LETTERS

Structure of the *S*-adenosylmethionine riboswitch regulatory mRNA element

Rebecca K. Montange[1] & Robert T. Batey[1]

Riboswitches are *cis*-acting genetic regulatory elements found in the 5′-untranslated regions of messenger RNAs that control gene expression through their ability to bind small molecule metabolites directly[1,2]. Regulation occurs through the interplay of two domains of the RNA: an aptamer domain that responds to intracellular metabolite concentrations and an expression platform that uses two mutually exclusive secondary structures to direct a decision-making process. In Gram-positive bacteria such as *Bacillus* species, riboswitches control the expression of more than 2% of all genes through their ability to respond to a diverse set of metabolites including amino acids, nucleobases and protein cofactors[1,2]. Here we report the 2.9-Å resolution crystal structure of an *S*-adenosylmethionine (SAM)-responsive riboswitch from *Thermoanaerobacter tengcongensis* complexed with *S*-adenosylmethionine, an RNA element that controls the expression of several genes involved in sulphur and methionine metabolism[3–6]. This RNA folds into a complex three-dimensional architecture that recognizes almost every functional group of the ligand through a combination of direct and indirect readout mechanisms. Ligand binding induces the formation of a series of tertiary interactions with one of the helices, serving as a communication link between the aptamer and expression platform domains.

Genetic regulation by the *S*-adenosylmethionine (SAM) riboswitch is driven through the binding of SAM to the aptamer domain, which consists of a series of stem and stem–loop structures centred on a four-way junction motif (Fig. 1a, Supplementary Fig. 1). Among the more than 150 identified SAM-I sequences[7], several nucleotides within and surrounding the junction are highly conserved phylogenetically, including a consensus kink-turn motif[8] and a genetically validated pseudoknot[9]. Stable formation of this domain through its interaction with SAM causes the downstream expression platform to create a rho-independent transcriptional terminator, turning off gene expression (Supplementary Fig. 1). In the absence of sufficient intracellular SAM, the 3′ side of the P1 helix forms an antiterminator element, allowing transcription of the mRNA.

To understand the structural basis for the specific recognition of SAM by a riboregulator and how binding is transduced into changes in gene expression, we have solved the structure of the SAM riboswitch RNA at 2.9 Å (Supplementary Table 1 and Supplementary Fig. 2). The global architecture of the SAM-I riboswitch aptamer domain is established through two sets of coaxially stacked helices. The first comprises helices P1 and P4 and the intervening J4/1 joining region (blue in Fig. 1b). J4/1 contains three conserved purine residues (A83–A85) stacked between the two helices, with A85 pairing with U64 of J3/4 and A24 of L2 to anchor helices P1/P4 against the other set of coaxial stacked helices, P2 and P3 (green in Fig. 1b). The two sets of helices pack together at a roughly 70° angle with the ligand-binding pocket situated between the minor grooves of the P1 and P3 helices. This arrangement of the P1/P4 and P2/P3 stack is architecturally similar to the P4–P6 and P3–P7 domains that form the conserved catalytic core of all group I introns[10–12] (Fig. 2). Each RNA consists of a pair of nearly coaxially stacked helices (P1/P4 and P4–P6; blue in Fig. 2) that serve to scaffold a second domain (P2/P3 and P3–P7; green in Fig. 2) that wraps around the first to complete the active site. The interactions between the two domains is stabilized by minor-groove–minor-groove packing near the point of covalent attachment by single-stranded joining regions along with a loop from the second domain forming tertiary interactions with

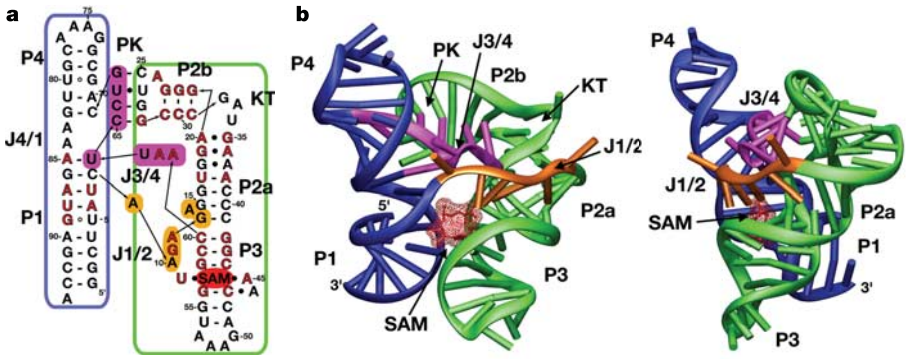

Figure 1 | Secondary and tertiary structure of the SAM-I riboswitch. **a**, Secondary structure of the SAM-I riboswitch aptamer domain reflecting the tertiary organization; nucleotides more than 95% conserved across phylogeny are highlighted in red. Solid arrows represent the direction of the RNA backbone, and elements of structure are labelled as follows: P, paired; J, joining; KT, kink-turn; PK, pseudoknot. Colours used to denote elements of structure are consistent with subsequent figures. **b**, Ribbon representation of the three-dimensional structure. The two views represent a front view (left) and a 90° clockwise rotation (right). SAM is in red with its surface represented as dots.

[1]Department of Chemistry and Biochemistry, 215 UCB, University of Colorado, Boulder, Colorado 80309, USA.

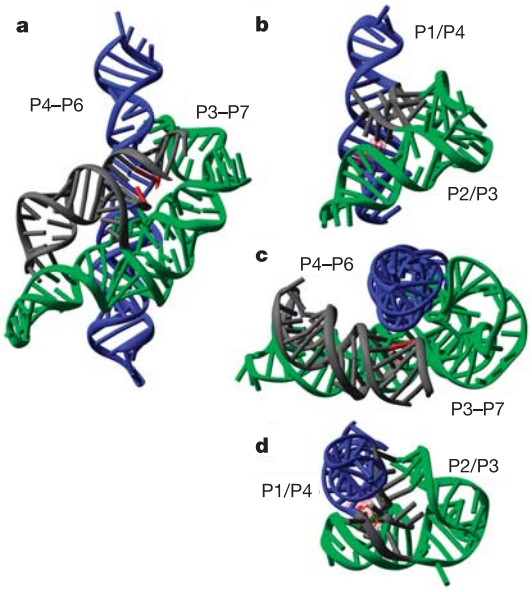

Figure 2 | Comparison of the global architecture of the *Azoarcus* group I intron and the SAM-I riboswitch. **a**, **c**, Side (**a**) and top (**c**) views of the group I intron with the two principal conserved domains P4–P6 and P3–P7 coloured in blue and green, respectively. The active site is highlighted by the two red nucleotides, which represent the site of cleavage. **b**, **d**, Side (**b**) and top (**d**) views of the SAM-I riboswitch in the same orientation as the group I intron, with the two primary domains P1/P4 and P2/P3 coloured in blue and green, respectively.

the first (L2 in the SAM riboswitch and a tetraloop in the group I intron).

A ligand-independent set of tertiary interactions between L2, J3/4 and J4/1 establish the global architecture of the RNA. Within helix P2 is a kink-turn motif that creates a roughly 100° bend in the helix; this element is identical in consensus sequence, pattern of non-canonical base pairing and three-dimensional fold to those observed in the ribosome[13], the Box C/D snoRNP[14] and the U4 snRNP[15], despite the lack of a bound protein. The kink-turn orients helix P2b back towards the P1/P4 stack, permitting a pseudoknot interaction between L2 and J3/4, as predicted from phylogenetic alignments[6] and genetic experiments[9]. The pseudoknot is tied against P1/P4 through the (A85–U64)·A24 triple (Fig. 3). J3/4 is further tied to helix P2b through two adenine-mediated triples in which the Watson–Crick face of A61 and A62 interact with the minor groove of the G22–C30 and G23–C29 pairs, respectively (Fig. 3a). This form of adenosine triple is not nearly as common as the A-minor triple motif[16] but is observed in the 16S rRNA[17]. In-line probing of several different SAM-I riboswitches revealed that key nucleotides in L2, J3/4 and J4/1 involved in the formation of this region of tertiary structure are protected from cleavage in the absence and presence of SAM[6,18] indicating that these structures are not ligand dependent. The SAM-I riboswitch, similarly to the purine riboswitch[19,20] and most probably the other riboswitches, has a pre-established architecture comprising the general coaxial stacking of helices and tertiary structural elements outside the ligand-binding pocket that serve to facilitate ligand recognition.

S-adenosylmethionine is bound within a pocket created between the P1 and P3 helices and J1/2 (Fig. 4a). The ligand adopts a compact conformation in which the methionine moiety stacks upon the adenine ring, stabilized in part by a π–cation interaction with the amino group pointing towards the adenine ring (Supplementary Fig. 3). This conformation of SAM is very different from that found in most proteins (Supplementary Fig. 4) in which SAM typically adopts an extended *trans* configuration, with the adenosyl and methionine moieties projecting away from one another[21]. Only a few methyltransferases and the repressor metJ bind SAM in a similarly compact fashion[21], although none possesses the degree of stacking between the methionine and adenine moieties observed in the RNA-bound form. Calculations of the most favoured conformations of *S*-adenosylhomocysteine (SAH) show that the RNA-bound configuration of SAM is close to the energetically most favourable form of SAH[22], indicating that the RNA probably does not distort the geometry of the ligand on binding.

The compact configuration of SAM creates two distinct faces that interact with the RNA in different fashions, one dominated by hydrogen-bonding interactions to helix P3 and the other using van der Waals surface complementarity with the minor groove of the P1 helix. The adenine ring of SAM is the central base of a triple between A45 and U57 (Fig. 4b). These two nucleotides are part of an asymmetric internal loop motif (5′-AA/U) in helix P3 (Fig. 1a) that is universally conserved among SAM-I riboswitch RNAs. The placement of the adenine ring is further stabilized by stacking with C47, which is part of a dinucleotide platform (A46·C47–G56)[23,24]. The main chain atoms of the methionine moiety are recognized through a series of hydrogen bonds with the G11·C44–G58 triple between helices P3 and J1/2 (Fig. 4c). The negatively charged carboxylate group is recognized through interactions with the Watson–Crick face of G11 (N1 and N2) in a commonly used mode of binding carboxylate and phosphate groups by RNAs[25]. Consistent with the structure is the observation that whereas SAH is tightly bound by the RNA ($K_d = 400$ nM in the *yitJ* homologue from *B. subtilis*), the analogue *S*-adenosylcysteine, containing one less side-chain methylene group, binds significantly more weakly ($K_d \approx 30$ μM)[6]. Thus, the length of the methionine side chain is sensed through an indirect readout mechanism requiring two methylene groups to place the carboxyl group within hydrogen-bonding distance of G11.

The other side of the binding pocket for SAM is created by the

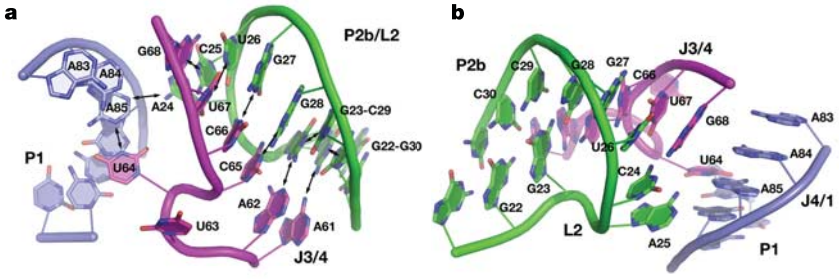

Figure 3 | Detailed view of the tertiary architecture of the pseudoknot region. **a**, Top view relative to Fig. 1c, left. The J3/4 strand ties the P1/P4 stack to the P2/P3 stack by means of base interactions. These comprise A-minor triples mediated by A61 and A62 on the 5′ end of J3/4 with helix P2b, a base triple between U64, A85 and A24, and a pseudoknot between C25–G28 of L2 and C65–G68 of J3/4. **b**, View of this region from the back relative to Fig. 1c, left, emphasizing the pseudoknot interaction and its orientation relative to the P1/P4 stack.

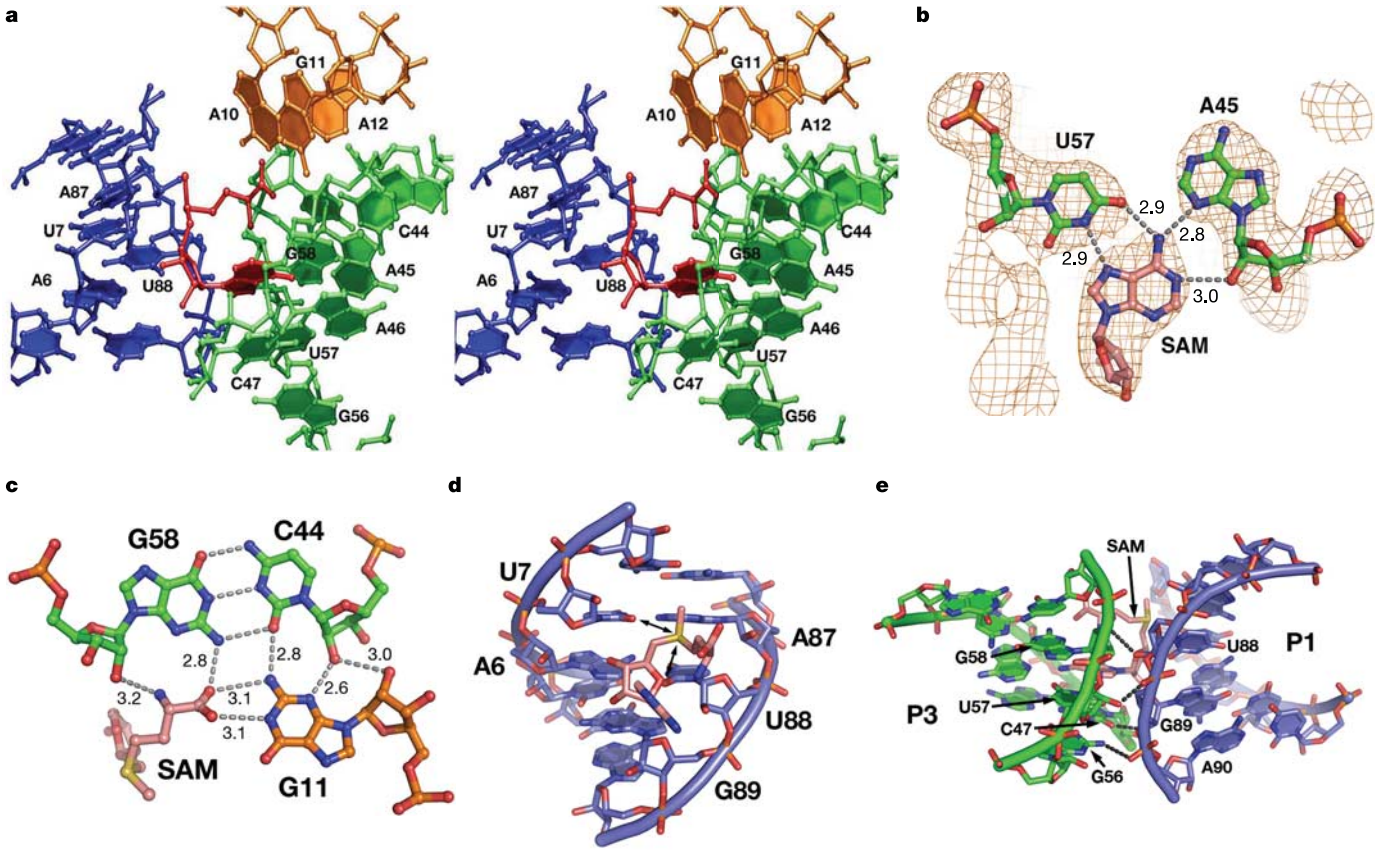

Figure 4 | Detailed view of SAM recognition by the riboswitch. **a**, Stereo view of the SAM-binding pocket comprising the P1 and P3 helices and J1/2. **b**, Hydrogen-bonding interactions between the adenine base of SAM, A45 and U57. The experimental electron density map is shown as an orange cage, contoured at 1.25σ. **c**, Interactions between the main chain atoms of methionine and the G58–C44 pair in P3 and G11 of J1/2. Distances in **b** and **c** are given in ångströms. **d**, SAM binding to the P1 helix; selectivity for SAM is probably mediated through electrostatic interactions with the carbonyl O2 of U7 and U88 (double-headed arrows). **e**, Ligand-induced interactions between the P3 helix and the 3′ side of the P1 helix (black dashed lines).

minor groove of the P1 helix, adjacent to the universally conserved A6–U88 and U7–A87 base pairs. The ribose sugar of SAM is situated in such a way that it forms a direct hydrogen bond only between its 2′-hydroxyl and O4 of C47 in helix P3. Interactions between SAM and helix P1 seem to be restricted to van der Waals packing with U7 and U88, which is consistent with binding studies with SAM analogues containing deletions of the 2′- and 3′-hydroxyl groups[18]. The sulphur atom is situated about 4 Å from the O2 carbonyl oxygens of U7 and U88 (Fig. 4d). This positioning serves as the basis for a 100-fold preference for SAM over SAH[6,18]; the positively charged sulphur is positioned to make favourable electrostatic interactions with the carbonyls of the minor groove of helix P1. This RNA has evolved to not use the anionic phosphate backbone for the recognition of either of the positively charged functional groups in SAM. Biochemical observations indicate that the identity of the charged moiety at this position is not important, but the presence of a formal positive charge or high partial positive charge is required[18]. The methyl group points towards a solvent cavity within the interior of the RNA and is not directly recognized, as previously predicted[6]. This structure is therefore consistent with all of the observed genetic, biochemical and phylogenetic data available for this class of riboswitches.

The mechanism by which the SAM-I riboswitch transduces a binding event into changes in gene expression is revealed by this structure. The interaction between SAM and the mRNA induces a series of tertiary interactions involving hydrogen-bonding interactions between backbone ribose–phosphate groups of U88–A90 in helix P1 and C47, G56–G58 of helix P3 and van der Waals interactions between the ribose sugars of C48 and A90 (Fig. 4e). These contacts are clearly ligand induced, as revealed by in-line probing, which shows strong protection of these nucleotides only on association with SAM[6,18]. In addition, the ligand is about 74% inaccessible to solvent (Supplementary Fig. 5); this degree of burial requires that helices P1 and P3 close around the ligand during the binding event, forming the new tertiary interactions.

Similarly to the purine riboswitch[19,20], the 3′ side of the P1 helix is buttressed through an extensive network of hydrogen-bonding, electrostatic and van der Waals interactions with both the ligand and other regions of the RNA, establishing the basis for gene regulation. In both cases, this side of the P1 helix is an integral part of a structural switch involving two mutually exclusive secondary structures in the expression platform (Supplementary Fig. 1). Stabilization of the P1 helix communicates the liganded state of the aptamer domain to the expression platform by preventing formation of the antiterminator element, causing the expression platform to form a terminator element. Because all of the riboswitches (with the exception of the *glmS* riboswitch-ribozyme[26]) have a similar secondary structural organization in which a P1-like stem directly links the aptamer domain with the expression platform, this mechanism is probably common to all. Thus, ligand-induced allosteric changes drive effective gene regulation by both RNA-based riboswitches and protein-based transcriptional repressors.

METHODS

Transcription vectors for the synthesis of RNA were constructed and transcribed with the use of techniques described previously[27], and purified with 12% denaturing polyacrylamide gels. After purification, the RNA was concentrated, refolded by heating and cooling, and SAM was added to a final concentration of

5 mM. Crystals were grown by mixing this solution in a 1:1 ratio with mother liquor (100 mM KCl, 5 mM $MgCl_2$, 10% 2-methyl-2,4-pentanediol 10 mM sodium cacodylate pH 7.0, 8 mM iridium hexa-ammine, 6 mM spermine-HCl) and incubated for 48 h at 30 °C. A two-wavelength anomalous diffraction experiment was performed on crystals cryoprotected in mother liquor plus 15% ethylene glycol, and data were collected to 2.8 Å resolution. The data were indexed, integrated and scaled with D*TREK[28], and all subsequent calculations and refinement were performed with CNS[29] to obtain a model containing all RNA atoms as well as the SAM ligand with final $R_{crystal}$ and R_{free} values of 26.7% and 28.8%, respectively. Additional experimental details and references are given in Supplementary Information.

Received 8 March; accepted 20 April 2006.

1. Nudler, E. & Mironov, A. S. The riboswitch control of bacterial metabolism. *Trends Biochem. Sci.* **29,** 11–17 (2004).
2. Winkler, W. C. & Breaker, R. R. Genetic control by metabolite-binding riboswitches. *ChemBioChem* **4,** 1024–1032 (2003).
3. Epshtein, V., Mironov, A. S. & Nudler, E. The riboswitch-mediated control of sulfur metabolism in bacteria. *Proc. Natl Acad. Sci. USA* **100,** 5052–5056 (2003).
4. Grundy, F. J. & Henkin, T. M. The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in Gram-positive bacteria. *Mol. Microbiol.* **30,** 737–749 (1998).
5. McDaniel, B. A., Grundy, F. J., Artsimovitch, I. & Henkin, T. M. Transcription termination control of the S box system: direct measurement of S-adenosylmethionine by the leader RNA. *Proc. Natl Acad. Sci. USA* **100,** 3083–3088 (2003).
6. Winkler, W. C., Nahvi, A., Sudarsan, N., Barrick, J. E. & Breaker, R. R. An mRNA structure that controls gene expression by binding *S*-adenosylmethionine. *Nature Struct. Biol.* **10,** 701–707 (2003).
7. Griffiths-Jones, S. *et al.* Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* **33,** D121–D124 (2005).
8. Winkler, W. C., Grundy, F. J., Murphy, B. A. & Henkin, T. M. The GA motif: an RNA element common to bacterial antitermination systems, rRNA, and eukaryotic RNAs. *RNA* **7,** 1165–1172 (2001).
9. McDaniel, B. A., Grundy, F. J. & Henkin, T. M. A tertiary structural element in S box leader RNAs is required for S-adenosylmethionine-directed transcription termination. *Mol. Microbiol.* **57,** 1008–1021 (2005).
10. Adams, P. L., Stahley, M. R., Kosek, A. B., Wang, J. & Strobel, S. A. Crystal structure of a self-splicing group I intron with both exons. *Nature* **430,** 45–50 (2004).
11. Golden, B. L., Kim, H. & Chase, E. Crystal structure of a phage Twort group I ribozyme–product complex. *Nature Struct. Mol. Biol.* **12,** 82–89 (2005).
12. Guo, F., Gooding, A. R. & Cech, T. R. Structure of the *Tetrahymena* ribozyme: base triple sandwich and metal ion at the active site. *Mol. Cell* **16,** 351–362 (2004).
13. Klein, D. J., Schmeing, T. M., Moore, P. B. & Steitz, T. A. The kink-turn: a new RNA secondary structure motif. *EMBO J.* **20,** 4214–4221 (2001).
14. Moore, T., Zhang, Y., Fenley, M. O. & Li, H. Molecular basis of box C/D RNA–protein interactions; cocrystal structure of archaeal L7Ae and a box C/D RNA. *Structure* **12,** 807–818 (2004).
15. Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R. & Ficner, R. Crystal structure of the spliceosomal 15.5 kD protein bound to a U4 snRNA fragment. *Mol. Cell* **6,** 1331–1342 (2000).
16. Nissen, P., Ippolito, J. A., Ban, N., Moore, P. B. & Steitz, T. A. RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc. Natl Acad. Sci. USA* **98,** 4899–4903 (2001).
17. Wimberly, B. T. *et al.* Structure of the 30S ribosomal subunit. *Nature* **407,** 327–339 (2000).
18. Lim, J., Winkler, W. C., Nakamura, S., Scott, V. & Breaker, R. R. Molecular-recognition characteristics of SAM-binding riboswitches. *Angew. Chem. Int. Edn Engl.* **45,** 964–968 (2006).
19. Batey, R. T., Gilbert, S. D. & Montange, R. K. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature* **432,** 411–415 (2004).
20. Serganov, A. *et al.* Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem. Biol.* **11,** 1729–1741 (2004).
21. Schubert, H. L., Blumenthal, R. M. & Cheng, X. Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem. Sci.* **28,** 329–335 (2003).
22. Ishida, T., Tanaka, A., Inoue, M., Fujiwara, T. & Tomita, K. Conformational studies of *S*-adenosyl-L-homocysteine, a potential inhibitor of *S*-adenosyl-L-methionine-dependent methyltransferases. *J. Am. Chem. Soc.* **104,** 7239–7248 (1982).
23. Cate, J. H. *et al.* RNA tertiary structure mediation by adenosine platforms. *Science* **273,** 1696–1699 (1996).
24. Leontis, N. B., Stombaugh, J. & Westhof, E. The non-Watson–Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res.* **30,** 3497–3531 (2002).
25. Auffinger, P., Bielecki, L. & Westhof, E. Anion binding to nucleic acids. *Structure* **12,** 379–388 (2004).
26. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. & Breaker, R. R. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428,** 281–286 (2004).
27. Kieft, J. S. & Batey, R. T. A general method for rapid and nondenaturing purification of RNAs. *RNA* **10,** 988–995 (2004).
28. Pflugrath, J. W. The finer things in X-ray diffraction data collection. *Acta Crystallogr. D* **55,** 1718–1725 (1999).
29. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Edwards for managing the CU Boulder X-ray crystallography facility, the staff at ALS beamline 8.2.1 for assistance, and A. Pardi, T. Cech, R. Breaker, W. Winkler, B. Golden, J. Kieft and A. Ke for discussions. This work was supported by grants from the National Institutes of Health and the W. M. Keck Foundation Initiative in RNA Science at the University of Colorado (R.T.B.) and a National Institutes of Health Predoctoral Training Grant (R.K.M.).

Author Information The atomic coordinates and structure factors have been deposited in the Protein Data Bank with the accession number 2GIS. Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.T.B. (robert.batey@colorado.edu).