQC spectrum of this mutant reveals an additional G imino group signal with a chemical shift of 9.3 ppm (Figure 6D) which is in line with the presence of a 'sheared' A:G base pair where the G imino group has no direct hydrogen bond acceptor (55).

Right underneath this base triple A6 and C19 form a base pair with a very unusual geometry (Figure 7A). Here, the amino group of C19 is involved in two hydrogen bonds with the N3 nitrogen and the 2'-OH group of A6 as acceptors (Figure 7B). The peculiar geometry of this base pair is supported by a cross-hydrogen bond scalar coupling between the C19 amino group resonances and the A6 N3 nitrogen observed in an amino group optimized HNN-COSY experiment (Figure 7C). Furthermore, a very strong NOE cross peak between the H5 resonance of C15 and the H2 resonance of A6 unequivocally confirms the conformation of this base pair (Figure 7D). Additional NMR data illustrating the signal assignment process for the A6 and C19 spin systems as well as additional evidence for the A6:C19 base pairing interaction are shown in the Supplementary material (Supplementary Figure S7). Both A6 and C19 are strictly conserved in the consensus sequence of the SAM/SAH riboswitches suggesting an important functional role of this unique A:C mismatch. However, the A6:C19 base pair is not directly involved in ligand binding or additional tertiary interactions (Figure 3). We tested the functional importance of its unusual geometry by replacing A6 with G thereby enforcing the formation of a normal stable G6:C19 Watson-Crick base pair at this position. The mutant RNA binds SAM with a significantly reduced affinity as reflected by the 27-fold higher K_D compared to the WT (Figure 7E). Apparently, one important structural consequence of the presence of the adjacent A6:C19 and A7:A18 base pairs in the extension of the P1 stem is a significant deviation from the typical A-form helical intrastrand base stacking interactions. In particular, the P1 helix is strongly underwound at the A6/A7 helical step. This departure from standard A-form geometry, however, enables the placement of C8 at the floor of the ligand binding pocket.

Binding of SAM/SAH analogs to the riboswitch

Overall, our structure of the SAM/SAH riboswitch RNA bound to SAH establishes a mode of ligand recognition, which relies on a more limited number of riboswitch-ligand interactions compared to the highly specific SAM riboswitches, and explains why both SAM and SAH bind with rather similar affinities. To gain further insights into ligand recognition and the importance of different functional groups in the ligand we tested different ligand analogs for binding. To assess the contribution of the aminocarboxypyl side chain of SAH to riboswitch binding, we measured the affinity of 5'-deoxy-5'-methylthioadenosine (Supplementary Figure S8 and Table S1). 5'-deoxy-5'-methylthioadenosine bound the riboswitch with a 6-fold reduced affinity compared to SAH in line with the loss of

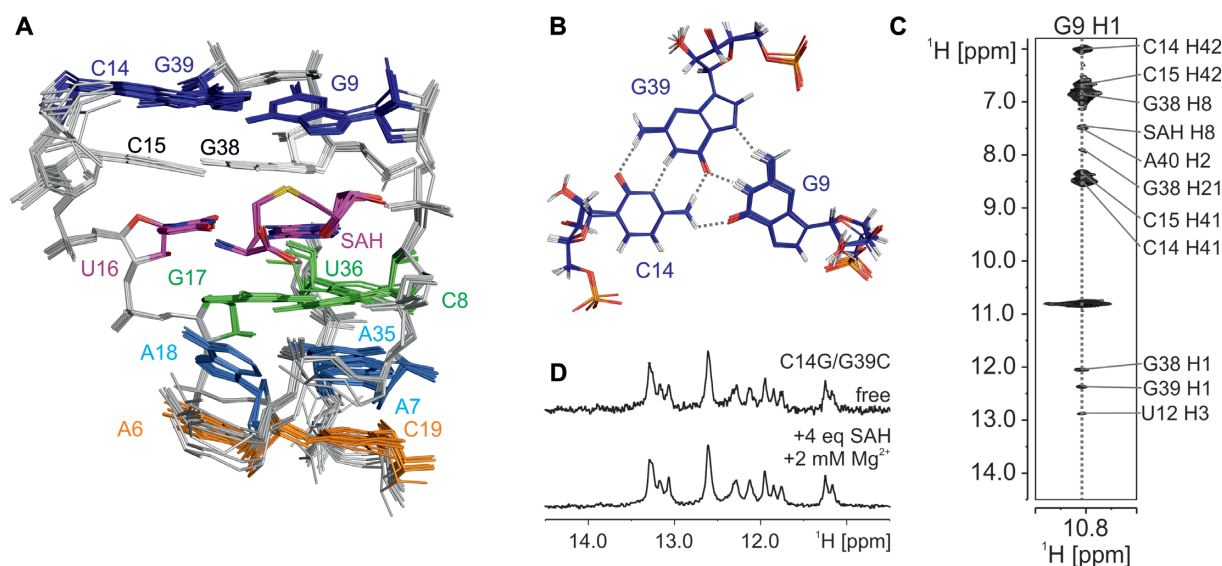


Figure 5. Additional tertiary interactions in the vicinity of the ligand binding pocket. (A) Tertiary structure elements surrounding the intermolecular U16:SAH base pair (magenta). The Watson–Crick base pair C15:G38 at the base of stem P2 is shown in white and the base triplet C8:G17:U36 closing P1 is shown in green. Directly underneath this base triplet is the A7:A18:A35 base triplet (light blue) and the A6:C29 base pair (orange). A major groove base triplet G9:C14:G39 (dark blue) is located directly above the P2 closing base pair C15:G38 (white). (B) The G9:C14:G39 base triplet (dark blue) in a bundle representation. (C) Strip from the ¹H, ¹H-NOESY at the H1 G9 resonance showing NOE contacts to neighboring nucleotides. (D) Imino proton spectra of the mutant env9bP1C14G/G39C without (top) and with 1.5-fold excess of SAH and 2 mM magnesium acetate (bottom). In the mutant with the inverted G:C base pairing orientation the base triplet with G9 cannot form and ligand binding is abrogated.

the electrostatic interaction of the aminocarboxypropyl side chain amino group with U16 O4. Adenosine, where the sulfur atom at the 5'-position involved in a chalcogen bond with O6 of G38 is replaced by oxygen, has a 4-fold weaker affinity compared to 5'-deoxy-5'-methylthioadenosine and a 22 times weaker affinity compared to SAH. Importantly, adenosine monophosphate (AMP), which is abundant in the cell, does not bind the riboswitch (Supplementary Figure S8 and Table S1). Surprisingly, the analogs of SAH and SAM that only lack the carboxyl group in the amino acid side chain—decarboxy-SAH and decarboxy-SAM—bind with better affinity compared to SAH and SAM, respectively (Supplementary Figure S8 and Table S1). This is particularly pronounced for decarboxy-SAM with a K_D of 0.06 μ M for the env9b riboswitch. This K_D is 25-fold lower compared to the K_D for SAM. In contrast, other SAM riboswitches strongly prefer binding to SAM over decarboxy-SAM binding (56,57). Structurally, the preference of the SAM/SAH riboswitch for the decarboxy-analogs of SAH and SAM can be easily rationalized. In the structure of the SAH–RNA complex, the carboxy group of the ligand points into the direction of the negatively charged phosphate backbone, which is energetically unfavorable. The 1D imino proton spectrum of the RNA bound to decarboxy-SAM resembles that of the SAH- and SAM-bound RNA demonstrating that the overall structure of the complexes is highly similar (Supplementary Figure S8). In order to investigate if this preference for decarboxy-SAM over SAM and SAH is a more general feature of the SAM/SAH-riboswitches we measured the affinities for the three ligands for SAM/SAH-riboswitches from *Paracoccus denitrificans* (Pde-1-1) and *Roseobacter* sp. SK209-2-6 (SK209-52-II) (21). The secondary structures for these two SAM/SAH

riboswitch variants are shown in Supplementary Figure S9. Both riboswitches bound to the three ligands with the same order of affinities as the env9b riboswitch demonstrating that the preference for decarboxy-SAM over SAM is maybe a general feature of the SAM/SAH-riboswitches (Supplementary Figure S9). However, their affinity for decarboxy-SAM is only ~3-fold (Pde-1-1) or ~2-fold higher (SK209-52-II) than their affinity for SAM and not 25-fold higher as in the env9b riboswitch (Supplementary Table S1).

The biologically relevant ligand of the SAM/SAH-riboswitch

Interestingly, decarboxy-SAM is a naturally occurring metabolite in bacterial cells used in the biosynthesis of spermine, spermidine and other polyamines. A general assumption in the riboswitch field is that the cognate ligand of the riboswitch has the highest affinity for the RNA. Accordingly, the higher affinity of decarboxy-SAM compared to SAM for the riboswitch would suggest that decarboxy-SAM is the 'real' ligand of the SAM/SAH-riboswitch and not SAM or SAH. A natural target gene for regulation by a decarboxy-SAM-binding 'OFF'-riboswitch would be SAM-decarboxylase. Thus, we conducted an extended bioinformatics search for additional occurrences of the SAM/SAH-riboswitch that might also reveal novel gene associations. While we found 552 novel sequences for this riboswitch compared to version 14.0 of the Rfam database (58), they were almost exclusively associated with the SAM-synthetase gene *metK* and never with SAM-decarboxylase (E.C. 4.1.1.50). Six of nine of the organisms with fully sequenced genomes containing the SAM/SAH riboswitches (Supplementary Table S2) do not even contain genes for SAM-decarboxylase and might therefore not

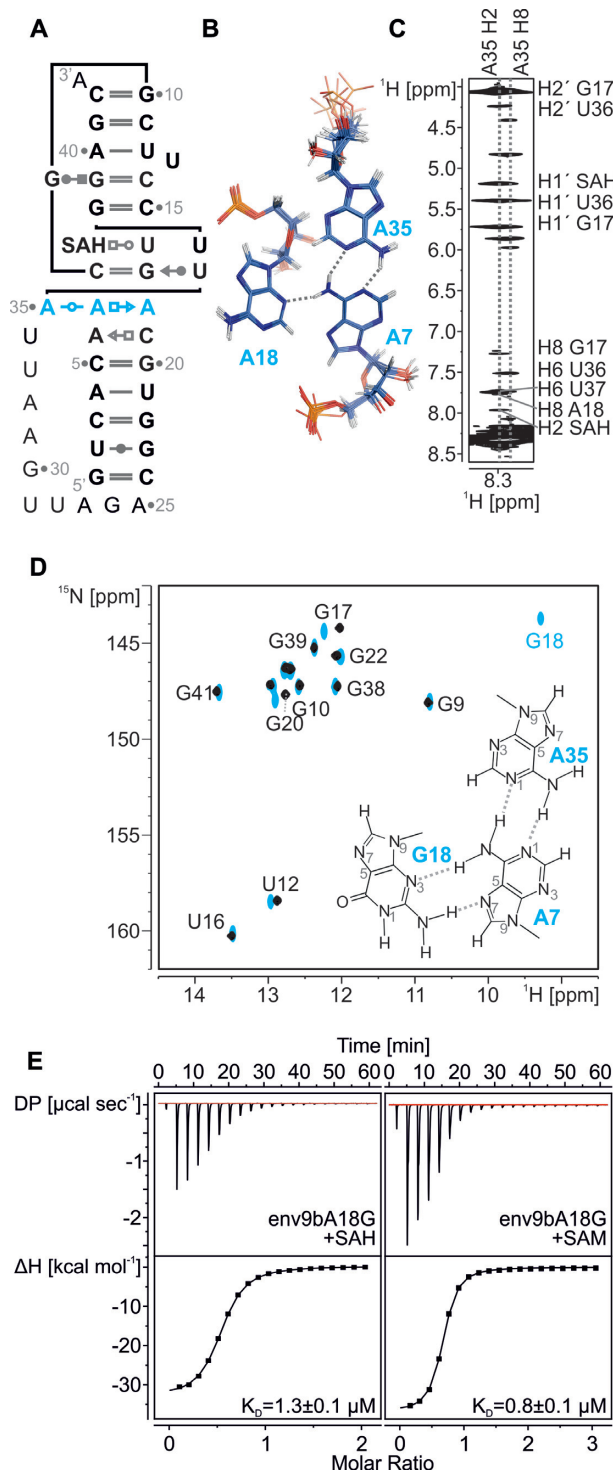


Figure 6. The adenine base triplet A7:A18:A35. (A) Structure diagram of the complex structure using the Leontis–Westhof notation. The base triplet A7:A18:A35 is colored in light blue. (B) The A7:A18:A35 base triplet in a bundle representation. (C) Strip from the ^1H , ^1H -NOESY-spectrum at the H2/H8 A35 resonances showing diagnostic NOE contacts between the A35 base protons and neighboring nucleotides. Signal assignments are indicated. (D) Overlay of the ^{15}N -HSQC spectra of the imino region of uniformly ^{15}N -labeled env9b RNA (black) and env9bA18G (light blue) in complex with SAH recorded at 20°C. The inset shows the geometry of the base triplet upon A18 replacement with G. (E) Representative ITC thermograms for env9bA18G titrated with SAH (left) or SAM (right). The resulting K_D values are indicated.

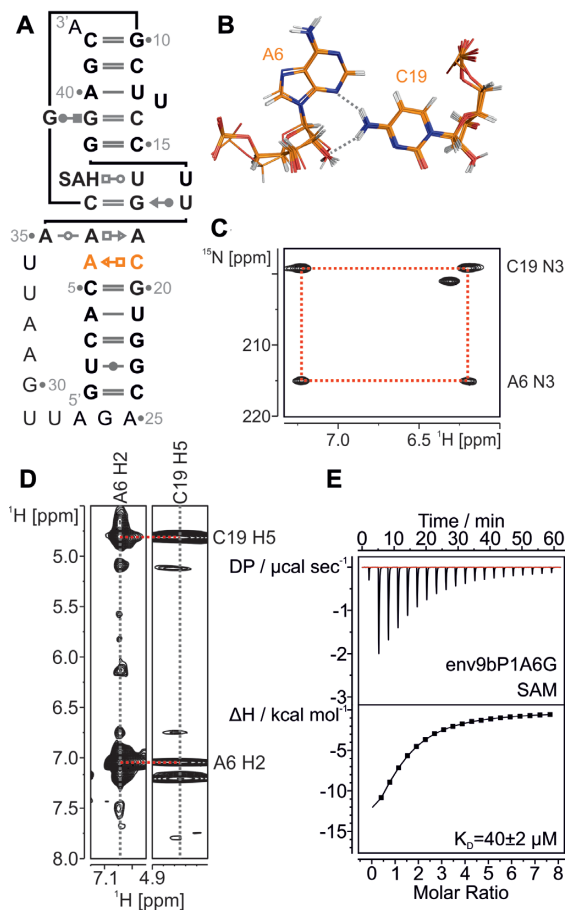


Figure 7. The non-canonical A6:C19 base pair. (A) Structure diagram of the complex structure using the Leontis–Westhof notation. The non-canonical base pair A6:C19 is colored in orange. (B) The A6:C19 base pair in a bundle representation. (C) HNN-COSY experiment for the identification of the hydrogen bond between the N4 amino group of C19 and the N3 nitrogen of A6. In this experiment the amino group protons of C19 simultaneously show cross peaks to the N3 nitrogen of C19 due to the intranucleotide through-bond $^2J_{\text{N4(C)N3(C)}}$ scalar coupling as well as to the A6 N3 nitrogen due to the internucleotide cross-hydrogen-bond $^2hJ_{\text{N4(C)N3(A)}}$ scalar coupling. (D) Strips from the ^1H , ^{13}C -NOESY-HSQC at the H2 A6 and H5 C19 resonance frequencies showing the strong NOEs between the H5 proton of C19 and the H2 proton of A6 supporting the unusual geometry of this A:C base pair. (E) Representative ITC thermogram for the titration of the env9bP1A6G riboswitch mutant with SAM. In this mutant a Watson–Crick G6:C19 base pair would replace the A:C base pair of the WT. The resulting K_D value is indicated.

even produce decarboxy-SAM. Overall, this argues against a role of decarboxy-SAM as the native cognate ligand for the SAM/SAH-riboswitch. Thus, the enhanced affinity for decarboxy-SAM compared to SAM and SAH might just be a side effect of the low level of structural complexity in the ligand binding site of the SAM/SAH riboswitch and the lack of ‘sophistication’ in the ligand recognition mode compared to the SAM-specific riboswitches. A similar argument can be made with regard to the question if SAM or SAH is the biologically relevant ligand for the SAM/SAH riboswitch. Our structure of the SAM/SAH riboswitch shows that there are fewer recognition elements that are able to distinguish a positively charged sulfonium

group in SAM from a neutral sulfur atom in SAH compared to the SAM-specific riboswitches. Furthermore, our structure clearly demonstrates that the SAM/SAH riboswitch is a translational 'OFF' switch since ligand binding helps to sequester the ribosome binding site. The dominating gene association for the SAM/SAH riboswitch is the SAM synthase gene *metK*. However, a shutdown in SAM-synthase expression induced by high concentrations of SAH signaling conditions where SAM consumption is apparently high seems to defy the logic of metabolic regulation. On the other hand, of the nine genome sequences from organisms with a SAM/SAH riboswitch that were practical to analyze (see 'Materials and Methods' section) only one apparently has no easily recognizable gene encoding an SAH hydrolase (Supplementary Table S2). Moreover, this organism, *Ketogulonicigenium vulgare* WSH-001, is closely related to *K. vulgare* Y25, which has a SAM/SAH-riboswitch as well as a SAH hydrolase. This situation suggests the possibility of a recent loss of SAH hydrolase in *K. vulgare* WSH-001. Thus, one might assume that in almost all organisms carrying the SAM/SAH riboswitch, the intracellular SAM levels are significantly higher than the SAH concentrations. Therefore, we tend to favor a scenario where SAM is the cognate, biologically important ligand for the SAM/SAH riboswitch. Its reduced structural complexity compared to the other SAM-specific riboswitches might thus represent a minimal solution for a SAM-based regulatory mechanism that would only work properly in a background where an SAH degradation system is already established. Moreover, all known examples for SAM/SAH-riboswitches are restricted to bacteria from the order Rhodobacterales. Taken together, this could imply that in contrast to other riboswitches the SAM/SAH riboswitch appeared late during bacterial evolution and is not a remnant of the RNA-world.

DATA AVAILABILITY

Atomic coordinates and structure factors for the reported solution NMR structures were deposited with the Protein Data bank under accession number 6HAG. The chemical shift assignments for the SAH/SAM-riboswitch bound to SAH were submitted to the BioMagResBank (BMRB) under accession number 27452.

SUPPLEMENTARY DATA

[Supplementary Data](#) are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Christian Richter and Manfred Strupf for maintenance of the NMR facility and to Kerstin Yacoub for help with sample preparation. Riboswitch searches used computer time provided by the Center for Information Services and High Performance Computing (ZIH) at TU Dresden.

FUNDING

Center for Biomolecular Magnetic Resonance (BMRZ) of the Goethe University Frankfurt; Deutsche Forschungsgemeinschaft (DFG) [CRC 902 'Molecular principles of

RNA-based regulation' B10]. Funding for open access charge: DFG.

Conflict of interest statement. None declared.

REFERENCES

- Fontecave, M., Atta, M. and Mulliez, E. (2004) S-adenosylmethionine. Nothing goes to waste. *Trends Biochem. Sci.*, **29**, 243–249.
- Struck, A.-W., Thompson, M.L., Wong, L.S. and Micklefield, J. (2012) S-adenosyl-methionine-dependent methyltransferases. Highly versatile enzymes in biocatalysis, biosynthesis and other biotechnological applications. *ChemBiochem.*, **13**, 2642–2655.
- Lee, J.E., Cornell, K.A., Riscoe, M.K. and Howell, P.L. (2003) Structure of *Escherichia coli* 5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase inhibitor complexes provide insight into the conformational changes required for substrate binding and catalysis. *J. Biol. Chem.*, **278**, 8761–8770.
- Parveen, N. and Cornell, K.A. (2011) Methylthioadenosine/S-adenosylhomocysteine nucleosidase, a crucial enzyme for bacterial metabolism. *Mol. Microbiol.*, **79**, 7–20.
- Kusakabe, Y., Ishihara, M., Umeda, T., Kuroda, D., Nakanishi, M., Kitade, Y., Gouda, H., Nakamura, K.T. and Tanaka, N. (2015) Structural insights into the reaction mechanism of S-adenosyl-L-homocysteine hydrolase. *Sci. Rep.*, **5**, 16641.
- Turner, M.A., Yang, X., Yin, D., Kuczera, K., Borchardt, R.T. and Howell, P.L. (2000) Structure and function of S-adenosylhomocysteine hydrolase. *Cell. Biochem. Biophys.*, **33**, 101–125.
- Meyer, B., Wurm, J.P., Sharma, S., Immer, C., Pogoryelov, D., Kötter, P., Lafontaine, D.L.J., Wöhrnt, J. and Entian, K.-D. (2016) Ribosome biogenesis factor Tsr3 is the aminocarboxypropyl transferase responsible for 18S rRNA hypermodification in yeast and humans. *Nucleic Acids Res.*, **44**, 4304–4316.
- Umitsu, M., Nishimasu, H., Noma, A., Suzuki, T., Ishitani, R. and Nureki, O. (2009) Structural basis of AdoMet-dependent aminocarboxypropyl transfer reaction catalyzed by tRNA-wybutosine synthesizing enzyme, TYW2. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 15616–15621.
- Pegg, A.E., Xiong, H., Feith, D.J. and Shantz, L.M. (1998) S-adenosylmethionine decarboxylase. Structure, function and regulation by polyamines. *Biochem. Soc. Trans.*, **26**, 580–586.
- Bale, S., Baba, K., McCloskey, D.E., Pegg, A.E. and Ealick, S.E. (2010) Complexes of *Thermotoga maritima* S-adenosylmethionine decarboxylase provide insights into substrate specificity. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 181–189.
- Tabor, C.W. and Tabor, H. (1984) Polyamines. *Ann. Rev. Biochem.*, **53**, 749–790.
- O'Hagan, D., Schaffrath, C., Cobb, S.L., Hamilton, J.T.G. and Murphy, C.D. (2002) Biochemistry: biosynthesis of an organofluorine molecule. *Nature*, **416**, 279.
- Schaffrath, C., Deng, H. and O'Hagan, D. (2003) Isolation and characterisation of 5'-fluorodeoxyadenosine synthase, a fluorination enzyme from *Streptomyces cattleya*. *FEBS Lett.*, **547**, 111–114.
- Holloway, C.T., Greene, R.C. and Su, C.H. (1970) Regulation of S-adenosylmethionine synthetase in *Escherichia coli*. *J. Bacteriol.*, **104**, 734–747.
- Wang, J.X. and Breaker, R.R. (2008) Riboswitches that sense S-adenosylmethionine and S-adenosylhomocysteine. *Biochem. Cell Biol.*, **86**, 157–168.
- Batey, R.T. (2011) Recognition of S-adenosylmethionine by riboswitches. *Wiley Interdiscip. Rev. RNA*, **2**, 299–311.
- Tucker, B.J. and Breaker, R.R. (2005) Riboswitches as versatile gene control elements. *Curr. Opin. Struct. Biol.*, **15**, 342–348.
- Gilbert, S.D., Montange, R.K., Stoddard, C.D. and Batey, R.T. (2006) Structural studies of the purine and SAM binding riboswitches. *Cold Spring Harb. Symp. Quant. Biol.*, **71**, 259–268.
- Gayán, M.A., Sherlock, M.E., Weinberg, Z. and Breaker, R.R. (2018) SAM-VI RNAs selectively bind S-adenosylmethionine and exhibit similarities to SAM-III riboswitches. *RNA Biol.*, **4**, 371–378.
- Wang, J.X., Lee, E.R., Morales, D.R., Lim, J. and Breaker, R.R. (2008) Riboswitches that sense S-adenosylhomocysteine and activate genes involved in coenzyme recycling. *Mol. Cell*, **29**, 691–702.

21. Weinberg,Z., Wang,J.X., Bogue,J., Yang,J., Corbino,K., Moy,R.H. and Breaker,R.R. (2010) Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol.*, **11**, R31.
22. Duchardt-Ferner,E., Weigand,J.E., Ohlenschläger,O., Schmidtke,S.R., Suess,B. and Wöhnert,J. (2010) Highly modular structure and ligand binding by conformational capture in a minimalistic riboswitch. *Angew. Chem. Int. Ed. Engl.*, **49**, 6216–6219.
23. Wunderlich,C.H., Spitzer,R., Santner,T., Fauster,K., Tollinger,M. and Kreutz,C. (2012) Synthesis of (6-(13)C)pyrimidine nucleotides as spin-labels for RNA dynamics. *J. Am. Chem. Soc.*, **134**, 7558–7569.
24. Alvarado,L.J., Longhini,A.P., LeBlanc,R.M., Chen,B., Kreutz,C. and Dayie,T.K. (2014) Chemo-enzymatic synthesis of selectively ¹³C/¹⁵N-labeled RNA for NMR structural and dynamics studies. *Methods Enzymol.*, **549**, 133–162.
25. Juen,M.A., Wunderlich,C.H., Nußbaumer,F., Tollinger,M., Kontaxis,G., Konrat,R., Hansen,D.F. and Kreutz,C. (2016) Excited states of nucleic acids probed by proton relaxation dispersion NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.*, **55**, 12008–12012.
26. Ottink,O.M., Nelissen,F.H.T., Derks,Y., Wijmenga,S.S. and Heus,H.A. (2010) Enzymatic stereospecific preparation of fluorescent S-adenosyl-L-methionine analogs. *Anal. Biochem.*, **396**, 280–283.
27. Chatterjee,D., Kudlinzki,D., Linhard,V., Saxena,K., Schieborr,U., Gande,S.L., Wurm,J.P., Wöhnert,J., Abele,R., Rogov,V.V. *et al.* (2015) Structure and biophysical characterization of the S-adenosylmethionine-dependent O-methyltransferase PAMTH1, a putative enzyme accumulating during senescence of *Podospora anserina*. *J. Biol. Chem.*, **290**, 16415–16430.
28. Weickmann,A.K., Keller,H., Duchardt-Ferner,E., Strebitzer,E., Juen,M.A., Kremser,J., Wurm,J.P., Kreutz,C. and Wöhnert,J. (2018) NMR resonance assignments for the SAM/SAH-binding riboswitch RNA bound to S-adenosylhomocysteine. *Biomol. NMR Assign.*, **12**, 329–334.
29. Keller,R. (2004) *The Computer Aided Resonance Assignment Tutorial*. CANTINA, Goldau.
30. Wishart,D.S., Bigam,C.G., Yao,J., Abildgaard,F., Dyson,H.J., Oldfield,E., Markley,J.L. and Sykes,B.D. (1995) ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR*, **6**, 135–140.
31. Mueller,L., Legault,P. and Pardi,A. (1995) Improved RNA structure determination by detection of NOE contacts to exchange-broadened amino protons. *J. Am. Chem. Soc.*, **117**, 11043–11048.
32. Grzesiek,S. and Bax,A. (1993) The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc.*, **115**, 12593–12594.
33. Dingley,A.J. and Grzesiek,S. (1998) Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide ²J_{NN} couplings. *J. Am. Chem. Soc.*, **120**, 8293–8297.
34. Wöhnert,J., Dingley,A.J., Stoldt,M., Gorchach,M., Grzesiek,S. and Brown,L.R. (1999) Direct identification of NH...N hydrogen bonds in non-canonical base pairs of RNA by NMR spectroscopy. *Nucleic Acids Res.*, **27**, 3104–3110.
35. Majumdar,A., Kettani,A. and Skripkin,E. (1999) Observation and measurement of internucleotide ²J_{NN} coupling constants between ¹⁵N nuclei with widely separated chemical shifts. *J. Biomol. NMR*, **14**, 67–70.
36. Hennig,M. and Williamson,J.R. (2000) Detection of N-H...N hydrogen bonding in RNA via scalar couplings in the absence of observable imino proton resonances. *Nucleic Acids Res.*, **28**, 1585–1593.
37. Dallmann,A., Simon,B., Duszczuk,M.M., Kooshapur,H., Pardi,A., Bermel,W. and Sattler,M. (2013) Efficient detection of hydrogen bonds in dynamic regions of RNA by sensitivity-optimized NMR pulse sequences. *Angew. Chem. Int. Ed. Engl.*, **52**, 10487–10490.
38. Konradi,R., Billeter,M. and Wüthrich,K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 29–32.
39. Weinberg,Z., Lünse,C.E., Corbino,K.A., Ames,T.D., Nelson,J.W., Roth,A., Perkins,K.R., Sherlock,M.E. and Breaker,R.R. (2017) Detection of 224 candidate structured RNAs by comparative analysis of specific subsets of intergenic regions. *Nucleic Acids Res.*, **45**, 10811–10823.
40. O’Leary,N.A., Wright,M.W., Brister,J.R., Ciufio,S., Haddad,D., McVeigh,R., Rajput,B., Robbertse,B., Smith-White,B., Ako-Adjei,D. *et al.* (2016) Reference sequence (RefSeq) database at NCBI. Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.*, **44**, D733–D745.
41. Kanehisa,M., Furumichi,M., Tanabe,M., Sato,Y. and Morishima,K. (2017) KEGG. New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.*, **45**, D353–D361.
42. Sklenár,V., Peterson,R.D., Rejante,M.R. and Feigon,J. (1994) Correlation of nucleotide base and sugar protons in a ¹⁵N-labeled HIV-1 RNA oligonucleotide by ¹H-¹⁵N HSQC experiments. *J. Biomol. NMR*, **4**, 117–122.
43. Simorre,J.P., Zimmermann,G.R., Pardi,A., Farmer,B.T. and Mueller,L. (1995) Triple resonance HNCCCH experiments for correlating exchangeable and nonexchangeable cytidine and uridine base protons in RNA. *J. Biomol. NMR*, **6**, 427–432.
44. Sklenár,V., Dieckmann,T., Butcher,S.E. and Feigon,J. (1996) Through-bond correlation of imino and aromatic resonances in ¹³C-, ¹⁵N-labeled RNA via heteronuclear TOCSY. *J. Biomol. NMR*, **7**, 83–87.
45. Wöhnert,J., Ramachandran,R., Görlach,M. and Brown,L.R. (1999) Triple-resonance experiments for correlation of H5 and exchangeable pyrimidine base hydrogens in (¹³C,¹⁵N)-labeled RNA. *J. Magn. Reson.*, **139**, 430–433.
46. Phan,A.T. (2000) Long-range imino proton-¹³C J-couplings and the through-bond correlation of imino and non-exchangeable protons in unlabeled DNA. *J. Biomol. NMR*, **16**, 175–178.
47. Nikonowicz,E.P. and Pardi,A. (1992) Three-dimensional heteronuclear NMR studies of RNA. *Nature*, **355**, 184–186.
48. Sklenár,V., Peterson,R.D., Rejante,M.R. and Feigon,J. (1993) Two- and three-dimensional HCN experiments for correlating base and sugar resonances in ¹⁵N,¹³C-labeled RNA oligonucleotides. *J. Biomol. NMR*, **3**, 721–727.
49. Tolbert,T.J. and Williamson,J.R. (1997) Preparation of specifically deuterated and ¹³C-labeled RNA for NMR studies using enzymatic synthesis. *J. Am. Chem. Soc.*, **119**, 12100–12108.
50. Gilbert,S.D., Rambo,R.P., van Tyne,D. and Batey,R.T. (2008) Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nat. Struct. Mol. Biol.*, **15**, 177–182.
51. Klein,D.J., Edwards,T.E. and Ferré-D’Amaré,A.R. (2009) Cocystal structure of a class I preQ1 riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. *Nat. Struct. Mol. Biol.*, **16**, 343–344.
52. Koebel,M.R., Cooper,A., Schmadeke,G., Jeon,S., Narayan,M. and Sirimulla,S. (2016) S...O and S...N sulfur bonding interactions in protein-ligand complexes. Empirical considerations and scoring function. *J. Chem. Inf. Model.*, **56**, 2298–2309.
53. Egli,M. and Sarkhel,S. (2007) Lone pair-aromatic interactions. To stabilize or not to stabilize. *Acc. Chem. Res.*, **40**, 197–205.
54. Chawla,M., Chermak,E., Zhang,Q., Bujnicki,J.M., Oliva,R. and Cavallo,L. (2017) Occurrence and stability of lone pair-π stacking interactions between ribose and nucleobases in functional RNAs. *Nucleic Acids Res.*, **45**, 11019–11032.
55. Chen,G., Znosko,B.M., Kennedy,S.D., Krugh,T.R. and Turner,D.H. (2005) Solution structure of an RNA internal loop with three consecutive sheared GA pairs. *Biochemistry*, **44**, 2845–2856.
56. Corbino,K.A., Barrick,J.E., Lim,J., Welz,R., Tucker,B.J., Puskarczyk,I., Mandal,M., Rudnick,N.D. and Breaker,R.R. (2005) Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria. *Genome Biol.*, **6**, R70.
57. Lim,J., Winkler,W.C., Nakamura,S., Scott,V. and Breaker,R.R. (2006) Molecular-recognition characteristics of SAM-binding riboswitches. *Angew. Chem. Int. Ed. Engl.*, **45**, 964–968.
58. Kalvari,I., Argasinska,J., Quinones-Olvera,N., Nawrocki,E.P., Rivas,E., Eddy,S.R., Bateman,A., Finn,R.D. and Petrov,A.I. (2018) Rfam 13.0. Shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res.*, **46**, D335–D342.