

REVIEW

Structure-based Insights into Recognition and Regulation of SAM-sensing Riboswitches

Luqian Zheng^{1,2,3}, Qianqian Song², Xiaochen Xu², Xin Shen², Chunyan Li², Hongcheng Li², Hao Chen², Aiming Ren^{2,1,*}

¹Department of Gastroenterology/Department of Cardiology of the Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310058, China;

²Life Sciences Institute, Zhejiang University, Hangzhou 310058, China;

³The Eighth Affiliated Hospital, Sun Yat-Sen University, Shenzhen, Guangdong 518033, China.

*Corresponding authors (Aiming Ren, email: aimingren@zju.edu.cn).

2022-5-12; 2022-7-17

Keywords: Riboswitches, SAM, SAH, tertiary structures, sensing domain, expression platform

Abstract

Riboswitches are highly conserved RNA elements that located in the 5'-UTR of mRNAs, which undergo real-time structure conformational change to achieve the regulation of downstream gene expression by sensing their cognate ligands. S-adenosylmethionine (SAM) is a ubiquitous methyl donor for transmethylation reactions in all living organisms. SAM riboswitch is one of the most abundant riboswitches that bind to SAM with high affinity and selectivity, serving as regulatory modules in multiple metabolic pathways. To date, seven SAM-specific riboswitch classes that belong to four families, **one SAM/SAH riboswitch and one SAH riboswitch** have been identified. Each SAM riboswitch family has a well-organized tertiary core scaffold to support their unique ligand-specific binding pocket. In this review, we summarize the current research progress on the distribution, structure, ligand recognition and gene regulation mechanism of these SAM-related riboswitch families, and further discuss their evolutionary prospects and potential applications.

Accepted

Introduction

Riboswitches are conserved RNA elements that exist in the 5'-UTR of mRNA, which usually contain two interactive functional domains, a sensing domain and the adjoining expression platform (Figure 1). The sensing domain can bind specifically to the cognate ligand and induce the structure rearrangement of the overall riboswitch, thereby changing the conformation of the expression platform and turning the downstream gene expression on or off (Serganov and Nudler, 2013; Tucker and Breaker, 2005). Since the first discovery of riboswitches in 2002 that sense the vitamin derivatives, more than 55 classes of riboswitches have been identified to date, which were distinguished on the basis of the conserved sequence and the unique structural feature of the RNA motif as well as the cognate metabolites (Breaker, 2022; Tucker and Breaker, 2005; Wang and Breaker, 2008). Riboswitches can sense and bind a variety of cellular metabolites, including enzyme cofactors and their derivatives, nucleobases and their derivatives, amino acids, ions and others (Breaker, 2010; Serganov and Nudler, 2013). Most riboswitches are found in bacteria, while a few representatives such as TPP riboswitches exist in eukaryotes as well (Wachter, 2010; Wachter et al., 2007). Usually, the regulation of gene expression by riboswitches can function in two stages, i.e., by forming terminator/anti-terminator stems in transcription process or sequester/anti-sequester stems in translation process to turn off/on the downstream gene expression (Figure 1). In addition, riboswitches can also regulate the expression of related gene by affecting the decay of mRNA (*Escherichia coli lysC* riboswitch (Caron et al., 2012), *glmS* riboswitch (Winkler et al., 2004)) or by modulating the splicing of specific pre-mRNA (most eukaryotic TPP riboswitches) (Cheah et al., 2007; Wachter, 2010; Wachter et al., 2007).

S-adenosylmethionine termed SAM or AdoMet is an essential sulfonium compound involved in many biological processes, whose biosynthesis is based on the reaction of ATP and methionine catalyzed by SAM synthetase (Fontecave et al., 2004). The major function of SAM is to act as a ubiquitous methyl donor in biochemical methylation reactions in all the living organisms, where the methyl group of SAM is transferred to a nucleophile substrate catalyzed by methyltransferases to generate the neutral compound *S*-adenosylhomocysteine (SAH or AdoHcy) (Figure 2A) (Struck et al., 2012). Furthermore, SAM also functions as an indispensable

activating group donor in polyamine synthesis and SAM radical-mediated vitamin synthesis (Parveen and Cornell, 2011).

Due to the importance of SAM metabolites, a large number of SAM riboswitches have been employed by different bacterial species to mediate SAM biosynthesis and ultimately maintain the normal cellular SAM levels (Figure 2B) (Price et al., 2014; Wang and Breaker, 2008). By selectively sensing SAM, these riboswitches are also enabled to detect the ratio of SAM to SAH in the cell, which is important for effective regulation of SAM biosynthesis (Montange et al., 2010). Unlike SAM, SAH is detrimental to cells and is usually broken down by SAH-nucleosidase or SAH-hydrolases to recycle adenine as well as the *S*-ribosylhomocysteine or homocysteine (Kusakabe et al., 2015; Parveen and Cornell, 2011). The intercellular concentration of SAM and SAH in *E. coli* was about 0.4 mM and 1.3 μ M respectively (Halliday et al., 2010). Accumulation of SAH was predicted to attenuate the reaction tendency of SAM-dependent cellular methylation, resulting in cell growth inhibition (Kusakabe et al., 2015; Parveen and Cornell, 2011). Therefore, strict and timely regulation of cellular concentration SAM and SAH is essential for all organisms (Wang and Breaker, 2008; Weickhmann et al., 2019).

To date, seven SAM-responsive riboswitch classes termed SAM-I (Winkler et al., 2003), SAM-II (Corbino et al., 2005), SAM-III (Fuchs et al., 2006), SAM-IV (Weinberg et al., 2007; Yao et al., 2007), SAM-I/IV (Weinberg et al., 2010), SAM-V (Poiata et al., 2009), and SAM-VI (Mirihana Arachchilage et al., 2018), have been identified and proven to bind specifically with SAM while distinguishing against SAH as well as other closely related cellular metabolites (Figure 2B). One SAH riboswitch class was reported with higher binding specificity and affinity to SAH against SAM (Wang and Breaker, 2008; Wang et al., 2008), while the SAM-SAH motif was discovered to bind with SAM and SAH with comparable affinity (Weinberg et al., 2010). Each SAM riboswitch class has its own conserved sequence, featured secondary structure and distinct tertiary fold, which provide the structural basis for the local architecture to support the precise ligand-binding pocket. Since these riboswitches sense SAM and its metabolic derivative SAH, some of them also share some similarities in the ligand recognition pattern and the overall structure folding. In this paper, we focus on the 3D structure and recognition mechanism of

each SAM riboswitch class. Furthermore, we compare these SAM riboswitch classes to reveal how these representative members with different structures selectively bind the same target ligand SAM. Moreover, we extend our discussion from the structural studies of SAM riboswitches to other the evolution of SAM-sensing RNAs and their potential applications.

SAM-I riboswitch

SAM-I riboswitch was first described as “S box regulon” by Frank J. Grundy and Tina M. Henkin in 1998 (Grundy and Henkin, 1998), which was discovered to be involved in the global transcriptional control of methionine and cysteine biosynthesis genes in gram-positive bacteria (Grundy and Henkin, 1998). Later, S-box RNA domain was found to possess high affinity and specificity for SAM and therefore was consequently identified as a SAM-responsive riboswitch (Winkler et al., 2003). The first tertiary structure of SAM-I riboswitch sensing domain bound to SAM was solved by Rebecca K. Montange and Robert T. Batey in 2006 (Montange and Batey, 2006), which was followed by a series of reported structures solved with improved diffraction resolution in different states or in different conditions for sequences originating from other species (Lu et al., 2010; Montange et al., 2010; Stoddard et al., 2010).

In this review, we start our description with the first reported structure of SAM-I riboswitch bound to SAM (PDB: 2GIS, 2.90 Å). The sequence of the SAM-I riboswitch construct used to determine the tertiary structure is shown along with the predicted secondary structure in Figure 3A. There are four stems P1, P2, P3 and P4 connected to one central four-way junctional internal loop L1, where both stems P2 and P3 contains a bubble each in the middle. Notably, the bubble within stem P2 of SAM-I riboswitch is a standard kink-turn (k-turn) (Montange and Batey, 2006; Schroeder et al., 2011) (Figure 3A), a common RNA motif that typically mediates long-range interactions involved in the stabilization of RNA tertiary structure (Schroeder et al., 2011). For clarity, we split stem P2 into stems P2a and P2b with the k-turn in the middle.

The schematic secondary structure and the solved tertiary structure are shown in Figure 3B and Figure 3C respectively. It is noted that in the three-dimensional

structure of the SAM-I riboswitch, the two co-axially stacked helices (Helix P1/4 and Helix P2/3) are formed by four stems (P1, P2a, P3 and P4). Stems P1 and P4 form helix P1/4 involving some residues from the internal loop L1, while stems P2 and P3 form helix P2/3. Between these two helices, stem P2a points from the k-turn, the terminal of helix P2/3 towards helix P1/4, and forms a long-range pairing interaction with residues C65-G68 within J3/4, thereby generating one additional pseudoknot stem (PK) between stem P2a and stem P4 (Figure 3B-C). Besides, other additional canonical (A85-U64) and non-canonical (A85·A24) long-range interacting base pairs are also formed in the tertiary fold as shown by the dashed line in Figure 3B. Taken together, all of these interactions define the overall structure of SAM-I riboswitch and therefore contribute to the folding of SAM binding pocket (Montange and Batey, 2006).

The SAM binding pocket is located between the minor grooves of stems P1 and P3, and is capped by the junctional region J1/2 between stems P1 and P2 (Figure 3C-E). The methionine tail and the adenine moiety of SAM are folded in parallel to fit into the binding pocket (Figure 3D-E). The adenine moiety is anchored by hydrogen-bonding interactions with A45 and U57 (Figure 3F) and stacking interactions with C47-G56 from stem P3 (Figure 3D), while the methionine tail is held in place by a triple base pair (G58-C44)·G11 that is formed by the adjacent G58-C44 from P3 and G11 from J1/2 (Figure 3G). Here it is noted that the ribose moiety of SAM adopts C3'-endo sugar pucker conformation and forms one hydrogen bond with the ribose of C47 (Figure 3F).

The diffraction resolution of the first reported SAM-I riboswitch structure was further improved by the introduction of mutations A94G and U34C to alter the lattice contacts, which provided more details on the tertiary structure folding and SAM recognition, and further revealed the involvement of magnesium ions in its structural stabilization (Montange et al., 2010). Another research work on the structure of SAM-I riboswitch from *Bacillus subtilis yitJ* investigated the roles of magnesium cations in ligand recognition (Lu et al., 2010), which was further illuminated by chemical probing of the RNA (Stoddard et al., 2010).

SAM-I riboswitches preferentially bind to SAM and discriminate against SAH (Winkler et al., 2003). The structures of SAM-I riboswitches in complex with SAH were also solved to explore the ligand discrimination (Montange and Batey, 2006; Montange et al., 2010). The positively charged sulfonium ion in SAM seems to form stronger electrostatic contacts with the adjacent carbonyl oxygens of U7 and U88 than the neutral sulfur atom in SAH (Figure 3H). The following mutation and ITC-binding experiments confirmed the contributions of two highly conserved base pairs U7-A87 and A6-U88 to the binding specificity of SAM (Figure 3D and 3H)(Montange et al., 2010). All these results suggested that electrostatic contacts between the positively charged sulfonium ion of SAM and adjacent carbonyl oxygens in SAM-I riboswitch contribute to the high selectivity for SAM and reinforce the discrimination between SAM and SAH.

Most SAM-I riboswitches were identified in Gram positive bacteria located before the genes involved in sulfur metabolism, methionine biosynthesis, cysteine biosynthesis and SAM biosynthesis (Winkler et al., 2003). One unanimous regulation mechanism of SAM-I riboswitch is through a metabolite-induced transcription termination mechanism. The entrance of the RNA polymerase will be blocked by the formation of a terminator stem upon SAM binding, which turns off the downstream gene expression. In the absence of SAM binding, the downstream gene expression will be turned on with formation of an anti-terminator stem (AT) (Boyapati et al., 2012; Eschbach et al., 2012; Manz et al., 2018; Manz et al., 2017; Montange and Batey, 2006; Winkler et al., 2003). Dynamic studies of the full-length SAM-I riboswitch revealed that terminator and anti-terminator conformations coexist, with SAM binding shifting the conformation to the terminator state (Manz 2017, NCB). Conformational changes depend on both the concentrations of Mg^{2+} ions and the ligand SAM, with Mg^{2+} ions being the key factor in global folding (Boyapati et al., 2012; Eschbach et al., 2012; Huang et al., 2012).

Recently, one exceptional SAM-I riboswitch termed SAM- I_{Xcc} was identified from the *Xanthomonas campestris*, a Gram-negative bacteria specie (Tang et al., 2020). In contrast to previously characterized SAM-I riboswitches, SAM- I_{Xcc} lacks the Rho-independent terminator that regulates gene expression at transcriptional level, but regulates methionine synthesis via the *met* operon at translational level (Tang et

al., 2020). Further studies indicated that SAM-I_{Xcc} riboswitch expression platform has a dual function, which not only represses the gene expression in response to SAM binding with the adjoining sensing domain, but also initiates the gene translation upon uncharged Met tRNA binding to the anti-Shine-Dalgarno (SD) sequence (Tang et al., 2020). SAM-I_{Xcc} adopts a previously unrecognized regulation mechanism. Although SAM-I riboswitch was the first SAM riboswitch to be identified, perhaps more SAM-I riboswitches with diverse characteristics are yet to be explored in different species.

SAM-II riboswitch

SAM-II riboswitch was identified as a novel class of SAM-binding riboswitch by Breaker's lab using the comparative sequence analysis method supplemented with structural probing (Corbino et al., 2005; Lim et al., 2006). Unlike SAM-I riboswitch, SAM-II riboswitch was found in proteobacteria, particularly alpha-proteobacteria (Gilbert et al., 2008). SAM-II riboswitch has a simpler, compact architecture composed of a hairpin and a pseudoknot. Compared to SAM-I riboswitch ($K_d = 4\sim 5$ nM) (Winkler et al., 2003), SAM-II has a weaker binding affinity for SAM approximately 0.67 μ M (Gilbert et al., 2008). The sequence (51 nt) used to determine SAM-II riboswitch structure is located in the upstream of the *metX* gene in the Sargasso Sea metagenome (Gilbert et al., 2008), which contains the functional core regulatory elements including both the aptamer domain and expression platform (Figure 4A) (Gilbert et al., 2008).

The overall structure of SAM-II riboswitch in complex with SAM (PDB: 2QWY, 2.8 Å) folds in H-type (hairpin type) pseudoknot scaffold (Gilbert et al., 2008; Hilbers et al., 1998), in which the hairpin loop forms a pseudoknot stem (P2) with the 3'-end of the single-stranded segment. The zipped L1 (shown in pale cyan, Figure 4A-B) ascends along the major groove of P2 and forms continuous non-canonical pairing interactions, resulting in a triple-helical structure above stem P1 (Figure 4A-B). L2 (shown in pale green) is stabilized by the minor groove of stem P1 (Figure 4A-B).

The ligand SAM, which adopts a splayed conformation, is bound to the major groove side of stem P2 (Figure 4C), where the adenine moiety of SAM intercalates into stem P2 and forms a Hoogsteen base pair with the highly conserved residue U44 (Figure 4C-E). The methionine moiety of SAM is recognized by A47, which is further

stabled by A19 (Figure 3F). The positively charged sulfur moiety is anchored by the carbonyl oxygen groups of U11 and U21 from a base triple U21-A45-U11 (Figure 4G). All these interactions endow SAM-II riboswitch with high ability to distinguish SAM from SAH or other SAM analogues (Lim et al., 2006) (Corbino et al., 2005; Gilbert et al., 2008), (Lim et al., 2006).

Compared to SAM-I riboswitch, SAM-II riboswitch forms fewer hydrogen bonding interactions with SAM inside the ligand binding pocket (Figure 3D-H & Figure 4C-G). Accordingly, the binding affinity of SAM-II riboswitch to SAM is about 1000-fold weaker than that of SAM-I riboswitch (Gilbert et al., 2008; Winkler et al., 2003). Another striking difference between SAM-I riboswitch and SAM-II riboswitch is that the bound SAM folds into a “C” shaped conformation in SAM-I riboswitch (Figure 3D-E), whereas in the SAM-II riboswitch it adopts an “L-like” shape conformation (Figure 4C-D).

Further research to investigate the free state of SAM-II riboswitch indicated that in the absence of SAM, the pseudoknot structure formation in SAM-II riboswitch (Figure 4A-B) is highly transient and dynamic, where multiple conformations and exposure of SD sequence facilitate the translation initiation (Chen et al., 2016; Chen et al., 2012; Doshi et al., 2012; Gilbert et al., 2008; Haller et al., 2011a; Xue et al., 2015). Upon ligand binding, SAM-II riboswitch undergoes a conformational change and generate a stable H-type pseudoknot structure with a sequestered SD sequence, which leads to the turnoff of downstream gene expression (Chen et al., 2016; Doshi et al., 2011; Haller et al., 2011a; Haller et al., 2011b).

SAM-III riboswitch

SAM-III riboswitch, also termed as SMK box, was the third new SAM-binding riboswitch, which was first discovered in the 5'-UTR of the *metX* gene in lactic acid bacteria and characterized as a conserved regulatory RNA element for SAM synthesis (Fuchs et al., 2006; Lu et al., 2008). Compared to SAH, SAM-III riboswitch prefers binding to SAM by at least 100-fold (Fuchs et al., 2006). In response to SAM binding, SAM-III riboswitch undergoes structural rearrangement and sequesters the SD sequence in the bound state structure, which results in translational repression of the downstream genes such as the *lacZ* reporter gene. Notably, mutants that lose SAM

binding affinity also lose their repressive ability (Fuchs et al., 2006). The crystal structure of SAM-III riboswitch from *Enterococcus faecalis* (PDB: 3E5C, 2.25 Å) was determined in complex with SAM at a resolution of 2.2 Å resolution by the Ke's group in 2008 (Lu et al., 2008) (Figure 5A-B). The construct used to determine the tertiary structure is shown in Figure 5B, in which the GAAA tetraloop was introduced to replace two flexible stem-loops (P3 and P4 stem-loops), respectively.

The tertiary structure of SAM-III riboswitch bound to SAM adopts an inverted “Y”-shaped architecture that is highly conforms to the predicted three-way junction secondary structure (Figure 5A-B). The bulge loop L2 (A29-G31) within stem P2 flips down and forms long-range interaction with the residues in L1 (Figure 5A-B), where the SAM binding pocket is formed at the intersection of P1, P2, P3, L1 and L2 (Figure 5B-C).

The bound SAM intercalates into the binding pocket, where the adenosine moiety of SAM is recognized by G7 and A38 in loop L1 and becomes stacked above G48-C6, which is the terminal base pair in stem P1 (Figure 5C-E). The ribose moiety of SAM adopts C2'-endo sugar pucker conformation and is stabilized by G47, the terminal residue in stem P3 (Figure 5E). Similar to other SAM riboswitches, electrostatic interactions are also involved between the positively charged sulfur moiety of SAM and the adjacent electronegatively charged atoms (2'-OH of G36 and O4 of U37) in the binding pocket, which enable SAM-III riboswitch with the specific recognition capacity for SAM (Figure 5F). Without this positively charged sulfonium moiety, SAH has a weaker affinity for SAM-III riboswitch and forms fewer contacts in the binding pocket (PDB: 3E5E, 2.90 Å)(Figure 5G)(Lu et al., 2008). The following molecular dynamics (MD) simulations studies also clarified that the specific binding of the adenosine moiety of SAM to SAM-III riboswitch as well as the involved electrostatic interactions are crucial to the binding discrimination between SAM and SAH (Priyakumar, 2010). It is noted that the methionine moieties of the bound SAM or SAH in the SAM-III structure are relatively flexible and form different conformations at the binding site, which is consistent with the observation of poor electron density beyond sulfur in the crystal structure (Figure 5E-G) (Lu et al., 2008). SAM adopts a sheared “C”-shaped structure to fit the binding pocket, while

SAH adopts “L”-shaped folding and only form one possible hydrogen bond with A38 using the terminal amino group (Figure 5E-G).

SAM-III riboswitch regulates downstream gene expression via translation initiation mechanism (Fuchs et al., 2006; Gong et al., 2016; Lu et al., 2008; Smith et al., 2010; Suresh et al., 2016). Sequence alignment and structure research indicated the Shine-Dalgarno sequence is involved in the formation of stems P1 and P3 in the SAM-binding state structure of SAM-III riboswitch, thus participating not only in SAM recognition, but also contributing to define SAM-binding pocket (Figure 5A-B)(Lu et al., 2008) (Fuchs et al., 2006). The free state of SAM-III riboswitch was detected by in-line probing in the absence of SAM, which showed that most SAM-III riboswitch molecules have an translation “ON” conformation with a free SD for ribosome binding and a minority group of SAM-III riboswitch molecules adopted a SAM-bound-like (READY) conformation with a fixed SD (Lu et al., 2011). Upon SAM binding, the “OFF” conformation of SAM-III riboswitch with sequestered SD was stabilized (Lu et al., 2011). Consistent results were obtained from systematic helix-based computational methods (Gong et al., 2016) and atomistic molecular dynamics (MD) simulations, where the open binding pocket switches to a stable duplex-like architecture induced by SAM exposure (Suresh et al., 2016).

SAM-IV riboswitch

SAM-IV riboswitch was the fourth SAM sensing RNA motif identified to directly modulates the transcription or translation of genes involved in sulfur metabolism (Weinberg et al., 2008). SAM-IV riboswitch shows similarities to SAM-I riboswitch in terms of conserved nucleotide positions and core domain alignment of the predicted secondary structures, but shows specific differences in the absence of stem P4 and the formation of stem P5 as well as an additional pseudoknot stem PK-2 in the overall scaffold (Figure 3A and 6A) (Trausch et al., 2014; Weinberg et al., 2008). Compared to SAM-I riboswitch ($K_d = 4\sim 5$ nM), SAM-IV riboswitch has a 30-folds weaker binding affinity to SAM ($K_d = 150$ nM), but retains the ability to discriminate against SAH ($K_d = 20$ μ M)(Weinberg et al., 2008).

The free form and bound form tertiary structures of SAM-IV riboswitch were solved by cryo-EM method at 3.7 Å (PDB: 6UES) and 4.1 Å (PDB: 6UET) resolution

respectively (Figure 6)(Zhang et al., 2019). Due to the present resolutions, these cryo-EM structures are limited to provide some detailed structural information such as the ligand binding interactions (Zhang et al., 2019). However, these structures show a global profile of SAM-IV riboswitch tertiary structure clearly, which adopts different scaffolds compared to SAM-I riboswitch and shares similar ligand binding module (Figure 3A and Figure 6B). Previous studies based on sequence alignment and biochemical analysis revealed that the SAM-binding pocket of SAM-IV riboswitch is similar to that of SAM-I riboswitch, as confirmed by these cryo-EM structures (Figure 6D-E) (Tausch et al., 2014; Weinberg et al., 2008). Many SAM-IV riboswitch sequences were identified near the predicted start site of the downstream genes and demonstrated the ability to repress reporter gene expression *in vivo*, so it is reasonable to infer that SAM-IV riboswitch is capable of sequestering the ribosome binding site upon ligand binding (Weinberg et al., 2008).

SAM-I/IV riboswitch

SAM-I/IV riboswitch is another new class of SAM sensing elements involved in gene regulation in sulfur metabolism (Weinberg et al., 2010). Compared to SAM-I and SAM-IV riboswitches, SAM-I/IV riboswitch has the common core region and distinguishable peripheral architecture (Figure 3A, 6A and 7A). The phylogenetic tree strongly suggests that SAM-IV riboswitch evolved from SAM-I/IV riboswitch (Tausch et al., 2014). However, to date, no definite conclusion about the evolutionary relationship between SAM-I and SAM-I/IV riboswitch have been drawn from the related studies. The ancestor of SAM-I/IV riboswitch might have been a SAM-I-like RNA, or they may have evolved independently (Breaker, 2010; Tausch et al., 2014).

The secondary structure of SAM-I/IV riboswitch is shown in Figure 7A, which is characterized by a central four-way junction consisting of five stems P1-P5, a predicted pseudoknot (PK). The three-dimensional structure of SAM-I/IV riboswitch was solved by X-ray crystallography method (PDB: 4OQU, 3.20 Å)(Tausch et al., 2014). To facilitate crystallization, the original variable stem-loops of P2, P4, and P5 were replaced with GAGA tetraloop (Figure 7A-B). Further deletion of U92 was undertaken to achieve high diffraction resolution (Tausch et al., 2014).

The finally determined tertiary structure of SAM-I/IV riboswitch consists of three helices with SAM ligand positioned at the center of the whole structure (Figure 7B-C). The long co-axial helix is constituted by continuously stacked P2, P3, and the pseudoknot stem PK, while the short co-axial helix is formed by stacking of P1 and P4 stem in parallel to the long co-axial helix (Figure 7B-C). Stem P5 is positioned almost perpendicular to stem P1 at the bottom and bridges these two parallel helices together (Figure 7B-C). SAM is recognized by the delicate pocket assembled by stems P1, P3 and J1/2, which is further wrapped and supported by the peripheral structure of stems P5, PK, P2, and P4 (Figure 7D-E).

The SAM-binding site of the *env87* SAM-I/IV riboswitch described above is almost identical to that of *T. tengcogensis* SAM-I riboswitch in superimposition of the two SAM-bound structures (Trausch et al., 2014). The adenine ring of SAM is recognized jointly by A25 and U47, while the 3'-OH of the ribose sugar contacts with A3 (Figure 7F). The methionine moiety of SAM is stabilized by forming a multiple hydrogen bonding network with G8 and G48, which are further stabilized by C24 (Figure 7G). Carbonyl oxygens of U4 and U69 are located nearby and have the capacity to form electrostatic contacts with the positively charged sulfonium ion of SAM (Figure 7H). Similar SAM binding pockets establish the same structural basis for SAM binding activity. It was noted that the binding affinity of the *env87* SAM-I/IV riboswitch to SAM is about 90 nM, which is comparable to that of SAM-I riboswitch (Trausch et al., 2014).

SAM-V riboswitch

SAM-V riboswitch was discovered as a structured RNA motif in the marine alpha-protobacteria and Bacteroidetes, and was later proven to be another riboswitch sensing SAM (Meyer et al., 2009; Poiata et al., 2009). The consensus sequence and secondary structure of SAM-V riboswitch share similarities with that of SAM-II riboswitch, but differ significantly from known families of SAM-binding riboswitches. The critical nucleotides that contribute to SAM binding in SAM-II riboswitch are highly conserved and located at the same positions in SAM-V riboswitch, therefore SAM-V riboswitch is predicted to adopt the same binding mode as SAM-II riboswitch (Figure 4A and 8A)(Poiata et al., 2009). However, most members of SAM-II

riboswitch class seem to function as transcription regulators, whereas SAM-V riboswitches are involved in translation process (Huang and Lilley, 2018).

The crystal structure was solved with the sequence of *metY* SAM-V riboswitch from *Candidatus Pelagibacter ubique*, in which a bromocytidine at position C4 was introduced for phase determination (Figure 8A)(Huang and Lilley, 2018). The overall structure of SAM-V riboswitch in complex with SAM (PDB: 6FZ0, 2.50 Å) is also folded in an H-type pseudoknot scaffold as SAM-II structure (Figure 8B). Stem P1 stacks with the pseudoknot structure comprised by stem P2 and L1 (Figure 8A-B). L2 connects to stem P1 by three continuous nucleotides A40-A42. Partially pairing interaction in L2 produces a short stem below stem P1 (Figure 8A-B).

The bound SAM lies in the narrow groove composed by loop L1 and stem P2 (Figure 8B-D), with the adenine moiety intercalated into the binding pocket assembled by U8, U20, G21, U47, and A48 (Figure 8D-G). The adenine moiety is specifically recognized by U47, which is located in the middle of P2 and stabilized by the formation of a non-canonical base pair with U8 in L1 (Figure 8E). Two base triple, (U20-A48)·U9 and (G21-C46)·A7 constitute the floor and ceiling of the binding pocket and sandwich the adenine moiety on both two sides (Figure 8D-E). In addition, the 2'-OH of SAM forms a hydrogen bond with N7 of G21 (Figure 8E). The aminoacyl group of SAM pairs with the nucleobase of A50 (Figure 8F). One Mg²⁺ cation was also found to be involved in ligand binding, which forms a water-mediated coordination with the carboxyl group of SAM and direct coordination to U10 (Figure 8F). The positively charged SAM sulfonium is positioned in the major side of a base triple (U20-A48)·U9 and may form electrostatic interactions with O4 atoms of U9 and U20 (Figure 8G), which was confirmed by the fact that the atom-specific substitution of O4 from U9 or U20 by sulfur leads to a nearly complete loss of binding activity to SAM (Huang and Lilley, 2018). The electrostatic interactions of the positively charged SAM sulfonium also contribute to the high binding selectivity of SAM-V riboswitch to SAM, but not to SAH. In ITC measurements, the binding affinity of *metY* SAM-V riboswitch to SAM was about 5 μM, while no binding to SAH is detectable (Huang and Lilley, 2018).

Notably, in most SAM-V riboswitch representatives such as the sequence used in tertiary structure determination, the SD element is located in the 3'-end of stem P2

and near the following start codon (Figure 8A)(Huang and Lilley, 2018; Meyer et al., 2009; Poiata et al., 2009). The formation of a short stem P2b ascending above the SAM binding site is anticipated to shade the SD sequence and prevent translation initiation (Huang and Lilley, 2018), which functions in a similar pattern to the *metX* SAM-II riboswitch (Gilbert et al., 2008).

Interestingly, it was found that in many cases SAM-II and SAM-V riboswitches exist in tandem across genomes, wherein SAM-II riboswitch followed by an intrinsic transcription terminator stem is always arrayed in front of SAM-V riboswitch that contains the purine-rich SD sequence as well as preceding the start codon of related genes (Poiata et al., 2009). It is speculated that independent SAM-II and SAM-V riboswitches arrayed in tandem are able to function in both transcription and translation steps to control the gene expression by sensing the same ligand SAM, which may represent some type of more complicated function derived from simple riboswitches combination (Poiata et al., 2009).

SAM-VI riboswitch

SAM-VI riboswitch was identified as another new class of SAM-responsive riboswitch originating from *Bifidobacterium* species (Mirihana Arachchilage et al., 2018). The consensus central sequence and the predicted secondary structure model of SAM-VI riboswitch are shown in Figure 9A, which is characterized by a three-way junctional architecture with SD sequence involved in the formation of stems P1 and P3. In comparison with other SAM-sensing riboswitches, some similarities were found between SAM-VI riboswitch and SAM-III riboswitch (Mirihana Arachchilage et al., 2018). Beside adopting an overall similar three-way junctional fold, five highly conserved nucleotides located in the junctional region are almost identical to the corresponding nucleotides in SAM-III riboswitch that are involved in SAM interaction (Figure 9A)(Mirihana Arachchilage et al., 2018). However, sufficient differences between SAM-III and SAM-VI riboswitches still exist. The numbers of nucleotides in J2/3 and the nucleotides in J1/2 flanking the junction region showed obvious variations, suggesting that SAM-VI riboswitch may have a different binding pattern to SAM (Figure 9A)(Mirihana Arachchilage et al., 2018). In addition, SAM-III and SAM-VI riboswitches are also phylogenetically distant (Mirihana Arachchilage et al., 2018).

The three-dimensional structure (PDB: 6LAS) was determined at 2.7 Å resolution with the construct originating from *B. angulatum* 59 *metK*, in which the variable P2 stem-loop was replaced by the U1A protein binding loop to assist the crystallization via co-crystallization with U1A protein (PDB: 6LAS, 2.71Å) (Figure 9B-C)(Sun et al., 2019b). As shown in Figure 9B, the overall fold of SAM-VI riboswitch in complex with SAM is consists mainly of two helices, with the long helix comprised of stems P1 and P2 and the short helix comprised of stem P3 (Figure 9B). These two helices are arranged almost perpendicularly. The U1A protein binds to the U1A binding loop (in light gray), which is positioned at the top terminal of the long helix, and has no interaction with other regions of the RNA (Figure 9B-C). The bound ligand SAM is located in the intersection of stems P1, P2 and P3, mainly recognized and stabilized by a compact binding pocket made up of J1/2 and J2/3 (Figure 9B-C). The ADP moiety of SAM intercalates into the binding pocket with the methionine moiety pointing outwards (Figure 9D). The adenine base of SAM is paired with U8 in J1/2 and becomes stacked between G9-C32 and G7-G33, which are further anchored by three continuously stacked nucleotides A36-A37-G38 (Figure 9E-F). Meanwhile, the 2'-OH of ribose sugar of SAM is hydrogen-bonded with the non-bridging phosphate oxygen of G33, and the amino group in the methionine moiety of SAM forms an additional hydrogen bond with the O4 of U8 (Figure 9E).

SAM-VI riboswitch selectively binds with SAM over SAH about 33-fold in the ITC titration (Sun et al., 2019b). Tertiary structure of SAM-VI riboswitch shows that the positively charged sulfonium moiety of the bound SAM is located adjacent to O4 atoms of U6 and U8 and may form potential electrostatic interactions to facilitate the binding discrimination of SAM over SAH (Figure 9G). Structure-based mutation of U6C resulted in a remarkable decrease of the binding discrimination between SAM and SAH by about 5-fold (Sun et al., 2019b), confirming the importance of the electrostatic interactions related with the positively charged sulfonium moiety of the bound SAM.

As mentioned before, SD sequence is involved in the formation of stem P1 and P3 in the bound form structure, which indicates that SAM-VI riboswitch is a translational regulator (Sun et al., 2019b). Structure-based analysis of the related sequence of the nascent SAM-VI riboswitch from different species revealed that the

key nucleotides involved in SAM recognition are able to form an alternative stem P0 with the 5'-end sequence before the RNA motif and disrupt the formation of stem P1, thus leading to release of the SD sequence for ribosome binding (Figure 9H)(Sun et al., 2019b). The following 2-aminopurine (Ap) based fluorescence assay and the *lacZ* reporter assay supplemented with related mutations confirmed the involvement of stem P0 in the translation regulation process, which produced a more accurate downstream gene expression regulation mechanism for SAM-VI riboswitch (Figure 9H)(Sun et al., 2019b). This is the first time to discover the 5'-end RNA elements of riboswitches participating in the gene expression regulation, which renovate the current regulation model of riboswitches.

SAM-SAH riboswitches

As previously described, seven SAM riboswitches specifically recognize SAM over SAH, while the SAH riboswitch works in the opposite manner and distinguishes SAH from SAM. Comparative genomic searches identified an RNA motif called SAM-SAH, which binds to SAM and SAH with comparable affinity (Weinberg et al., 2010). SAM-SAH riboswitches are located upstream of *metK* genes in the order Rhodobacterales of α -proteobacteria (Weinberg et al., 2010). Since they are associated with *metK* genes involved in SAM biosynthesis and the intercellular concentration of SAM is significantly higher than SAH, it is considered that SAM-SAH riboswitches function partially as a SAM-sensing riboswitch class (Weinberg et al., 2010).

The tertiary structure of SAM-SAH riboswitch in complex with SAH was first determined using NMR-spectroscopy by Wohnert's group in 2019 (Weickhmann et al., 2019). Later, the crystal structures of SAM-SAH riboswitch (*Roseobacter* sp. SK209-2-6) complexed with SAH or SAM were solved by Lilley's group at resolution 1.70-2.50 Å (Huang et al., 2020). The NMR structure and the crystal structure showed similar overall folding and ligand recognition pattern. Like SAM-II and SAM-V riboswitches, SAM-SAH riboswitch also adopts a H-type pseudoknot architecture (Figure 10A). Stem P1, the zippered L1 and stem PK form a coaxial long helix, with the ligand-binding pocket located between L1 and PK and being shaped by the related residues in tertiary folding (Figure 10A-B). The adenine moiety of SAH forms Hoogsteen pairing interaction with U23 and one hydrogen bond with 2'-OH of

U44, which are further stacked by C21-G46 and C15-G24 (Figure 10C-D). The 2'-OH of SAH forms hydrogen bonds with N7 of G46 and O4' of ribose-G16 (Figure 10E). The α -amine of SAH is hydrogen-bonded to O4 of U23, while no such interaction is observed in the SAM complex structure (Figure 10E-F). Unlike SAM riboswitches, it seems that the methyl group and the positive charged sulfonium of SAM do not form additional interaction with the RNA molecule in SAM-SAH riboswitch structure. Thus, SAM-SAH riboswitch exhibits comparable binding capacity for SAH and SAM. Since the recognition of SAM and SAH mainly involves the interaction of the adenine moiety of SAM-SAH riboswitch, SAM-SAH riboswitch can also bind to some other variant ligands that containing the adenine moiety such as dCSAH, 5DMA and AMP (Huang et al., 2020).

Incidentally, it was found that a second SAH molecule binding site was also formed at the apical interface of two adjacent molecules in the crystal structure of SAM-SAH riboswitch in complex with SAH (Huang et al., 2020). However, according to previous biochemical results, the second SAH binding site does not exist in solution and should be formed casually in the crystal packing (Huang et al., 2020). Therefore, we will not extend our discussion of the second SAH ligand-binding pocket here.

For gene expression regulation, SAM-SAH riboswitch plays a role in the translation process. NMR spectroscopy of a SAM-SAH riboswitch variant indicated that SAH binding would induce the formation of a pseudoknot stem, which would sequester the SD sequence from ribosome binding and turn off the downstream gene expression (Weickhmann et al., 2019). The conformational change upon ligand binding is consistent with the results of in-line probing (Weinberg et al., 2010) and the single-molecule FRET (Huang et al., 2020).

SAH riboswitch

SAH riboswitches possessed higher binding specificity and affinity for SAH over SAM in both Gram-positive and Gram-negative bacteria, which regulated the expression of genes associated with the production of expended SAM coenzymes (Wang and Breaker, 2008; Wang et al., 2008). SAH motif from the 5'-UTR of *ahcY* mRNA (*P. syringae*) has been shown to activate the expression of downstream

reporter genes *in vivo* in the presence of SAH (Wang et al., 2008). Another representative SAH motif called 68 *metH* was found to be located upstream of the *metH* gene in *Dechloromonas aromatica*. In-line probing and the equilibrium dialysis experiments indicated that 68 *metH* could strongly distinguish SAH from SAM by three orders of magnitude (Wang et al., 2008). In contrast to SAM riboswitch families, SAH riboswitch class contained distinct consensus sequence and was featured with the specific secondary structure scaffold. SAH riboswitches adopt a “LL-type” pseudoknot architecture that contains a central junction and surrounded stems P1 and P2 and a pseudoknot stem P4, where the highly conserved nucleotides are almost exclusively distributed in the junctional region and the pseudoknot stem P4 (Wang and Breaker, 2008; Wang et al., 2008). Notably, for more than 60% of SAH motifs, there is one additional non-conserved stem P3 is present between stem P2 and P4 (Wang et al., 2008).

The tertiary structure of SAH riboswitch (without P3 stem) in complex with SAH was determined by Batey's group in 2010 (Edwards et al., 2010). They used a sequence from *Ralstonia solanacearum* (Rso), a plant pathogen in the structure determination. To improve the crystal diffraction, a U13C point mutation and a truncated stem P2 were introduced in crystallization process. Further optimization was performed by replacing the variable L2 loop with a more stable GAGA-tetraloop, resulting in high diffraction resolution (2.18 Å) (Figure 11A). In the tertiary structure, stems P2, P4 and the intermediate packed non-canonical base pairs from L1 form a co-axial long helix, while stem P1 is perpendicular to the above long helix (Figure 11B). Several terminal nucleotides in stem P1 didn't form regular intramolecular base-pairing interaction, but were swapped with stem P1 from another molecules, thus forming dimer in the crystal structure.

The bound SAH intercalates into the long helix and is stabilized by highly conserved residues from L1 at the intersection of stem P1, P2 and P4 (Figure 11B). Within the binding pocket, the adenine ring of SAH interacts with G15 and forms a non-canonical G-A base pair, which is further sandwiched by U14-A29 and C16-G31 (Figure 11C-D). The α -amine of SAH forms an extensive interaction with the phosphate of G30, the N3 and 2'-OH of G47 (Figure 11E). The carboxylate oxygen of SAH is hydrogen-bonded with the 2'-OH of G31 and O4' of C32 ribose. Notably,

these interactions between the aminoacyl group of SAH and the RNA molecule exhibit some minor variations in other molecules in the asymmetric unit.

SAH riboswitch makes a clear distinction between SAH and SAM, which is supported by the observation in the tertiary structure. The distance between sulfoether moiety of SAH and the ribose of A29 was found to be 3.2 Å, where the additional methyl in SAM would cause a steric clash (Figure 11D). The binding affinity of wild-type Rso SAH riboswitch with SAH generated from ITC experiments is around 32 nM (Edwards et al., 2010). Since the adenine moiety is involved in the recognition of SAH, other metabolites containing the adenine moiety, such as ATP and NAD⁺, were also found to have a weaker binding affinity for SAH riboswitch at about 100 μM and 150 μM, respectively. However, the typical cellular concentration of these metabolites are significantly higher than these equilibrium dissociation constants, so these adenine-containing compounds may act as competitive inhibitors of SAH binding and weaken the regulation of SAH riboswitch (Edwards et al., 2010).

Discussion

To date, there are nine SAM-related riboswitches have been identified, seven of which termed SAM-I, SAM-II, SAM-III, SAM-IV, SAM-I/IV, SAM-V and SAM-VI selectively bind to SAM and discriminate against SAH, SAH riboswitch sense SAH and SAH/SAM bind with SAH and SAM with similar binding affinity. Based on the conserved sequence, characterized secondary structure and the solved tertiary structure, these SAM-sensing riboswitches are assigned to four SAM riboswitch families termed SAM-I (containing SAM-I, SAM-IV and SAM-I/IV riboswitches), SAM-II (containing SAM-II and SAM-V riboswitches), SAM-III and SAM-VI (Figure 12). For comparison, we summarize the ligand recognition patterns and the key residues involved in these SAM-related riboswitches in Table 1.

In SAM-I riboswitch family, SAM-I, SAM-IV and SAM-I/IV riboswitches adopt nearly identical SAM-sensing core architecture composed of stem P1, P2, P3 and the junction loop J1/2, which are further stabilized and supported by different peripheral structures such as regular and pseudoknot stems P4, P5, PK1 and PK2 (Figure 12A). Stem P4 resides in the junction region J3/1, while stem P5 resides at the 3'-end of stem P1. The pseudoknot stem PK1 is formed between the 3'-end of stem P1 and the

stem-loop of stem P3, while PK2 is formed between the junction region J3/1 and the stem-loop of stem P2. SAM-I riboswitch is characterized by the presence of stem P5 and pseudoknot stem PK2. SAM-IV riboswitch is characterized by the presence of stem P4 and pseudoknot stem PK1, PK2. SAM-I/IV riboswitch is characterized by the presence of stem P4, P5 and pseudoknot stem PK1. Different combination of peripheral stem structures (P4, P5, PK1, PK2) generate different SAM-I riboswitch family members. It is an intriguing question whether there are more SAM-I riboswitch family members to be identified in the future. The adenine base of SAM in SAM-I riboswitch superfamily interacts with a sheared base pair A and U and form a base triple intercalating into stem P3, while the methionine moiety is recognized by a base triple (G-C)·G formed between stems P3 and J1/2 (Figure 3F-G and 7F-G). The carbonyl oxygens of the two uridine bases in stem P1 are involved in the electrostatic interaction of the positively charged sulfonium group in SAM, which contribute to the high binding selectivity of SAM over SAH (Figure 3H and 7H).

SAM-II and SAM-V riboswitches belonging to SAM-II riboswitch family, which both employ similar global H-type pseudoknot folding and highly consistent binding modules (Figure 12B). In SAM-II riboswitch, the pairing region P2e is formed coaxially above stem P2. Whereas in SAM-V riboswitch, stem P1 is extended by the forming the pairing region P1e below P1. The adenine base of SAM in SAM-II riboswitch superfamily pairs with a conserved U and become stacked between two base triples (U-A)·U and (G-C)·A (Figure 4D-E and 8D-E). Inside the binding pocket, G from (G-C)·A forms another hydrogen bond with the 2'-OH of the ribose, and the two uridine bases from (U-A)·U are involved in the interaction with the sulfonium group in SAM (Figure 4D-E,G and 8D-E,G). The methionine moiety of the SAM-II riboswitch superfamily is recognized by one adenine above the binding pocket (Figure 4F and 8F). As SAM-II riboswitch family, SAM-SAH riboswitch also adopts H-type pseudoknot architecture and the adenine moiety of SAH or SAM also forms a reverse Hoogsteen base pair with one U in the ligand recognition (Figure 4D-E, 8D-E and 10E-F, Table 1). However, there's no (U-A)·U base triple formed in the binding pocket of SAM-SAH riboswitch that contribute to the recognition of SAM in SAM-II riboswitch family (Figure 4G and 8G, Table 1), thus resulting in no selectivity between SAM and SAH for SAM-SAH riboswitch.

Though the conserved sequence and secondary structure of SAM-III riboswitch and SAM-VI riboswitch exhibit some similarities (Mirihana Arachchilage et al., 2018), the tertiary structural research confirm that they belong to distinct riboswitch families (Lu et al., 2008; Sun et al., 2019a). As shown in Figure 12C-D, both SAM-III riboswitch and SAM-VI riboswitch adopt similar overall architecture; however, stems P1 and P3 switch positions to stack with stem P2 (Figures 5, 9). Besides, the ligand interaction patterns are totally different (Figures 5, 9). SAH riboswitch adopts a unique overall folding and the ligand recognition pattern shares no similarities with other SAM-sensing riboswitch (Figure 11 and Table 1).

All the tertiary structural studies of SAM-sensing riboswitches provided the comprehensive structural information to investigate the specific ligand recognition and discrimination mechanisms of each SAM-sensing riboswitch and established the essential structural basis for comparing and further classifying SAM riboswitches. Given the importance of SAM, these researches may also facilitate the discovery of new types of SAM riboswitches. Since all the currently identified SAM riboswitches originated from bacteria, it is an interesting question to investigate whether there are more SAM riboswitches exist in the eukaryotic genome as well as TPP riboswitch (Wachter, 2010; Wachter et al., 2007).

As mentioned previously, all these SAM-sensing riboswitches, except SAM-IV riboswitch, were solved using the X-ray crystallography method. SAM-IV riboswitch was solved with cryo-EM method. Currently, more than 62% of the RNA structures were solved with X-ray crystallography method in the PDB bank (<https://www.rcsb.org>). The use of cryo-EM method in SAM-IV riboswitch also provides a promising way to determine the structure of RNA molecules (Zhang et al., 2019). RNA structure determination and prediction are the fundamental scientific questions in RNA biology, which remains challenging today. Within each SAM superfamily, different types of RNA molecules with solved tertiary structures possess the same SAM-binding pockets, which offer an ideal models to study the folding principles of RNA molecules. More extensive structural research work on RNA molecules including SAM riboswitches, will shed light on RNA folding theory and advance the development of RNA structure prediction.

The high affinity and selectivity of SAM riboswitches for SAM binding has enabled the development of RNA devices for SAM detection *in vivo and in vitro*. By fusion with Spinach aptamer or Corn aptamer, fluorescence imaging of cellular SAM was achieved via SAM-III riboswitch (Kim and Jaffrey, 2019; Paige et al., 2012). In another case, SAM-I riboswitch was also integrated with the Spinach aptamer to provide an alternative topology when designing fluorescent biosensors (Truong et al., 2018). These SAM riboswitches-based potential applications of SAM detection would greatly propel the research in SAM metabolism.

As ancient highly conserved biological molecules, riboswitches have evolved with high binding capacity to interact with small molecules during their long-term biological history, which enable riboswitches as an ideal RNA-drug targets. It is notable that most riboswitches were identified in bacteria, and thus since the discovery of riboswitches, they have been proposed as promising new antibacterial drug targets with many advantaging prospects (Blount and Breaker, 2006; Deigan and Ferre-D'Amare, 2011; Penchovsky et al., 2021). In this regard, a compound termed Ribocil was discovered to target FMN riboswitch and successfully suppress the bacterial growth (Howe et al., 2015). Due to the importance of SAM, SAM riboswitches are among the most abundant riboswitches and are widely distributed in bacteria, especially in many human pathogens (Breaker, 2010; Corbino et al., 2005; Fuchs et al., 2006; Grundy and Henkin, 1998; Hickey and Hammond, 2014; Meyer et al., 2009; Mirihana Arachchilage et al., 2018; Weinberg et al., 2007; Weinberg et al., 2010; Winkler et al., 2003). These SAM riboswitches regulate crucial metabolism pathways and expression of essential genes in pathogens, thus making them perfect targets for new antibiotics. The current structural information of SAM riboswitches complexed with SAM provides a comprehensive structural basis for rational design of SAM analogues or other compounds targeting SAM riboswitches, which will promote the development of riboswitch-targeting antibiotics and increase the possibility for the split-new drug design.

As the ubiquitous methyl donor in cell, SAM is utilized by methyltransferase proteins as protein cofactor for methylation modification (Kozbial and Mushegian, 2005). To investigate the possibility of SAM being used as cofactor for ribozyme, in a recent study, one SAM-dependent methyltransferase ribozyme RNA motif was

identified from a random RNA pool using SELEX method (Jiang et al., 2021). Further comparisons indicate that the RNA motif also exists in natural sequences, which reinforces the possibility that the SAM-dependent methyltransferase ribozyme may also contribute to the cellular methylation (Jiang et al., 2021). Currently, only one natural RNA motif, termed *glmS*, has been shown to have a dual role as both riboswitch and ribozyme (Winkler et al., 2004). *glmS* motif can bind to glucosamine-6-phosphate (GlcN6P), thereby promoting the self-cleavage activity of the *glmS* motif, which in turn reduces the expression of downstream genes (Klein and Ferre-D'Amare, 2006; Winkler et al., 2004). SAM is the one of the most ubiquitous cofactors in nature. It is an intriguing scientific question to detect whether more RNA motifs can both sense SAM (SAM riboswitches) and use SAM as a substrate to achieve RNA methylation (SAM-dependent methyltransferase ribozyme) or to perform new biological functions.

Riboswitches regulate the expression of downstream genes through the conformational change of RNA molecules, which provide valuable hints to demystify the working mechanism of other RNA molecules, especially for some non-coding RNAs. Non-coding RNA (ncRNA) is usually referred to the RNA segment that does not code proteins. A rising number of ncRNA molecules have been identified playing fundamental roles in diverse processes in life, such as microRNA (miRNA) (Chang et al., 2022), small interfering RNA (siRNA) (Saw and Song, 2020), circular RNAs (circRNAs) (Li et al., 2021), long non-coding RNAs (lncRNAs) (WEI et al., 2013), small nuclear RNAs (snRNAs) and piwi-interacting RNAs (piRNAs) (Dai et al., 2020; Saw et al., 2021; Xu et al., 2022; Xue et al., 2020). Tertiary structural investigation of these ncRNA molecules or their protein complex may advance the progress from function study to clinical application (Yang et al., 2020).

Compliance and ethics

The author(s) declare that they have no conflict of interest.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (32022039, 31870810, 91940302, and 91640104 to A.R.), the National Key Research

and Development Project of China (2021YFC2300300), the China Postdoctoral Science Foundation(2022M713637 to L.Z.), the outstanding youth fund of Zhejiang Province (LR19C050003 to A.R.) and the Fundamental Research Funds for the Central Universities (2017QN81010 to A.R.).

Accepted

Reference

- Blount, K.F., and Breaker, R.R. (2006). Riboswitches as antibacterial drug targets. *Nature biotechnology* 24, 1558-1564.
- Boyapati, V.K., Huang, W., Spedale, J., and Aboul-Ela, F. (2012). Basis for ligand discrimination between ON and OFF state riboswitch conformations: the case of the SAM-I riboswitch. *Rna* 18, 1230-1243.
- Breaker, R. (2022). The Biochemical Landscape of Riboswitch Ligands. *Biochemistry* 61.
- Breaker, R.R. (2010). Riboswitches and the RNA World. *Cold Spring Harbor perspectives in biology* 4, a003566-a003566.
- Caron, M.P., Bastet, L., Lussier, A., Simoneau-Roy, M., Masse, E., and Lafontaine, D.A. (2012). Dual-acting riboswitch control of translation initiation and mRNA decay. *Proceedings of the National Academy of Sciences of the United States of America* 109, E3444-3453.
- Chang, G., Ziliang, C., Yaping, Y., Jun, S., Yingying, Z., Li, L., Wanyi, L., Yuzhi, G., Zhihua, L., Weige, T., Chushan, Z., Wenbo, Z., Jiajie, Z., Xiang, Z., Yunjie, Z., Qiang, L., Stephanie, H., L., K.A., S., Y.E., François, B., Francesco, R., Armando, O., Gianluca, F., Kazuaki, T., Suzanne, K., Naohiro, I., Angela, T., P., T.M., A, C.M., and Erwei, S. (2022). A 10-miRNA risk score-based prediction model for pathological complete response to neoadjuvant chemotherapy in hormone receptor-positive breast cancer. *SCIENCE CHINA Life Sciences*.
- Cheah, M.T., Wachter, A., Sudarsan, N., and Breaker, R.R. (2007). Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature* 447, 497-500.
- Chen, B., LeBlanc, R., and Dayie, T.K. (2016). SAM-II Riboswitch Samples at least Two Conformations in Solution in the Absence of Ligand: Implications for Recognition. *Angewandte Chemie* 55, 2724-2727.
- Chen, B., Zuo, X., Wang, Y.X., and Dayie, T.K. (2012). Multiple conformations of SAM-II riboswitch detected with SAXS and NMR spectroscopy. *Nucleic acids research* 40, 3117-3130.
- Corbino, K.A., Barrick, J.E., Lim, J., Welz, R., Tucker, B.J., Puskarz, 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 biology* 6, R70.
- Dai, P., Wang, X., and Liu, M.-F. (2020). A dual role of the PIWI/piRNA machinery in regulating mRNAs during mouse spermiogenesis. *SCIENCE CHINA Life Sciences* 63, 447.
- Deigan, K.E., and Ferre-D'Amare, A.R. (2011). Riboswitches: discovery of drugs that target bacterial gene-regulatory RNAs. *Accounts of chemical research* 44, 1329-1338.
- Doshi, U., Kelley, J.M., and Hamelberg, D. (2011). Atomic-level insights into metabolite recognition and specificity of the SAM-II riboswitch. *Rna* 18, 300-307.
- Doshi, U., Kelley, J.M., and Hamelberg, D. (2012). Atomic-level insights into metabolite recognition and specificity of the SAM-II riboswitch. *Rna* 18, 300-307.
- Edwards, A.L., Reyes, F.E., Heroux, A., and Batey, R.T. (2010). Structural basis for recognition of S-adenosylhomocysteine by riboswitches. *Rna* 16, 2144-2155.
- Eschbach, S.H., St-Pierre, P., Penedo, J.C., and Lafontaine, D.A. (2012). Folding of the SAM-I riboswitch: a tale with a twist. *RNA biology* 9, 535-541.
- Fontecave, M., Atta, M., and Mulliez, E. (2004). S-adenosylmethionine: nothing goes to waste. *Trends in Biochemical Sciences* 29, 243-249.

- Fuchs, R.T., Grundy, F.J., and Henkin, T.M. (2006). The S(MK) box is a new SAM-binding RNA for translational regulation of SAM synthetase. *Nature structural & molecular biology* 13, 226-233.
- Gilbert, S.D., Rambo, R.P., Van Tyne, D., and Batey, R.T. (2008). Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nature structural & molecular biology* 15, 177-182.
- Gong, S., Wang, Y., Wang, Z., Wang, Y., and Zhang, W. (2016). Reversible-Switch Mechanism of the SAM-III Riboswitch. *The journal of physical chemistry B* 120, 12305-12311.
- Grundy, F.J., and Henkin, T.M. (1998). The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in gram-positive bacteria. *Molecular microbiology* 30, 737-749.
- Haller, A., Rieder, U., Aigner, M., Blanchard, S.C., and Micura, R. (2011a). Conformational capture of the SAM-II riboswitch. *Nature chemical biology* 7, 393-400.
- Haller, A., Souliere, M.F., and Micura, R. (2011b). The dynamic nature of RNA as key to understanding riboswitch mechanisms. *Accounts of chemical research* 44, 1339-1348.
- Halliday, N.M., Hardie, K.R., Williams, P., Winzer, K., and Barrett, D.A. (2010). Quantitative liquid chromatography-tandem mass spectrometry profiling of activated methyl cycle metabolites involved in LuxS-dependent quorum sensing in *Escherichia coli*. *Analytical biochemistry* 403, 20-29.
- Hickey, S.F., and Hammond, M.C. (2014). Structure-guided design of fluorescent S-adenosylmethionine analogs for a high-throughput screen to target SAM-I riboswitch RNAs. *Chemistry & biology* 21, 345-356.
- Hilbers, C.W., Michiels, P.J.A., and Heus, H.A. (1998). New developments in structure determination of pseudoknots. *Biopolymers* 48, 137-153.
- Howe, J.A., Wang, H., Fischmann, T.O., Balibar, C.J., Xiao, L., Galgoci, A.M., Malinverni, J.C., Mayhood, T., Villafania, A., Nahvi, A., Murgolo, N., Barbieri, C.M., Mann, P.A., Carr, D., Xia, E., Zuck, P., Riley, D., Painter, R.E., Walker, S.S., Sherborne, B., de Jesus, R., Pan, W., Plotkin, M.A., Wu, J., Rindgen, D., Cummings, J., Garlisi, C.G., Zhang, R., Sheth, P.R., Gill, C.J., Tang, H., and Roemer, T. (2015). Selective small-molecule inhibition of an RNA structural element. *Nature* 526, 672-677.
- Huang, L., Liao, T.W., Wang, J., Ha, T., and Lilley, D.M.J. (2020). Crystal structure and ligand-induced folding of the SAM/SAH riboswitch. *Nucleic acids research* 48, 7545-7556.
- Huang, L., and Lilley, D.M.J. (2018). Structure and ligand binding of the SAM-V riboswitch. *Nucleic acids research* 46, 6869-6879.
- Huang, W., Kim, J., Jha, S., and Aboul-Ela, F. (2012). Conformational heterogeneity of the SAM-I riboswitch transcriptional ON state: a chaperone-like role for S-adenosyl methionine. *Journal of molecular biology* 418, 331-349.
- Jiang, H., Gao, Y., Zhang, L., Chen, D., Gan, J., and Murchie, A.I.H. (2021). The identification and characterization of a selected SAM-dependent methyltransferase ribozyme that is present in natural sequences. *Nature Catalysis* 4, 872-881.
- Kim, H., and Jaffrey, S.R. (2019). A Fluorogenic RNA-Based Sensor Activated by Metabolite-Induced RNA Dimerization. *Cell chemical biology* 26, 1725-1731 e1726.
- Klein, D.J., and Ferre-D'Amare, A.R. (2006). Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. *Science* 313, 1752-1756.
- Kozbial, P.Z., and Mushegian, A.R. (2005). Natural history of S-adenosylmethionine-binding proteins. *BMC structural biology* 5, 19.

- 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. *Scientific reports* 5, 16641.
- Li, X., Zhang, J.-L., Lei, Y.-N., Liu, X.-Q., Xue, W., Zhang, Y., Nan, F., Gao, X., Zhang, J., Wei, J., Yang, L., and Chen, L.-L. (2021). Linking circular intronic RNA degradation and function in transcription by RNase H1. *SCIENCE CHINA Life Sciences* 64, 1795.
- Lim, J., Winkler, W.C., Nakamura, S., Scott, V., and Breaker, R.R. (2006). Molecular-recognition characteristics of SAM-binding riboswitches. *Angewandte Chemie* 45, 964-968.
- Lu, C., Ding, F., Chowdhury, A., Pradhan, V., Tomsic, J., Holmes, W.M., Henkin, T.M., and Ke, A. (2010). SAM recognition and conformational switching mechanism in the *Bacillus subtilis* yitJ S box/SAM-I riboswitch. *Journal of molecular biology* 404, 803-818.
- Lu, C., Smith, A.M., Ding, F., Chowdhury, A., Henkin, T.M., and Ke, A. (2011). Variable sequences outside the SAM-binding core critically influence the conformational dynamics of the SAM-III/SMK box riboswitch. *Journal of molecular biology* 409, 786-799.
- Lu, C., Smith, A.M., Fuchs, R.T., Ding, F., Rajashankar, K., Henkin, T.M., and Ke, A. (2008). Crystal structures of the SAM-III/S(MK) riboswitch reveal the SAM-dependent translation inhibition mechanism. *Nature structural & molecular biology* 15, 1076-1083.
- Manz, C., Kobitski, A.Y., Samanta, A., Jaschke, A., and Nienhaus, G.U. (2018). The multi-state energy landscape of the SAM-I riboswitch: A single-molecule Forster resonance energy transfer spectroscopy study. *The Journal of chemical physics* 148, 123324.
- Manz, C., Kobitski, A.Y., Samanta, A., Keller, B.G., Jaschke, A., and Nienhaus, G.U. (2017). Single-molecule FRET reveals the energy landscape of the full-length SAM-I riboswitch. *Nature chemical biology* 13, 1172-1178.
- Meyer, M.M., Ames, T.D., Smith, D.P., Weinberg, Z., Schwalbach, M.S., Giovannoni, S.J., and Breaker, R.R. (2009). Identification of candidate structured RNAs in the marine organism 'Candidatus Pelagibacter ubique'. *BMC genomics* 10, 268.
- Mirihana Arachchilage, G., 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 biology* 15, 371-378.
- Montange, R.K., and Batey, R.T. (2006). Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature* 441, 1172-1175.
- Montange, R.K., Mondragon, E., van Tyne, D., Garst, A.D., Ceres, P., and Batey, R.T. (2010). Discrimination between closely related cellular metabolites by the SAM-I riboswitch. *Journal of molecular biology* 396, 761-772.
- Paige, J.S., Nguyen-Duc, T., Song, W., and Jaffrey, S.R. (2012). Fluorescence imaging of cellular metabolites with RNA. *Science* 335, 1194.
- Parveen, N., and Cornell, K.A. (2011). Methylthioadenosine/S-adenosylhomocysteine nucleosidase, a critical enzyme for bacterial metabolism. *Molecular microbiology* 79, 7-20.
- Penchovsky, R., Pavlova, N., and Kaloudas, D. (2021). RSwitch: A Novel Bioinformatics Database on Riboswitches as Antibacterial Drug Targets. *IEEE/ACM Trans Comput Biol Bioinform* 18, 804-808.
- Poiata, E., Meyer, M.M., Ames, T.D., and Breaker, R.R. (2009). A variant riboswitch aptamer class for S-adenosylmethionine common in marine bacteria. *Rna* 15, 2046-2056.
- Price, I.R., Grigg, J.C., and Ke, A. (2014). Common themes and differences in SAM recognition among SAM riboswitches. *Biochimica et biophysica acta* 1839, 931-938.
- Priyakumar, U.D. (2010). Atomistic details of the ligand discrimination mechanism of S(MK)/SAM-III riboswitch. *The journal of physical chemistry B* 114, 9920-9925.

- Saw, P.E., and Song, E.-W. (2020). siRNA therapeutics: a clinical reality. *SCIENCE CHINA Life Sciences* 63, 485.
- Saw, P.E., Xu, X., Chen, J., and Song, E.-W. (2021). Non-coding RNAs: the new central dogma of cancer biology. *SCIENCE CHINA Life Sciences* 64, 22.
- Schroeder, K.T., Daldrop, P., and Lilley, D.M. (2011). RNA tertiary interactions in a riboswitch stabilize the structure of a kink turn. *Structure* 19, 1233-1240.
- Serganov, A., and Nudler, E. (2013). A decade of riboswitches. *Cell* 152, 17-24.
- Smith, A.M., Fuchs, R.T., Grundy, F.J., and Henkin, T.M. (2010). The SAM-responsive S(MK) box is a reversible riboswitch. *Molecular microbiology* 78, 1393-1402.
- Stoddard, C.D., Montange, R.K., Hennelly, S.P., Rambo, R.P., Sanbonmatsu, K.Y., and Batey, R.T. (2010). Free state conformational sampling of the SAM-I riboswitch aptamer domain. *Structure* 18, 787-797.
- Struck, A.W., Thompson, M.L., Wong, L.S., and Micklefield, J. (2012). S-adenosylmethionine-dependent methyltransferases: highly versatile enzymes in biocatalysis, biosynthesis and other biotechnological applications. *Chembiochem : a European journal of chemical biology* 13, 2642-2655.
- Sun, A., Gasser, C., Li, F., Chen, H., Mair, S., Krasheninina, O., Micura, R., and Ren, A. (2019a). SAM-VI riboswitch structure and signature for ligand discrimination. *Nat Commun* 10, 5728.
- Sun, A., Gasser, C., Li, F., Chen, H., Mair, S., Krasheninina, O., Micura, R., and Ren, A. (2019b). SAM-VI riboswitch structure and signature for ligand discrimination. *Nature Communications* 10.
- Suresh, G., Srinivasan, H., Nanda, S., and Priyakumar, U.D. (2016). Ligand-Induced Stabilization of a Duplex-like Architecture Is Crucial for the Switching Mechanism of the SAM-III Riboswitch. *Biochemistry* 55, 3349-3360.
- Tang, D.J., Du, X., Shi, Q., Zhang, J.L., He, Y.P., Chen, Y.M., Ming, Z., Wang, D., Zhong, W.Y., Liang, Y.W., Liu, J.Y., Huang, J.M., Zhong, Y.S., An, S.Q., Gu, H., and Tang, J.L. (2020). A SAM-I riboswitch with the ability to sense and respond to uncharged initiator tRNA. *Nat Commun* 11, 2794.
- Trausch, J.J., Xu, Z., Edwards, A.L., Reyes, F.E., Ross, P.E., Knight, R., and Batey, R.T. (2014). Structural basis for diversity in the SAM clan of riboswitches. *Proceedings of the National Academy of Sciences of the United States of America* 111, 6624-6629.
- Truong, J., Hsieh, Y.F., Truong, L., Jia, G., and Hammond, M.C. (2018). Designing fluorescent biosensors using circular permutations of riboswitches. *Methods* 143, 102-109.
- Tucker, B.J., and Breaker, R.R. (2005). Riboswitches as versatile gene control elements. *Current opinion in structural biology* 15, 342-348.
- Wachter, A. (2010). Riboswitch-mediated control of gene expression in eukaryotes. *RNA biology* 7, 67-76.
- Wachter, A., Tunc-Ozdemir, M., Grove, B.C., Green, P.J., Shintani, D.K., and Breaker, R.R. (2007). Riboswitch control of gene expression in plants by splicing and alternative 3' end processing of mRNAs. *The Plant cell* 19, 3437-3450.
- Wang, J.X., and Breaker, R.R. (2008). Riboswitches that sense S-adenosylmethionine and S-adenosylhomocysteine. *Biochemistry and cell biology = Biochimie et biologie cellulaire* 86, 157-168.
- 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. *Molecular cell* 29, 691-702.

- WEI, P., HAN, B., and CHEN, Y. (2013). Role of long non-coding RNAs in normal and malignant hematopoiesis. *SCIENCE CHINA Life Sciences* 56, 867.
- Weickhmann, A.K., Keller, H., Wurm, J.P., Strebiter, E., Juen, M.A., Kremser, J., Weinberg, Z., Kreutz, C., Duchardt-Ferner, E., and Wohnert, J. (2019). The structure of the SAM/SAH-binding riboswitch. *Nucleic acids research* 47, 2654-2665.
- Weinberg, Z., Barrick, J.E., Yao, Z., Roth, A., Kim, J.N., Gore, J., Wang, J.X., Lee, E.R., Block, K.F., Sudarsan, N., Neph, S., Tompa, M., Ruzzo, W.L., and Breaker, R.R. (2007). Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic acids research* 35, 4809-4819.
- Weinberg, Z., Regulski, E.E., Hammond, M.C., Barrick, J.E., Yao, Z., Ruzzo, W.L., and Breaker, R.R. (2008). The aptamer core of SAM-IV riboswitches mimics the ligand-binding site of SAM-I riboswitches. *Rna* 14, 822-828.
- 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 biology* 11, R31.
- Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A., and Breaker, R.R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 428, 281-286.
- Winkler, W.C., Nahvi, A., Sudarsan, N., Barrick, J.E., and Breaker, R.R. (2003). An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nature structural biology* 10, 701-707.
- Xu, B., Zhu, Y., Cao, C., Chen, H., Jin, Q., Li, G., Ma, J., Yang, S.L., Zhao, J., Zhu, J., Ding, Y., Fang, X., Jin, Y., Kwok, C.K., Ren, A., Wan, Y., Wang, Z., Xue, Y., Zhang, H., Zhang, Q.C., and Zhou, Y. (2022). Recent advances in RNA structure. *Science China Life Sciences* 65, 1285-1324.
- Xue, X., Yongjun, W., and Zhihong, L. (2015). Folding of SAM-II riboswitch explored by replica-exchange molecular dynamics simulation. *Journal of theoretical biology* 365, 265-269.
- Xue, Y., Chen, R., Qu, L., and Cao, X. (2020). Noncoding RNA: from dark matter to bright star. *SCIENCE CHINA Life Sciences* 63, 463.
- Yang, J., Li, J., Wang, J., Sheng, G., Wang, M., Zhao, H., Yang, Y., and Wang, Y. (2020). Crystal structure of Cas1 in complex with branched DNA. *SCIENCE CHINA Life Sciences* 63, 516.
- Yao, Z., Barrick, J., Weinberg, Z., Neph, S., Breaker, R., Tompa, M., and Ruzzo, W.L. (2007). A computational pipeline for high-throughput discovery of cis-regulatory noncoding RNA in prokaryotes. *PLoS computational biology* 3, e126.
- Zhang, K., Li, S., Kappel, K., Pintilie, G., Su, Z., Mou, T.C., Schmid, M.F., Das, R., and Chiu, W. (2019). Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. *Nat Commun* 10, 5511.

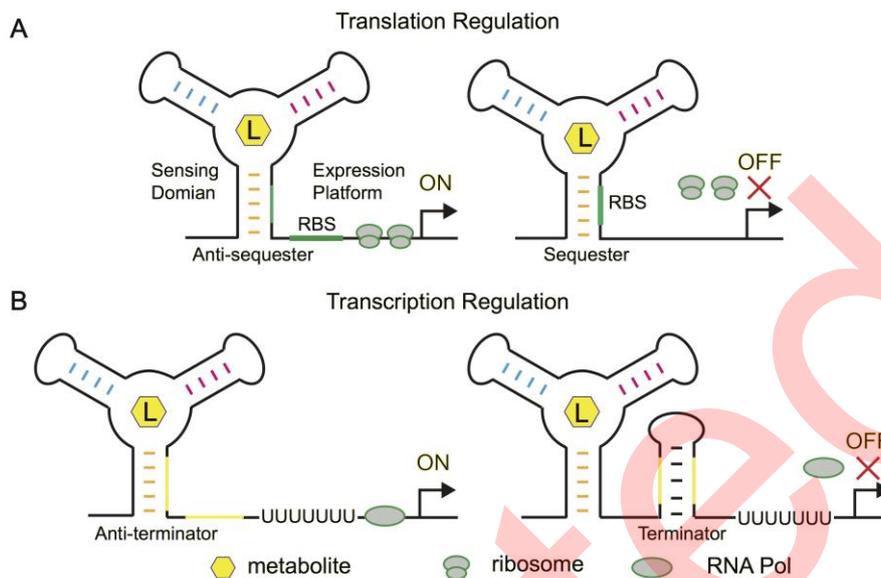


Figure 1 Gene expression regulation mechanisms of riboswitches. **A**, Downstream gene expression in translation process is regulated by ligand-induced conformational change, in which the anti-sequester stem or sequester stem is formed to initiate translation or inhibit translation by releasing or shading ribosome binding site (RBS) respectively. **B**, Upon ligand binding, the downstream gene expression in transcription process is regulated by forming anti-terminator stem or terminator stem to continue transcription or stop transcription by keeping or blocking RNA polymerase (RNA Pol) binding respectively.

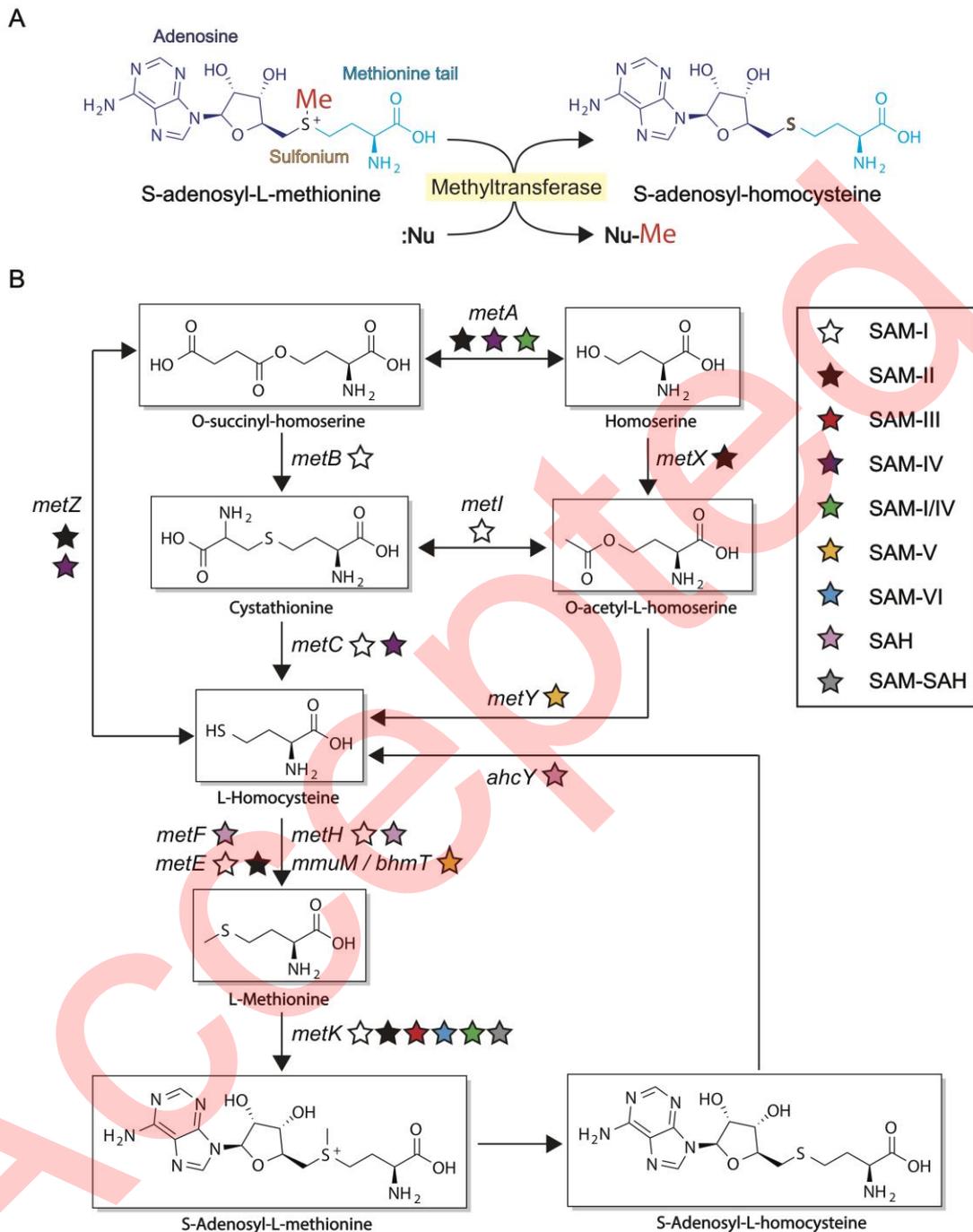


Figure 2 SAM-dependent methylation and SAM biosynthesis pathways. **A**, In the methylation reactions, the methyl group of SAM (Me) is transferred to a nucleophile substrate (:Nu) under catalysis of methyltransferases, in which the neutral compound *S*-adenosylhomocysteine is generated. **B**, Summary of SAM-sensing riboswitches and the downstream genes that are involved in SAM biosynthesis pathways.

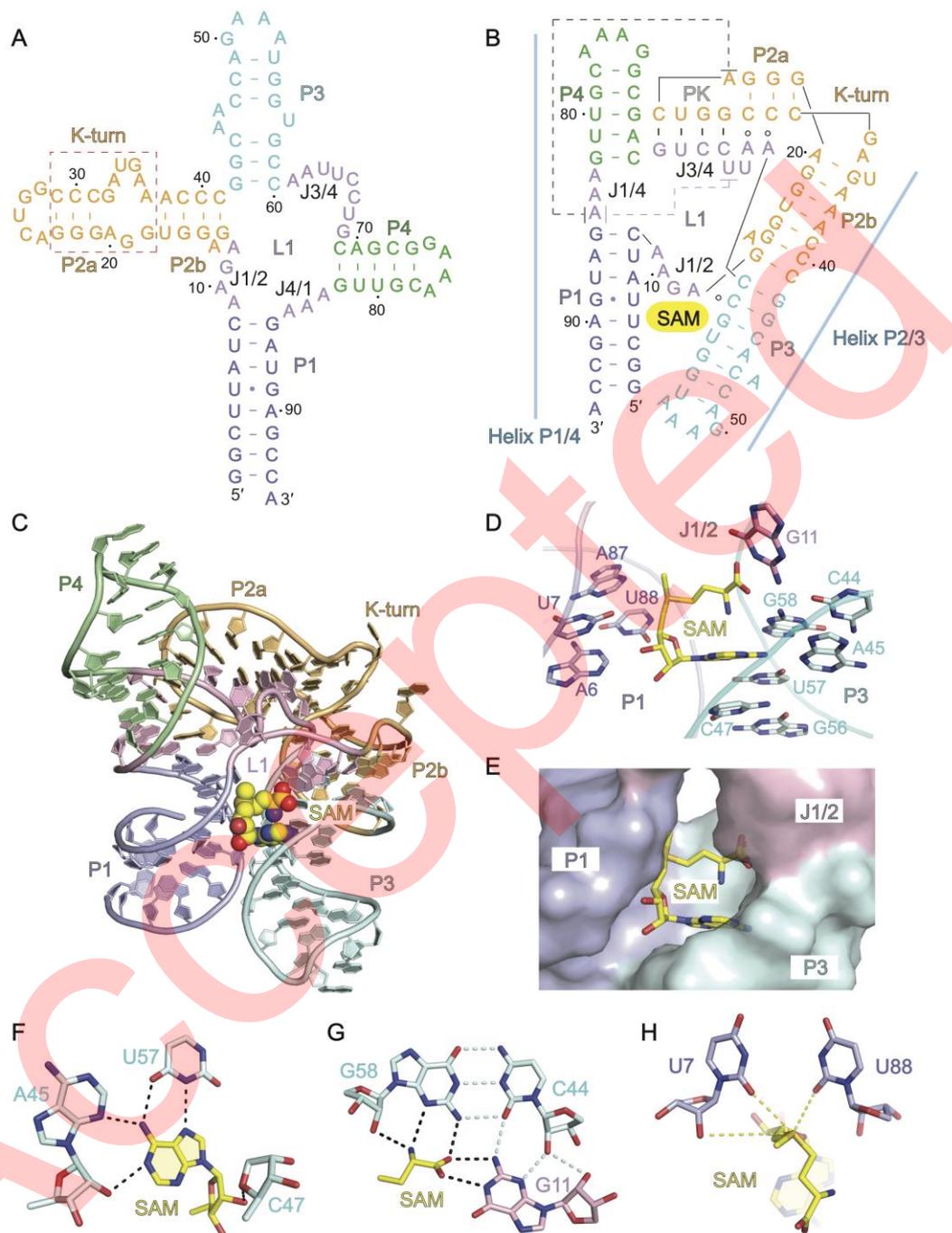


Figure 3 Structure and SAM-binding pocket of the *T. tengcogensis* SAM-I riboswitch. **A**, The sequence and secondary structure of SAM-I riboswitch used in structure determination. The color of all the residues is coded to the tertiary structure. **B**, The schematic folding topology of SAM-I riboswitch in complex with SAM. **C**, The tertiary structure of SAM-I riboswitch (shown in cartoon representation) complexed with SAM with SAM (shown as spheres). **D**, The composition of SAM binding pocket, in which the bases and SAM are shown in stick. **E**, SAM binding pocket shown in surface representation with SAM shown in stick. **F**, The adenine ring of SAM is recognized by A45 and U57 from stem P2. The ribose moiety of SAM forms one hydrogen bond with the ribose

of C47. **G**, The methionine moiety of SAM is recognized by (G58-C44)-G11 base triple. **H**, Electrostatic interaction between the positively charged sulfonium moiety of SAM and U7 and U88. Figures in panels **B-H** are depicted based on PDB 2GIS.

Accepted

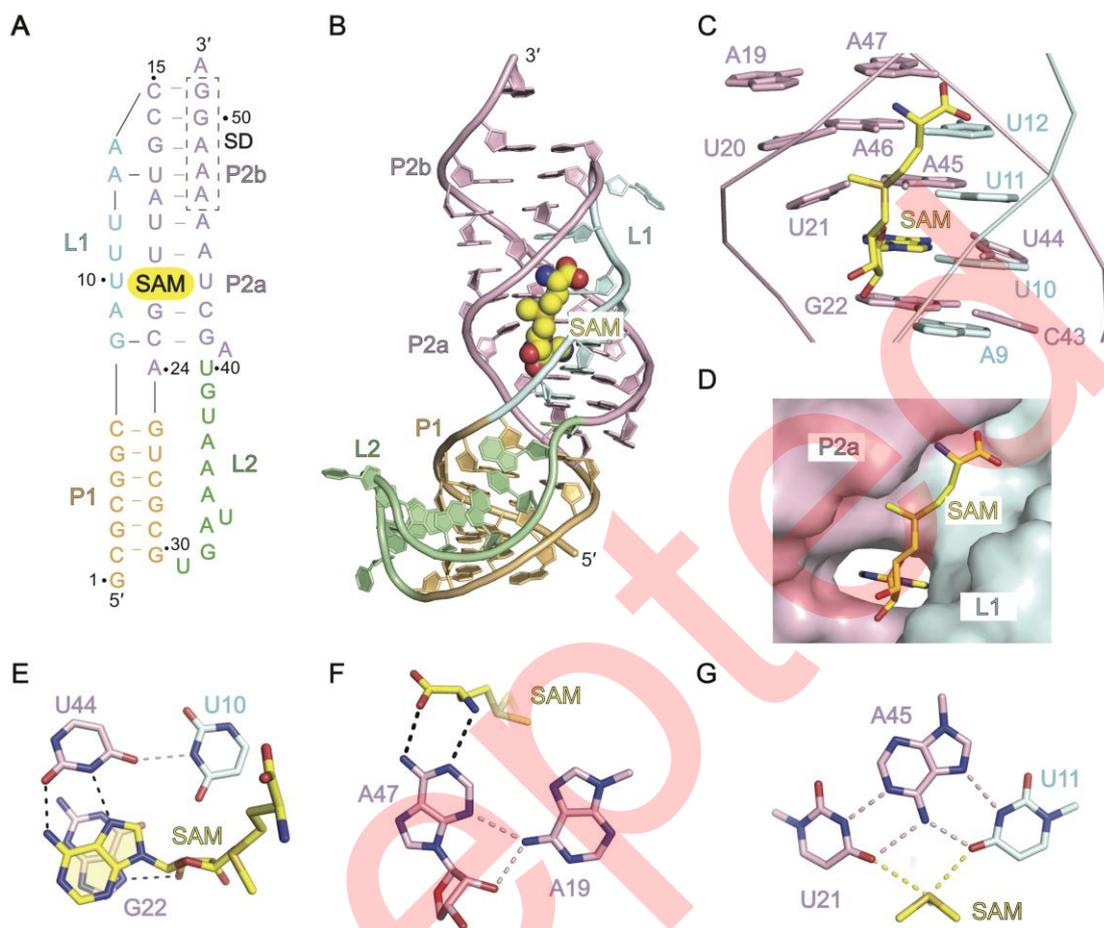


Figure 4 Structure and SAM-binding pocket of the *metX* SAM-II riboswitch. **A**, The Schematic folding topology of SAM-II riboswitch in complex with SAM. **B**, Overall structure of SAM-II riboswitch (shown in cartoon representation) in complex with SAM, in which SAM is shown as spheres. **C**, The composition of SAM binding pocket, in which the bases and SAM are shown in stick. **D**, SAM binding pocket shown in surface representation with SAM shown in stick. **E**, The adenine ring and ribose of SAM are recognized by U44 and G22. **F**, The methionine moiety of SAM is recognized by A47. A47 also forms hydrogen bonds with A19. **G**, Electrostatic interaction between the positively charged sulfonium moiety of SAM and O4 of U11 and U21. All the figures of SAM-II riboswitch are depicted based on PDB 2QWY.

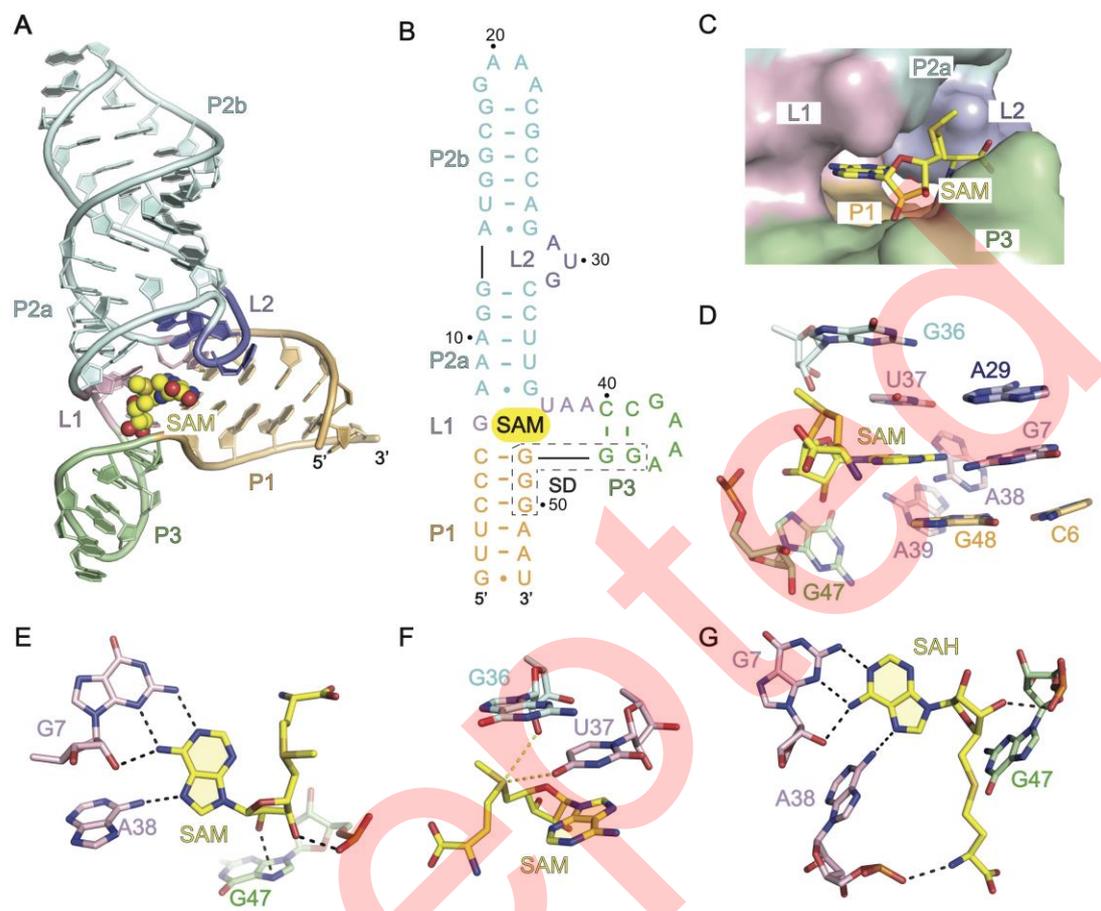


Figure 5 Structure and SAM-binding pocket of the *E. faecalis* SAM-III riboswitch. **A**, Tertiary structure of SAM-III riboswitch (shown in cartoon representation) complexed with SAM, in which SAM is shown as spheres. **B**, Schematic secondary structure of SAM-III riboswitch used in structure determination. **C-D**, The composition of SAM binding pocket of SAM-III riboswitch shown in surface representation (**C**) and sticks (**D**), in which SAM is shown as stick. **E**, The adenine ring of SAM is recognized by G7 and A38. The 2'-OH and 3'-OH of the ribose in SAM form hydrogen-bonding interactions with G47. **F**, Electrostatic interaction between the positively charged sulfonium moiety of SAM and O4 of U37 and 2'-OH of G36. **G**, SAH is stabilized by forming hydrogen bonds with G7, A38, and G47. All the figures in Panels **A-F** are depicted based on PDB 3E5C. Figure in panel **G** is depicted based on PDB 3E5E.

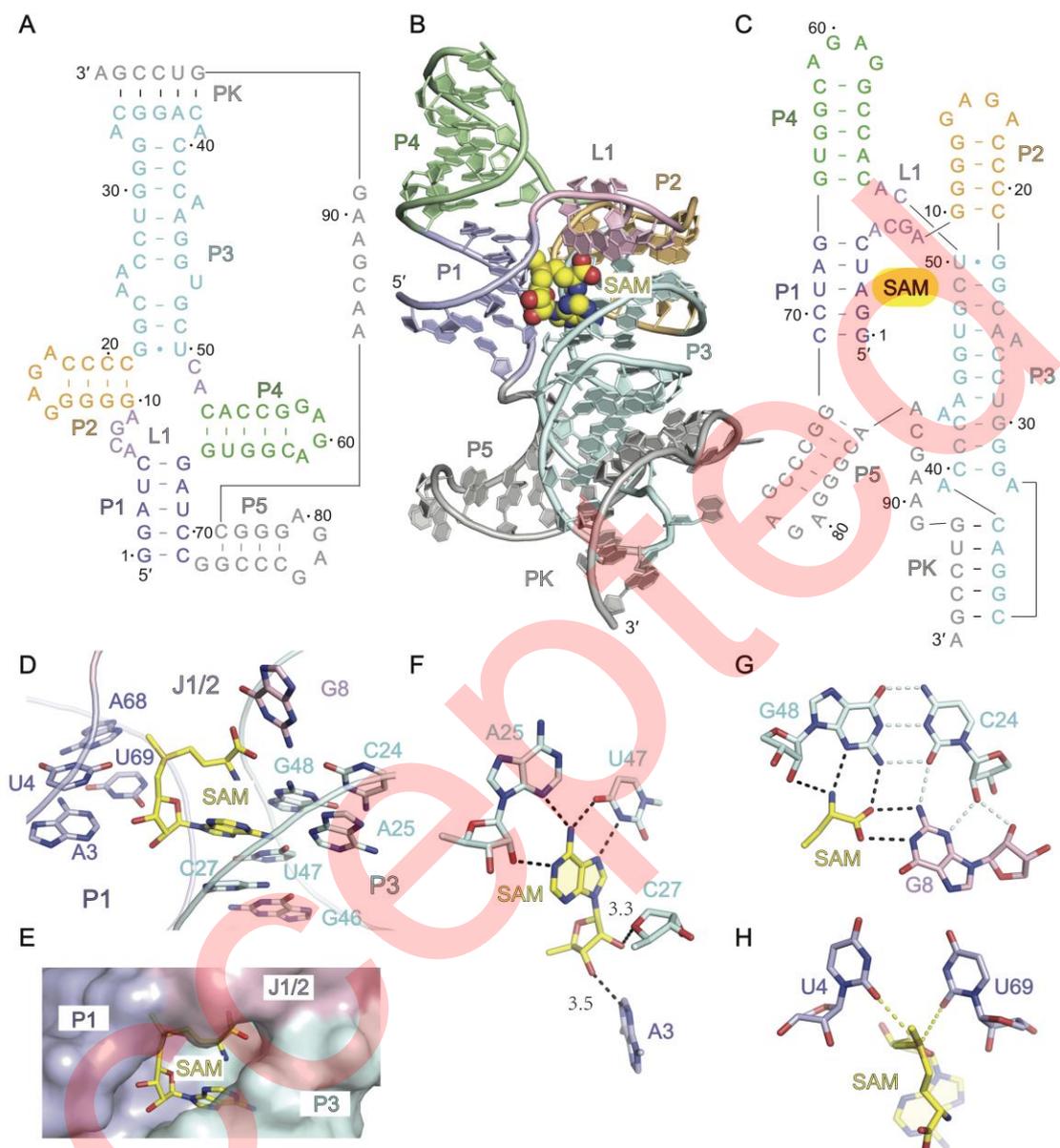


Figure 7 Structure and SAM-binding pocket of the *env87* SAM-I/IV riboswitch. **A**, The sequence and secondary structure of SAM-I/IV riboswitch used in structure determination. **B**, Tertiary structure of SAM-I/IV riboswitch (shown in cartoon representation) in complex with SAM (shown as spheres). **C**, The schematic topology of SAM-I/IV riboswitch bound with SAM. **D-E**, The composition of SAM binding pocket shown in surface representation (**D**) and sticks (**E**), in which SAM is shown as stick. **F**, The adenine ring of SAM is recognized by A25 and U47 from stem P3. The 2'-OH and 3'-OH of the ribose in SAM form hydrogen bonds with A3 and C27. **G**, The methionine moiety of SAM is recognized by (G48-C24)·G8 base triple. **H**, Electrostatic interaction between U4, U69, and the positively charged sulfonium moiety of SAM. All the figures of SAM-I/IV riboswitch are depicted based on PDB 4OQU.

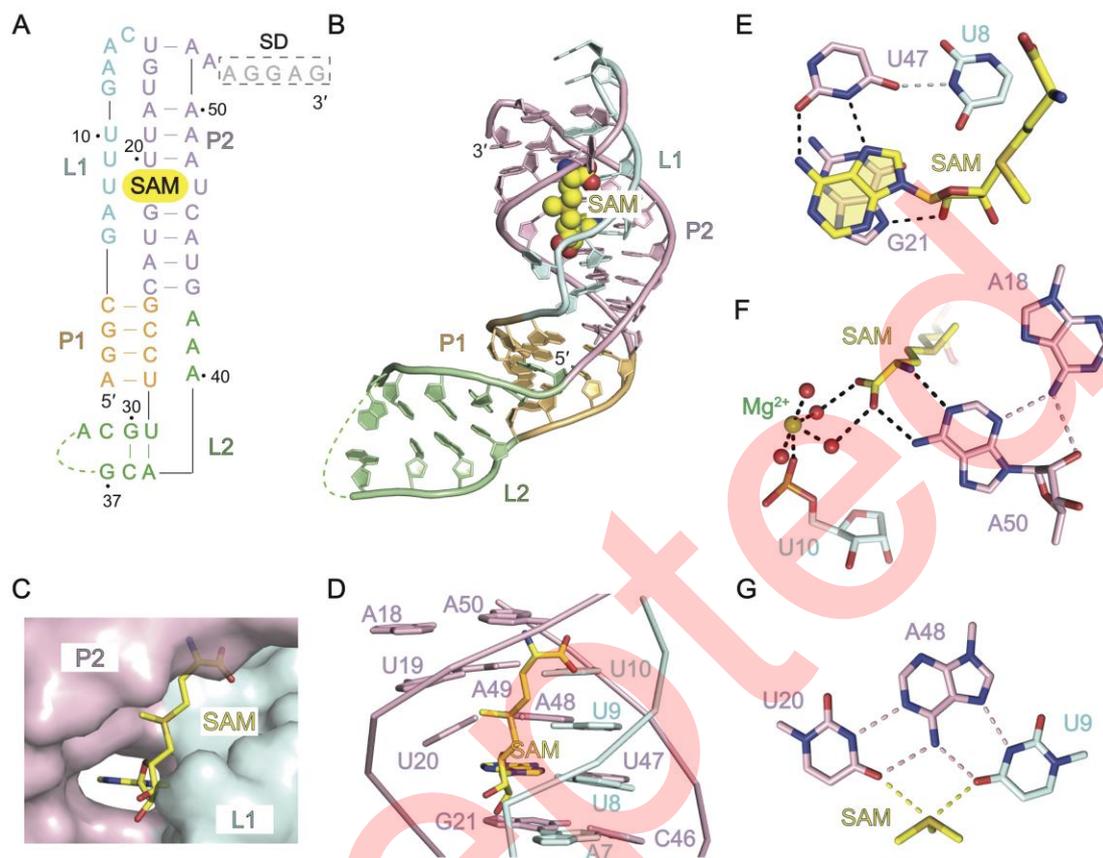


Figure 8 Structure and SAM-binding pocket of the *metY* SAM-V riboswitch. **A**, The schematic folding topology of SAM-V riboswitch in complex with SAM is depicted based on tertiary structure, in which the color of each residue is coded to the tertiary structure. **B**, Global folding of SAM-V riboswitch (shown in cartoon representation) in complex with SAM (shown as spheres). **C**, The surface representation of the cavity for SAM binding. **D**, The composition of SAM binding pocket with SAM shown in sticks. **E**, The adenine ring of SAM is recognized by U47. The 2'-OH of the ribose forms interaction with G21. **F**, The aminoacyl group of SAM forms hydrogen bonds with A50. One Mg^{2+} and the coordinated water molecules are also involved in interactions between methionine moiety of SAM and the riboswitch RNA. **G**, Electrostatic interaction between positively charged sulfonium moiety of SAM and O4 of U9 and U20. All the figures of SAM-II riboswitch are depicted based on PDB 6FZ0.

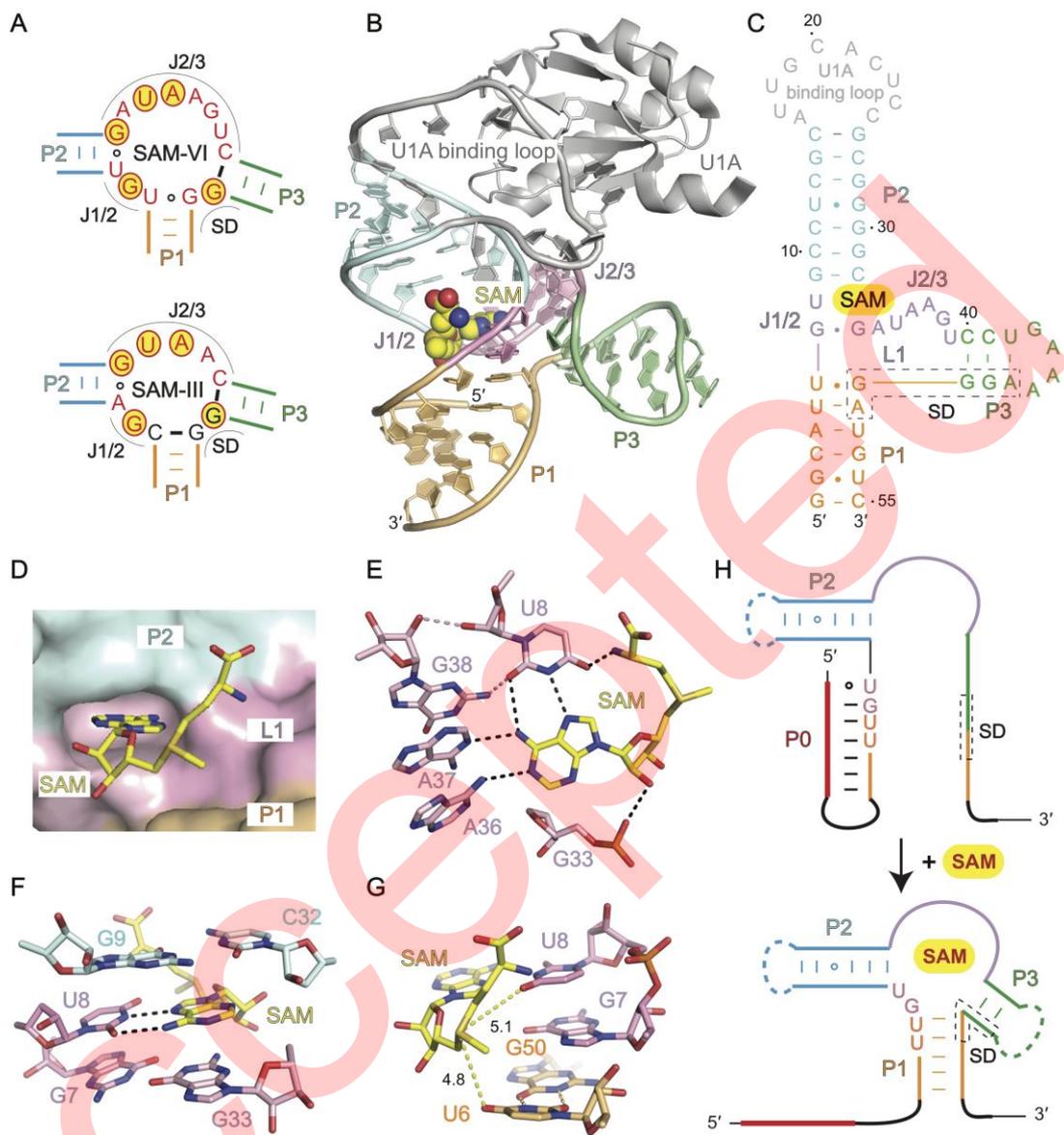


Figure 9 Structure and regulation mechanism of the SAM-VI riboswitch. **A**, Comparison of the three-way junction composition between SAM-III and SAM-VI riboswitches. All the highly conserved residues in junction are colored in red. The residues involved in SAM binding in SAM-III riboswitch and the predicted corresponding nucleotides in SAM-VI riboswitch are highlighted by yellow shadow circles. **B**, The tertiary structure of SAM-VI riboswitch (shown in cartoon representation) in complex with SAM (shown as spheres). The U1A-binding loop and the bound U1A protein to facilitate the crystallization are shown in gray. **C**, The schematic secondary structure of SAM-VI riboswitch depicted based on tertiary structure. **D**, The surface representation of the cavity for SAM binding. **E-F**, SAM forms hydrogen bonding interaction of with U8, A37, A38 and G33 in the binding pocket, in which the adenine of SAM pairs with U8 and become stacked between G9-C23 and G7·G33. **G**, Electrostatic interaction between the positively charged sulfonium moiety of the bound SAM and O4 of U6 and U8. **H**, Downstream gene expression regulation model of SAM-VI riboswitch. In absence of SAM, formation of P0 release the SD

sequence and activate the translation. While the binding of SAM induces the formation of stems P1 and P3, therefore inhibit translation by shading the SD sequence. Figures in panels **B-G** are depicted based on PDB 6LAS.

Accepted

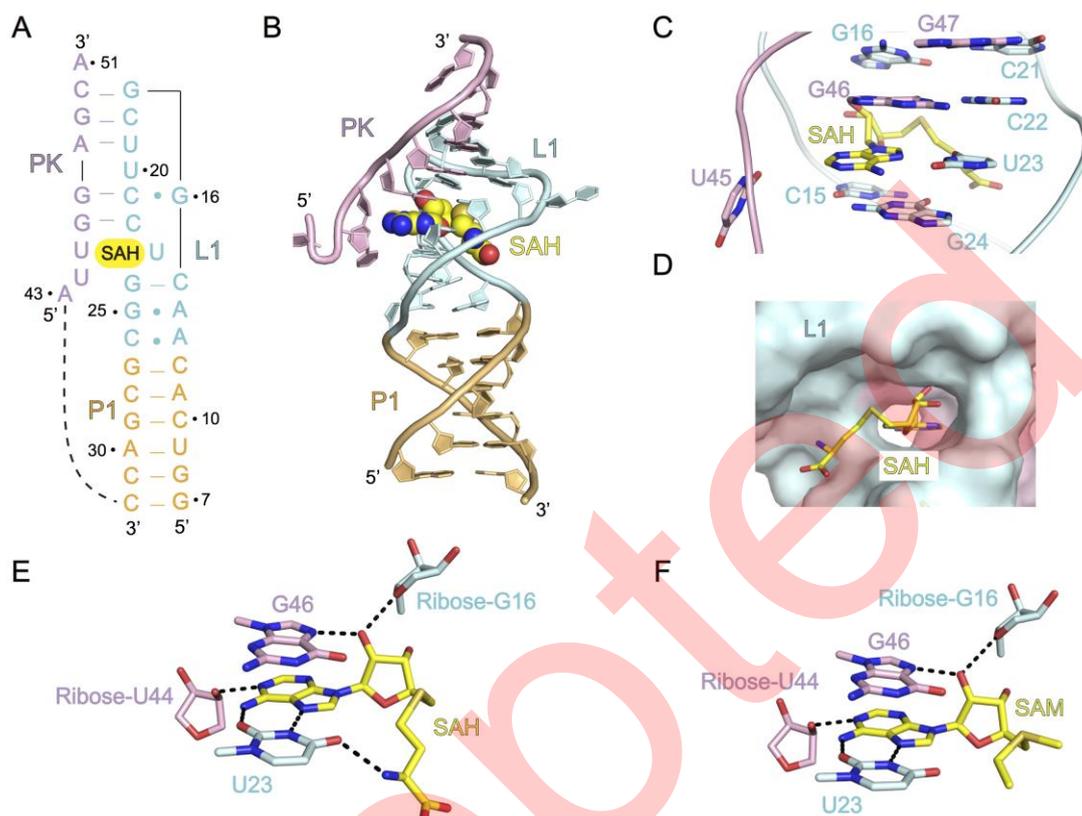


Figure 10 Structure and SAH/SAM-binding pocket of the SAM-SAH riboswitch. **A**, The schematic folding topology of SAM-SAH riboswitch in complex with SAH is depicted based on tertiary structure, in which the color of each residue is corresponding to the tertiary structure. **B**, Global folding of SAM-SAH riboswitch (shown in cartoon representation) in complex with SAH (shown as spheres). **C-D**, The composition of SAH binding pocket shown in sticks (**C**) and surface representation (**D**), in which SAH is shown as stick. **E**, The adenine ring of SAH pairs to U23 and the N1 of the adenine ring interacts with 2'-OH of ribose-U44. The 2'-OH of SAH form hydrogen bonds with N7 of G46 and O4' of ribose-G16. The α -amine of SAH is hydrogen-bonded to O4 of U23. **F**, The adenine ring of SAM pairs to U23 and the N1 of the adenine ring interacts with 2'-OH of ribose-U44. The 2'-OH of SAH make hydrogen bonds with N7 of G46 and O4' of ribose-G16. All the figures in panels **A-E** are depicted based on PDB 6YL5. Figure in panel **F** is depicted based on PDB 6YMM.

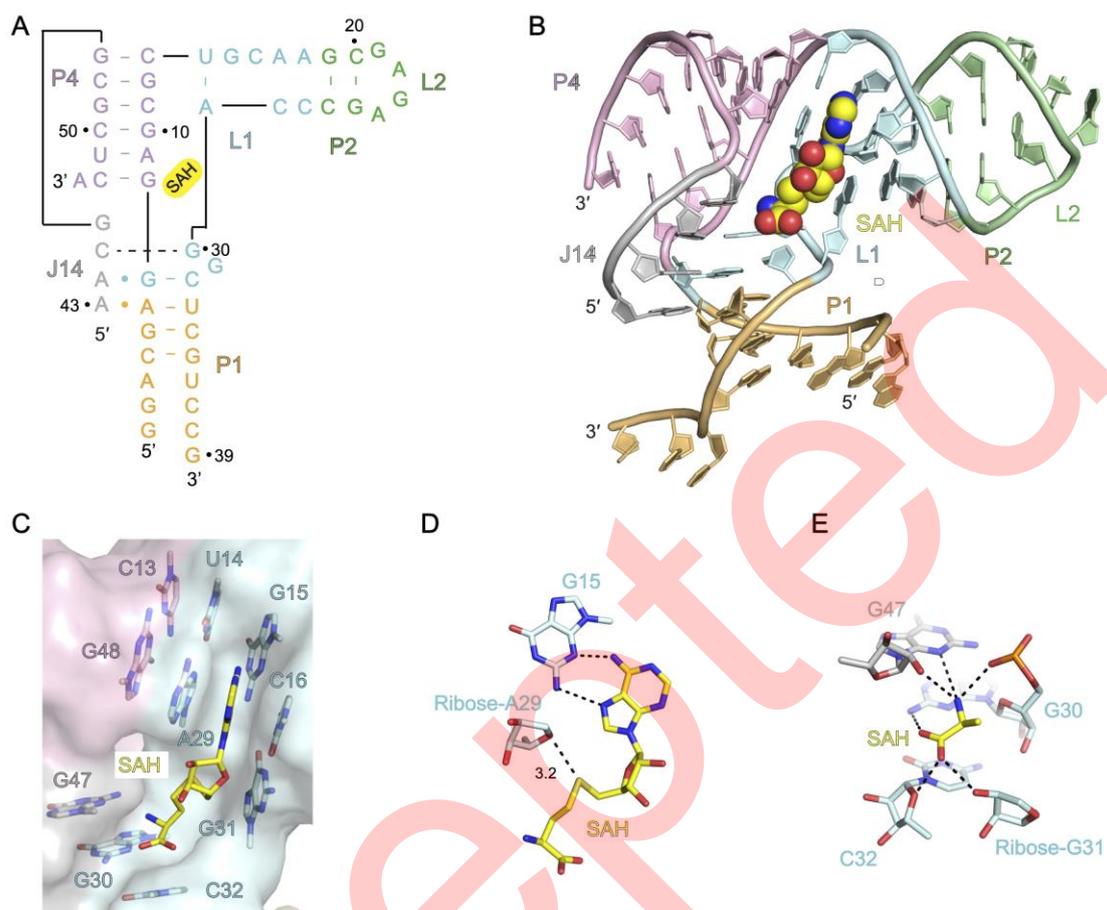


Figure 11 Structure and SAH-binding pocket of the SAH riboswitch. **A**, The schematic folding topology of SAH riboswitch in complex with SAH is depicted based on tertiary structure, in which the color of each residue is corresponding to the tertiary structure. **B**, Global folding of SAH riboswitch (shown in cartoon representation) in complex with SAH (shown as spheres). **C**, The surface representation and composition of the SAH binding cavity. **D**, The adenine ring of SAH pairs to G15. The distance between the sulfoether moiety of SAH and O4' of ribose-A29 is 3.2 Å. **E**, The carboxylate oxygens are hydrogen-bonded to the N2 of G30, the 2'-OH group of G31 and O4' of C32 ribose. The α -amine of SAH can contact with the phosphate of G31, the 2'-OH and the N3 of G47. All the figures of SAH riboswitch are depicted based on PDB 3NPQ.

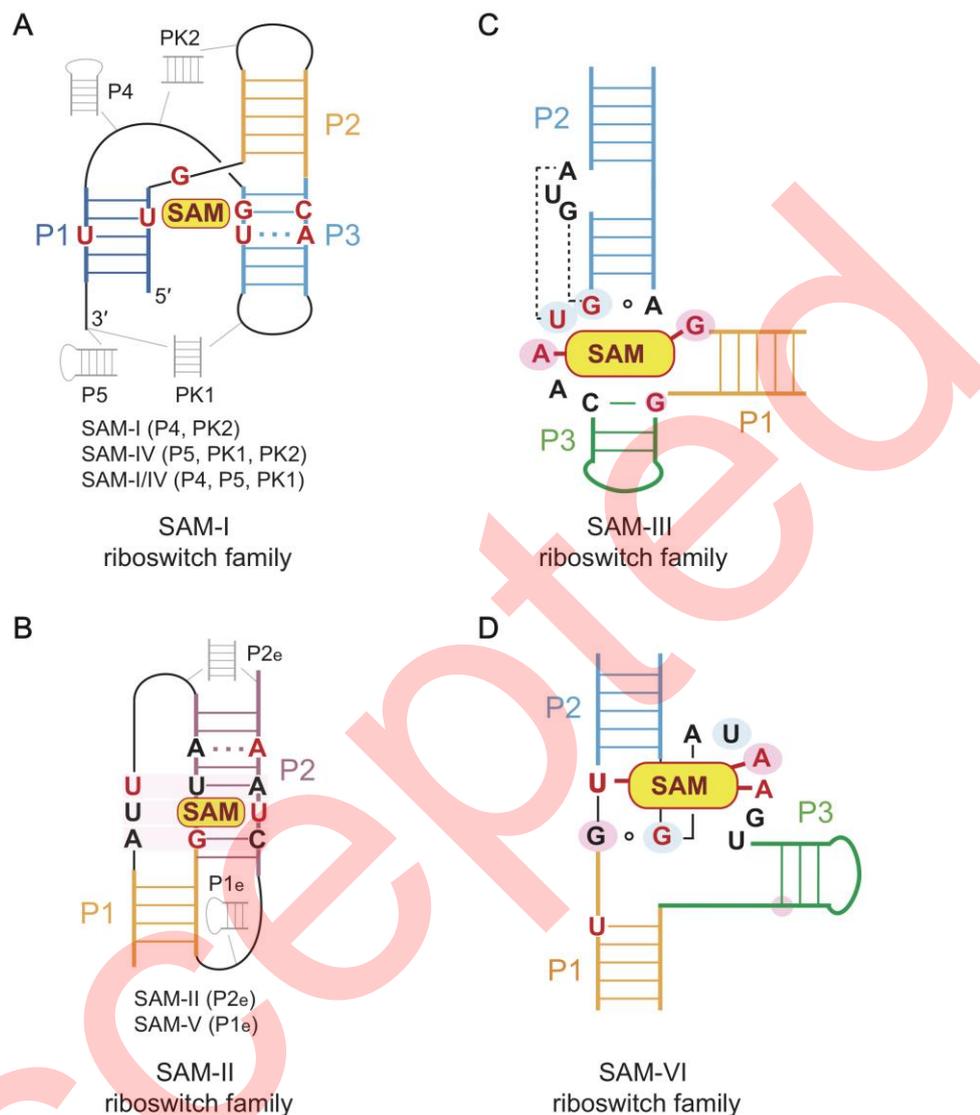


Figure 12 Common architectures and ligand recognition patterns of four SAM riboswitch families. **A**, The schematic secondary structure scheme for SAM-I riboswitch family. SAM-I, SAM-IV and SAM-I/IV share nearly identical binding pocket for SAM. All the nucleotides involved in ligand recognition are colored in red. **B**, The schematic secondary structure scheme for SAM-II riboswitch family, including SAM-II and SAM-V riboswitches. Residues involved in ligand recognition are assigned, in which nucleotides that directly interact with SAM are colored in red. **C-D**, The schematic secondary structure schemes for SAM-III and SAM-VI riboswitch families. Residues involved in ligand recognition are assigned, nucleotides that directly interact with SAM are colored in red. Here the ligand-interacted residues in SAM-III riboswitch and the equivalent residues in SAM-VI riboswitch are shown with light red shadow (hydrogen-bonding interaction) and light blue shadow (electrostatic interaction).

Table 1 Ligand Recognition Pattern of SAM-sensing Riboswit

Ligand Interaction	SAM-I Riboswitch Family		SAM-II Riboswitch Family		SAM-III Riboswitch Family	SAM-VI Riboswitch Family	SAM/SAH Riboswitch Family	SAH Riboswitch Family
	SAM-I Riboswitch	SAM-I/IV Riboswitch	SAM-II Riboswitch	SAM-V Riboswitch	SAM-III Riboswitch	SAM-VI Riboswitch	SAM/SAH Riboswitch	SAH Riboswitch
Ligand Conformation	Compact	Compact	Extend	Extend	Compact	Extend	Extend	Extend
Adenine ring	Sugar-Edge of A45; Watson-Crick Edge of U57	Sugar-Edge of A25; Watson-Crick Edge of U47	Watson-Crick Edge of U44	Watson-Crick Edge of U47	Sugar-Edge of G7; A38	Watson-Crick Edge of U8; A36 and A37	Watson-Crick Edge of U23; Ribose of U44	Sugar-Edge of G15
2'- and 3'-OH of the ribose	C47	A3, C27	G22	G21	G47	G33	G46; Ribose of G16	NO
Methionine	(G58-C44)·G11 base-triple	(G48-C24)·G8 base-triple	Watson-Crick Edge of A47	Watson-Crick Edge of A50; U10	NO	U8	NO	NO
Sulfonium moiety	Carbonyl oxygens of U7 and U88	Carbonyl oxygens of U4 and U69	Carbonyl oxygens U11 and U21	Carbonyl oxygens of U9 and U20	Carbonyl oxygen of U37 and 2'-OH of G36	Carbonyl oxygens U6 and U8	NO	NO

Accepted