

Structural basis of 7SK RNA 5'- γ -phosphate methylation and retention by MePCE

Yuan Yang^{1,3}, Catherine D. Eichhorn^{1,3}, Yaqiang Wang¹, Duilio Cascio^{1,2} and Juli Feigon^{1*}

Among RNA 5'-cap structures, γ -phosphate monomethylation is unique to a small subset of noncoding RNAs, 7SK and U6 in humans. 7SK is capped by methylphosphate capping enzyme (MePCE), which has a second nonenzymatic role as a core component of the 7SK ribonuclear protein (RNP), an essential regulator of RNA transcription. We report 2.0- and 2.1-Å X-ray crystal structures of the human MePCE methyltransferase domain bound to S-adenosylhomocysteine (SAH) and uncapped or capped 7SK substrates, respectively. 7SK recognition is achieved by protein contacts to a 5'-hairpin-single-stranded RNA region, thus explaining MePCE's specificity for 7SK and U6. The structures reveal SAH and product RNA in a near-transition-state geometry. Unexpectedly, binding experiments showed that MePCE has higher affinity for capped versus uncapped 7SK, and kinetic data support a model of slow product release. This work reveals the molecular mechanism of methyl transfer and 7SK retention by MePCE for subsequent assembly of 7SK RNP.

5' capping is a tightly regulated step in RNA processing that is required for RNA stability, localization, and function^{1,2}. All caps, which vary in an RNA- and organism-dependent manner, feature one or more methyl groups added by a SAM-dependent methyltransferase (MTase; Fig. 1a). A unique monomethyl- γ -phosphate 5' cap (mpppG)³ (Fig. 1b) has been identified in a small subset of RNA polymerase III transcripts, i.e., human 7SK long noncoding RNA and U6 spliceosomal RNA; rodent B2 short interspersed element RNA; and plant U3 small nucleolar RNA³⁻⁵. 7SK and U6 associate with the methylphosphate capping enzyme MePCE (also called BCDIN3 in mammals and Bin3 in *Drosophila*)^{6,7}, but only 7SK has been conclusively shown to be capped by MePCE^{6,8}.

MePCE also has a nonenzymatic function as a component of the core 7SK RNP, and knockdown significantly decreases 7SK levels in vivo^{6,9}. 7SK regulates RNA polymerase II transcription elongation in higher eukaryotes by sequestering and inactivating positive transcription elongation factor b (P-TEFb)^{10,11}. P-TEFb phosphorylates the RNA polymerase II C-terminal domain and negative transcription elongation factors, thus causing RNA polymerase II to transition from a promoter-proximal paused state to productive transcription elongation¹². 7SK constitutively assembles with MePCE at the 5' end and with La-related protein group 7 (Larp7) at the 3' end¹³⁻¹⁵. After binding 7SK, Larp7 and MePCE interact, forming a core RNP in which MePCE is inactivated⁷⁻⁹. Hexim1/2 and P-TEFb subsequently assemble onto 7SK RNP and P-TEFb is sequestered in an inactive form^{10,11,16,17}.

Little is known about 7SK biogenesis, and two alternative secondary structures for 7SK have been proposed^{18,19}, a 'linear 7SK model' with four stem-loops^{8,18} (Fig. 1c), which is supported by chemical mapping and enzymatic footprinting experiments, and a 'circular 7SK model' based on sequence conservation, with up to eight stem-loops, in which the first 10 nt at the 5' end base-pair with nucleotides near the 3' end^{19,20} (Supplementary Fig. 1). Both RNA secondary-structure models have an identical terminal stem-loop 4 (SL4), where Larp7 binds^{7,21-23}. However, the secondary structures differ at the 5' end, where MePCE binds and caps (Supplementary Fig. 1). For U6, the consensus motif for the determinants of

capping comprises a 5'-triphosphate, a hairpin beginning at the 5' end, and adjacent 3' single-stranded RNA (ssRNA) with the sequence AUAUAC²⁴. In 7SK, the sequence and secondary structure at the 5' end corresponding to the linear 7SK model are important for MePCE assembly in vivo⁷. Whether one or both of the proposed secondary structures participate in MePCE capping and/or core 7SK RNP assembly remains to be determined.

MePCE contains a C-terminal MTase domain that is highly conserved among eukaryotes and a highly variable N-terminal region of low sequence complexity²⁵. The structure of the MePCE MTase domain bound to S-adenosylhomocysteine (SAH) (MePCE-SAH) was determined by the Structural Genomics Consortium (PDB 5UNA). Here, we report structures of the MePCE MTase domain crystallized in the presence of 7SK substrate RNA and either SAH or SAM. Unexpectedly, despite transfer of the methyl group from SAM to the 5'- γ -phosphate of the 7SK substrate during crystallization, the RNA remains bound to MePCE, thereby providing a unique example of a product-bound RNA MTase structure. Together with analysis of enzyme kinetics and reactant and product binding affinities, the X-ray crystal structures of MePCE-SAH in complex with uncapped (MePCE-SAH-7SK) and capped (MePCE-SAH-me7SK) RNAs reported here reveal the mechanism of methyl transfer to a phosphate group and the determinants of 7SK recognition and retention, and define the first steps of assembly of 7SK core RNP.

Results

Crystal structures of the MePCE MTase domain bound to 7SK.

On the basis of studies indicating that MePCE binds the proximal end of the 7SK SL1 and adjacent ssRNA in vivo⁷ (Fig. 1c), we designed and synthesized an RNA hairpin-ssRNA construct comprising the 5' end of 7SK SL1 with a UUCG tetraloop and an 8-nt ssRNA 3' overhang (SL1p) (Fig. 1c). SAM or SAH did not copurify with the MePCE MTase domain (residues 400-689, MePCE_{MT}), as determined by LC-MS (Supplementary Fig. 2). To verify that SL1p is capped by MePCE_{MT}, we incubated [*methyl*-¹³C]SAM, SL1p, and MePCE_{MT} under single-turnover conditions to completion and

¹Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA, USA. ²UCLA-U.S. Department of Energy (DOE) Institute of Genomics and Proteomics, University of California, Los Angeles, Los Angeles, CA, USA. ³These authors contributed equally: Yuan Yang, Catherine D. Eichhorn. *e-mail: feigon@mbi.ucla.edu

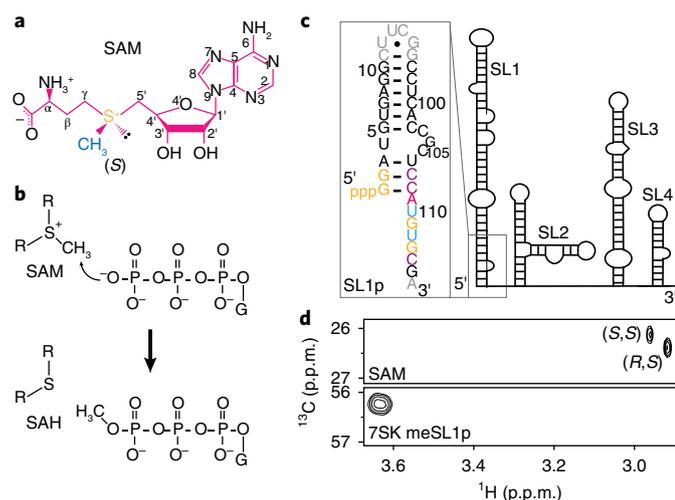


Fig. 1 | MePCE binds and caps 7SK. **a**, SAM cofactor, shown with a methyl sulfonium group as an (S) stereoisomer. Pink, chemical bonds; yellow, sulfur; blue, methyl group. **b**, Proposed transfer reaction mechanism of the SAM methyl to RNA 5'- γ -phosphate oxygen by MePCE. **c**, Cartoon schematic of the proposed secondary structure of the 7SK linear model. Inset, sequence and secondary structure of SL1p used in this study. Residues involved in MePCE interactions are colored by nucleotide (pink, A; gold, G; cyan, U; purple, C). Native residues not involved in protein binding are black, and non-native residues are gray. **d**, ^1H - ^{13}C HMQC spectra of free [methyl- ^{13}C]SAM (top) and the quantitatively capped 7SK [methyl- ^{13}C]SL1p. The two peaks for free [methyl- ^{13}C]SAM are from the (S,S) and (R,S) diastereomers. The unique resonance of meSL1p methyl has chemical-shift values of 3.635 p.p.m. (^1H) and 56.23 p.p.m. (^{13}C).

analyzed the subsequently purified SL1p by using NMR spectroscopy (Fig. 1d). A unique resonance attributed to a methyl group on SL1p ([methyl- ^{13}C]SL1p) was observed, with a substantial chemical-shift difference relative to [methyl- ^{13}C]SAM (Fig. 1d). High-resolution electrospray ionization mass spectrometry confirmed that SL1p was fully monomethylated (meSL1p) (Supplementary Fig. 3). These assays were performed in the absence of divalent cations (Fig. 1d and Supplementary Fig. 3), thus indicating that MePCE does not require divalent cations for catalysis.

MePCE_{MT} was crystallized in complex with SL1p in the presence of either SAH or SAM, and structures were solved from crystals that diffracted to 2.0-Å and 2.1-Å resolution, respectively (Supplementary Fig. 4 and Supplementary Table 1). The crystal structure of the sample with added SAM lacked the electron density expected for the SAM methyl moiety, whereas clear additional density was observed at an RNA 5'- γ -phosphate oxygen, a result indicating that methyl transfer occurred during crystallization (Supplementary Fig. 4). Thus, crystal structures of MePCE_{MT} in complex with SAH and SL1p in uncapped (MePCE-SAH-7SK) and capped (MePCE-SAH-me7SK) forms were determined (Supplementary Fig. 4). The structures are nearly identical (all-atom r.m.s. deviation 0.21 Å) and show substantial differences only at the SAH and 5'-triphosphate (Supplementary Figs. 4 and 5). In both structures, MePCE_{MT} has the expected Rossmann-like fold common to RNA methyltransferases^{1,2}, with an $\alpha\beta\alpha\beta\alpha\beta\alpha\beta\beta$ topology (Fig. 2a,b) also seen in RNA-free MePCE-SAH (Fig. 2c). Helices $\alpha 1$, $\alpha 2$, and $\alpha 3$ lie underneath the central β -sheet ($\uparrow 3\uparrow 2\uparrow 1\uparrow 4\uparrow 5\downarrow 7\uparrow 6$), and $\alpha 4$, $\alpha 5$, and $\alpha 6$ lie above the β -sheet. Noncanonical helices $\alpha 5'$ and $\alpha 6'$ are located adjacent to $\alpha 5$ and $\alpha 6$, respectively, at the β -sheet edge (Fig. 2b,c). In MePCE-SAH-me7SK (Fig. 2b) and MePCE-SAH-7SK (Supplementary Fig. 4a), MePCE_{MT} has two additional helices, $\alpha 0$ and $\alpha 7$, located

in the structure between $\alpha 3$ and $\alpha 5'$. SAH and the RNA 5'-triphosphate are buried in the active site (Fig. 2b,d). Notably, no divalent or trivalent cations are observed in the active site, even for crystals soaked with Mg^{2+} or Sm^{3+} (Methods). SL1p forms a 10-bp hairpin with an asymmetric (1–3) internal loop separating the bottom 3 bp and the top 7 bp, and a UUCG tetraloop; eight single-stranded nucleotides stack on one another below the terminal base pair (Figs. 1c and 2b and Supplementary Fig. 4). The internal loop, which is outside of the MePCE_{MT}-binding site, has crystal contacts with nucleotides G113–A116 from another molecule in the asymmetric unit, which form non-native interactions (Supplementary Fig. 6).

A comparison of MePCE-SAH-me7SK (Fig. 2b) and MePCE-SAH (PDB 5UNA; Fig. 2c) highlights the conformational changes in MePCE_{MT} that occur after RNA binding. The structures of MePCE-SAH in the absence and presence of SL1p are globally similar (all-atom r.m.s. deviation of 0.29 Å), but there are major differences around the active site and RNA-protein interface (Fig. 2e,f). In the absence of RNA, the N terminus (residues 418–431) and C-terminal $\beta 6$ – $\beta 7$ loop (residues 665–676) are partially disordered (Fig. 2c,f), whereas in the presence of meSL1p, a helix $\alpha 0$ forms near the N terminus, loop residues that connect to $\alpha 1$ become ordered, and the $\beta 6$ – $\beta 7$ loop adopts an ordered structure with a short helix $\alpha 7$ in the middle (Fig. 2b,e). Together with helices $\alpha 5'$ and $\alpha 6'$ —whose conformation does not change significantly when RNA binds, except for the Lys625 and Arg626 side chains—these elements form a tunnel that encloses the RNA triphosphate (hereafter called the triphosphate-binding tunnel) and consequently position it near the cofactor in the active site (Fig. 2d,g,h). These new structured elements are stabilized by hydrogen bonds and stacking interactions to both RNA substrate and one another, and sequester SAH in the cofactor-binding pocket with only its sulfur group exposed to the γ -phosphate in the tunnel (Fig. 2d,g,h). In contrast, in the absence of RNA, the SAH homocysteine is solvent accessible (Fig. 2i). In summary, binding to SL1p induces conformational changes in MePCE_{MT} that enclose SAH in the active site.

Extensive interaction network between MePCE_{MT} and 7SK. In MePCE-SAH-me7SK, there is an extensive interaction network involving a highly basic surface on MePCE and the SL1p 5'-triphosphate, two terminal base pairs, and the following six ssRNA nucleotides (Figs. 2b and 3 and Supplementary Fig. 4). G1 and G2 in the first 2 bp are recognized by MePCE_{MT} helices $\alpha 5'$, $\alpha 6'$, and $\alpha 7$; G1–C108 and the first two ssRNA nucleotides straddle helix $\alpha 0$; and the following four ssRNA nucleotides interact primarily with helix $\alpha 3$ (Figs. 2b and 3a). The interior of the triphosphate-binding tunnel (Fig. 2d and Supplementary Fig. 5) is accessible to solvent at the guanosine and α -phosphate and narrows at the β - and γ -phosphates, thereby excluding solvent. The oxygens on the 5'-triphosphate form multiple direct hydrogen bonds (to all three phosphates) and indirect hydrogen bonds (to α -phosphate) from residues on helices $\alpha 7$, $\alpha 6'$, $\alpha 5'$, and $\alpha 0$, and adjoining loops (Fig. 3a–e and Supplementary Table 2) that form the tunnel. Additional hydrogen bonds from $\alpha 6'$ and $\alpha 7$ recognize the attached G1 base and ribose, respectively (Fig. 3e and Supplementary Table 2). These hydrogen bonds fix the otherwise highly flexible 5'-triphosphate, thereby positioning the γ -phosphate oxygen for methyl transfer.

The terminal 2 bp and the following four ssRNA nucleotides (A109–U112) of SL1p are sequence-specifically recognized by MePCE_{MT}, and G113 and C114 form backbone and ribose interactions to MePCE_{MT} (Fig. 3a). MePCE_{MT} interacts in the major groove at the first 2 bp, forming hydrogen bonds to G1 and G2, as well as the phosphates (Fig. 3e–h). Phe674 ($\alpha 7$) stacks between the G1 ribose and the Arg425 ($\alpha 0$) guanidinium plane, which then stacks on Arg433 ($\alpha 1$), thereby encompassing the 'right wall' of the triphosphate-binding tunnel (Figs. 2e and 3e,g). The terminal G1–C108 base pair and ssRNA nucleotides A109 and U110 straddle

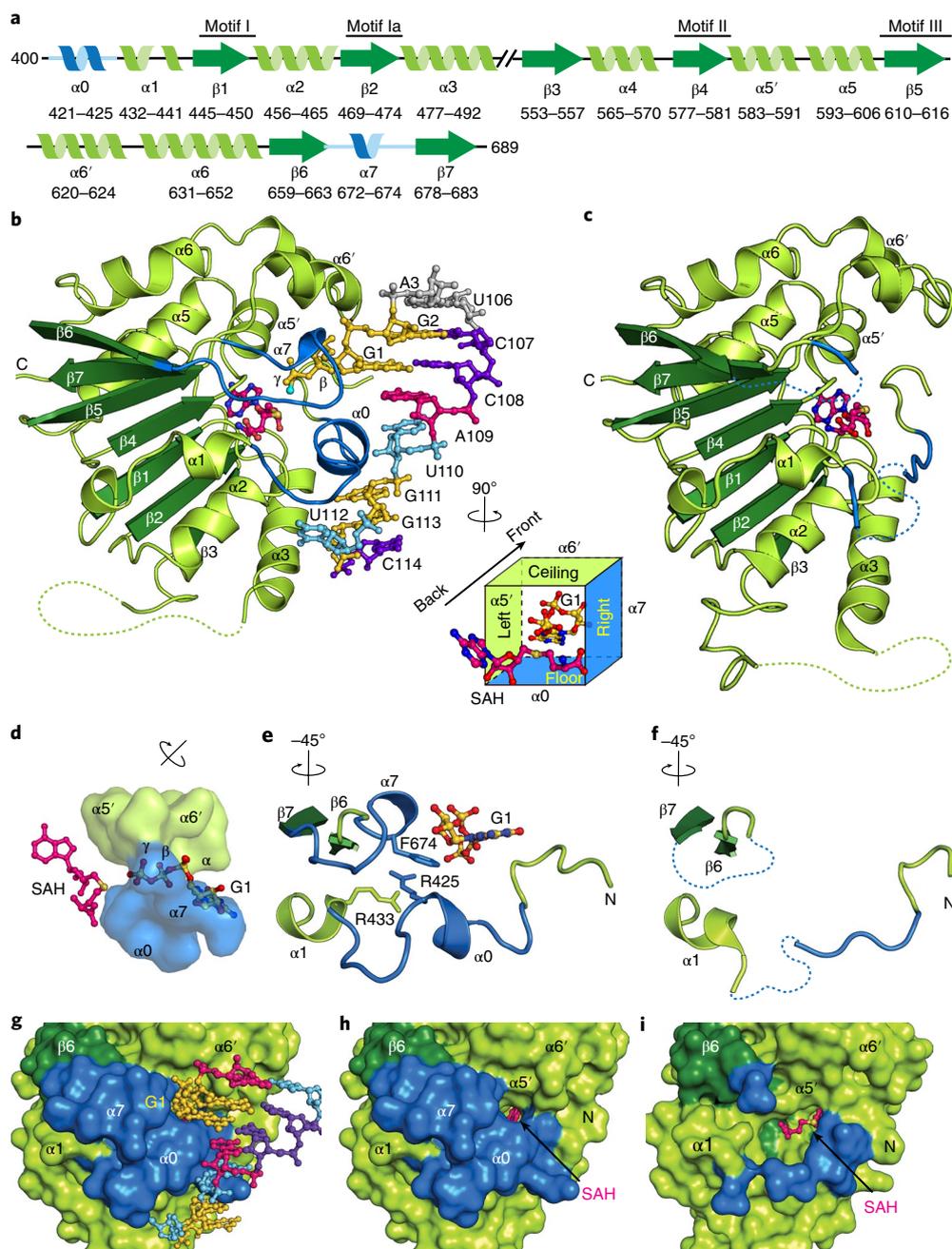


Fig. 2 | Structural changes in MePCE_{MT} upon binding RNA. **a**, Secondary-structure boundaries of MePCE_{MT}. Dark green, β -strands; lime, α -helices and loops; blue, loops that become ordered upon binding RNA. **b**, Crystal structure of MePCE-SAH-me7SK, with SAH in pink and meSL1p colored as in Fig. 1a. The methyl carbon is cyan. meSL1p and SAH are shown in ball-and-stick representation, and MePCE_{MT} is shown as a cartoon. Inset, cube representation of the active site topology, defined by helix $\alpha 0$ at the floor, helix $\alpha 5'$ to the left, helix $\alpha 6'$ at the ceiling, the β -sheet edge at the back, helix $\alpha 7$ to the right, and SL1p at the front, as seen from the back wall. **c**, Crystal structure of MePCE-SAH (PDB 5UNA). **d**, The triphosphate-binding tunnel formed by helix $\alpha 0$ and $\alpha 7$ (blue) and helix $\alpha 5'$ and $\alpha 6'$ (green) in the MePCE-SAH-7SK structure. Protein residues are shown in transparent surface representation. RNA residue G1 and SAH are shown in ball-and-stick representation. **e, f**, MePCE_{MT} N- and C-terminal loops in the presence (**e**) of SL1p are structured and form $\alpha 0$ and $\alpha 7$ helices (blue), but in the absence (**f**) of SL1p are disordered (represented as dashed blue lines). **g, h, i**, Surface representations of MePCE-SAH-7SK. In **g**, SAH is buried by MePCE_{MT} N- and C-terminal regions and SL1p. In **h**, removal of SL1p RNA visually reveals the sulfur atom of SAH at the base of the triphosphate-binding tunnel. **i**, SAH is solvent accessible in the absence of bound RNA for MePCE-SAH (PDB 5UNA).

helix $\alpha 0$, which is enriched in tyrosines whose aromatic rings splay out around $\alpha 0$ (Fig. 3f). This 'tyrosine wheel' forms intricate interactions with ssRNA nucleotides outside the active site, the 5'-triphosphate within the active site, and helix $\alpha 7$ residues (Fig. 3a–c, f, g, i–k). Specifically, Tyr424 has π - π stacking and hydrogen-bonding interactions with the A109 base and ribose (Fig. 3f, i).

C108, A109, and U110 have continuous base stacking (Fig. 3f, j). The Lys420 side chain on the tyrosine wheel stacks between U110 and G111 and forms hydrogen bonds to U110 (Fig. 3f, j, k). G111 forms a hydrogen bond to Ser482 ($\alpha 3$) (Fig. 3k) and stacks on G113 (Fig. 3f), with the U112 base flipped 120° out about the G111 ribose. The U112 base forms stacking and hydrogen-bonding interactions with

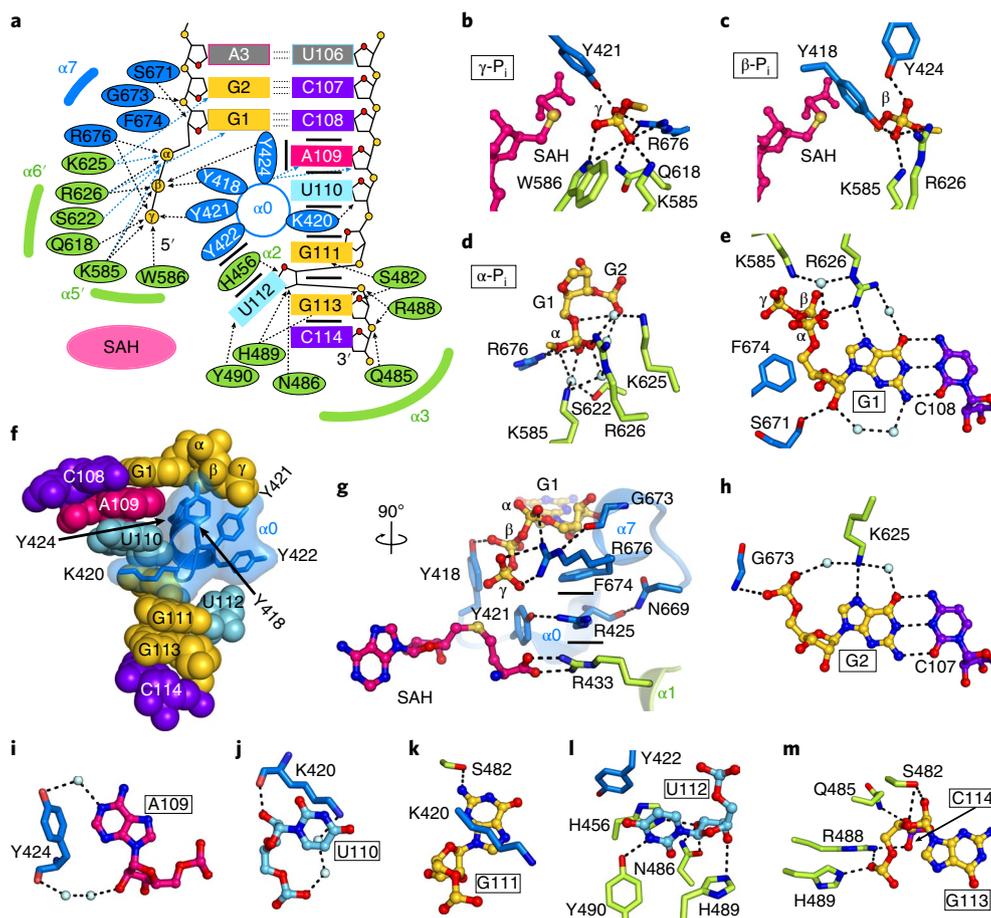


Fig. 3 | Sequence- and structure-specific recognition of 7SK by MePCE_{MT}. **a**, Schematic showing the specific interactions between MePCE_{MT} and 7SK SL1p. $\alpha 0$ is indicated as a blue circle with RNA-interacting residues as spokes around the helix (tyrosine wheel), the locations of other helices are shown with curved lines, and SAH is highlighted with a magenta oval. Dashed lines represent hydrogen bonds, and blue dashed lines represent water-mediated hydrogen bonds. Bold lines represent stacking interactions. **b–e**, Interactions between MePCE_{MT} and pppG1 for γ -phosphate (**b**), β -phosphate (**c**), α -phosphate (**d**), and G1 nucleotide (**e**). P_i denotes phosphate. **f**, The tyrosine wheel in helix $\alpha 0$ and the adjacent loop recognizes the helix–ssRNA junction of SL1p. Surface, stick, and cartoon representation are shown for the tyrosine residues and Lys420, and RNA residues are shown in ball-and-stick representation. **g**, Active site residues that participate in specific stacking and hydrogen-bonding interactions that stabilize helices $\alpha 0$ and $\alpha 7$. **h**, Protein residues that interact with the G2 nucleotide and backbone. **i–m**, Interactions between MePCE_{MT} and the ssRNA nucleotides A109 (**i**); U110 (**j**); G111 (**k**); U112 (**l**); and G113 and C114 (**m**).

four residues from $\alpha 2$ and $\alpha 3$ (Fig. 3l). U112 is the last RNA residue that is base-specifically recognized by MePCE_{MT}; G113 and C114 form only backbone interactions with helix $\alpha 3$ residues (Fig. 3m). The residues that participate in triphosphate and base-pair recognition are highly conserved in eukaryotes, whereas helix $\alpha 3$ is highly conserved among vertebrates but poorly conserved in fission yeast and plants (Supplementary Fig. 7). The large MePCE–RNA interaction surface explains the *in vivo* mutagenesis data indicating that a 5'-hairpin structure followed by ssRNA nucleotides is required for the mpppG cap^{7,24}.

The cofactor and triphosphate are in a near-transition state. The positions of SAH in MePCE–SAH–7SK and MePCE–SAH–me7SK are nearly identical, and there are only minor rotations of the γ - and β -phosphates and shifting of the SAH sulfur (Fig. 4a,b and Supplementary Fig. 5). The SAH adenosine and homocysteine moieties are anchored in a bidentate binding cleft at the back wall of the active site, and the sulfur atom points away from the β -sheet toward the G1 γ -phosphate oxygen atom, which peeks out of the triphosphate-binding tunnel (Figs. 2b and 4a–c). The SAH adenosine is recognized by multiple hydrogen bonds and stacking interactions

primarily from motif I and Ia residues (Fig. 4d, Supplementary Fig. 7 and Supplementary Table 2). The SAH homocysteine amide forms a hydrogen bond to the Gly451 ($\beta 1$ – $\alpha 2$ loop, motif I) and Leu581 ($\beta 4$, motif II) backbone carbonyl groups, and the SAH carboxylate group forms three hydrogen bonds with the Arg433 guanidinium ($\alpha 1$) and Tyr422 hydroxyl ($\alpha 0$) (Fig. 4e and Supplementary Table 2). Notably, in RNA-free MePCE–SAH, Tyr422 and Arg433 are partially disordered (Supplementary Fig. 8). In contrast, they are highly ordered in the RNA-bound structures (Supplementary Fig. 5), thus indicating that RNA binding to MePCE_{MT} further stabilizes SAH in its binding cleft.

In MePCE–SAH–me7SK, the methyl carbon has been transferred from SAM to the G1 γ -phosphate oxygen (Figs. 1b and 4g and Supplementary Fig. 4) with an O–C bond distance of 1.40 Å (Supplementary Fig. 4). Superposition of the MePCE–SAH–7SK and MePCE–SAH–me7SK reveals a local shift in the SAH sulfur by 0.89 Å, which would allow room for the methyl group (Fig. 4b). For MePCE–SAH–7SK, the distance between the SAH sulfur and G1 γ -phosphate oxygen (3.37 Å) equals the sum of their van der Waals (vdW) radii (3.32 Å) (Fig. 4f) (<http://www.rsc.org/periodic-table/>), whereas for MePCE–SAH–me7SK, the distance between the

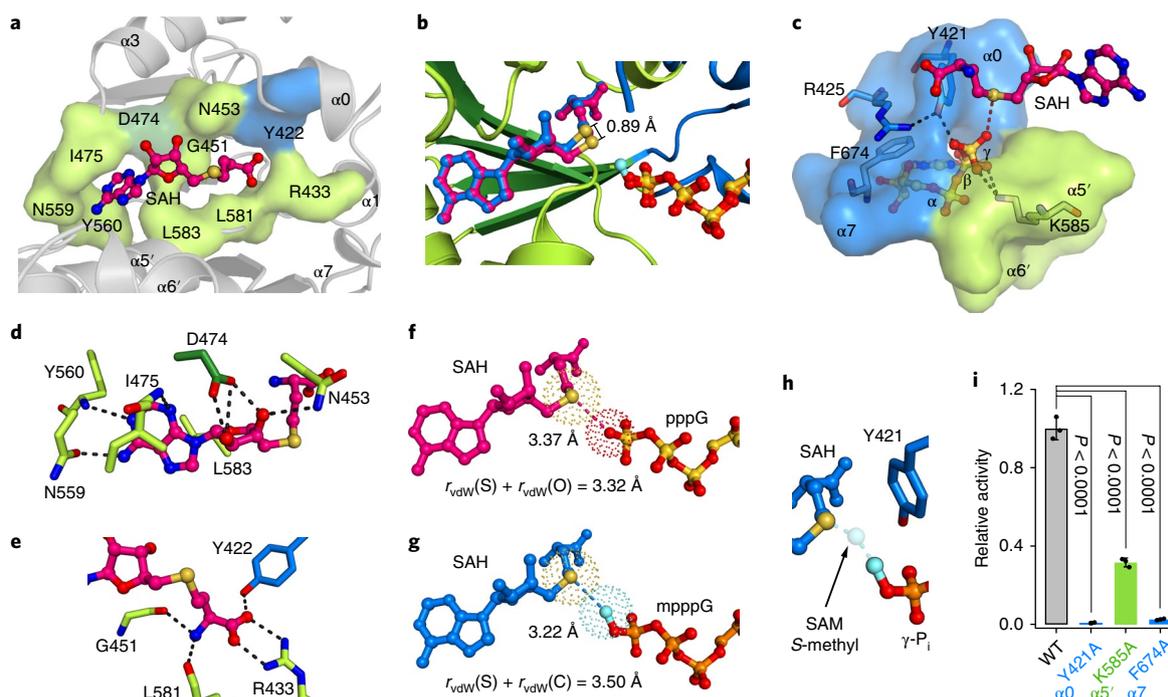


Fig. 4 | The active site of MePCEMT organizes the cofactor and triphosphate in near transition state. **a**, Surface representation of the cofactor-binding cleft in MePCE-SAH-7SK. MePCE_{MT} is shown in gray cartoon. The surface coloring is identical to that in Fig. 2. **b**, Overlay of MePCE-SAH-7SK and MePCE-SAH-me7SK structures, highlighting the positions of SAH and γ -phosphate in the active site. In MePCE-SAH-7SK, SAH is magenta, and phosphorus atoms are gold; in MePCE-SAH-me7SK, SAH is blue, phosphorus atoms are orange, and the methyl group is cyan. **c**, Triphosphate-binding residues position the γ -phosphate oxygen in line with SAH sulfur in MePCE-SAH-7SK. The triphosphate-binding tunnel is shown in surface representation and is colored as in Fig. 2. **d**, Specific hydrogen-bonding and stacking interactions between the SAH adenosyl group and MePCE_{MT}. **e**, Hydrogen bonds between the SAH homocysteine group and MePCE_{MT}. **f**, vdW radii of the sulfur atom and γ -phosphate oxygen atom in MePCE-SAH-7SK, showing the direct contact between the two. **g**, vdW radii of the sulfur atom and methyl carbon atom in MePCE-SAH-me7SK, showing the slight overlap between the two atoms. **h**, Positions of SAH sulfur, γ -meP_i, and Tyr421 in MePCE-SAH-me7SK, with the hypothetical position of the SAM methyl group shown as a transparent ball. Tyr421 stacks above the path of methyl transfer. Atoms are colored as in **g**. **i**, Methyltransferase activities of wild type (WT) MePCE_{MT} and point substitutions Y421A ($\alpha 0$), K585A ($\alpha 5'$), and F674A ($\alpha 7$) with SL1p as the RNA substrate. All activities are within one turnover and are scaled relative to that of wild type. Dots indicate values for each of the three independent reactions; bars indicate mean values; error bars, s.d. of three independent reactions performed for each protein construct. The single-factor ANOVA1 test was used to compute the *P* values, and the exact *P* values are 0.000041 (Y421A), 0.0000077 (K585A), and 0.0000072 (F674A).

SAH sulfur and the G1 γ -phosphate *O*-methyl carbon (3.22 Å) is slightly smaller than the sum of the vdW radii (3.50 Å) (Fig. 4g). These distances and geometries show a remarkable resemblance to a transition state, in which the methyl group would be located at an intermediate distance between the sulfur and γ -phosphate for in-line transfer. This proximity of reacting atoms at the active site is achieved by both the bidentate cofactor-binding cleft (Fig. 4a), which fixes the sulfur atom in place, and the triphosphate-binding tunnel (Fig. 4c), through which the γ -phosphate is anchored and positioned toward the cofactor sulfur atom.

Although we do not have structural information for the starting state of the methyl-transfer reaction (SAM and uncapped RNA), the near-transition-state geometry of both the MePCE-SAH-7SK and MePCE-SAH-me7SK structures suggests that a similar geometry would be present in a MePCE-SAM-7SK complex, as modeled in Fig. 4h. Thus, we infer the methyl-transfer reaction mechanism to be the following: SAM cofactor first binds MePCE_{MT}, and subsequent RNA binding further stabilizes and sequesters SAM in its binding cleft. The negatively charged G1 γ -phosphate oxygen is positioned in close proximity to the SAM methyl group, thus promoting nucleophilic attack at the positively charged methyl sulfonium moiety (Figs. 1a,b and 4h). Notably, the γ -phosphate oxygens are fully deprotonated and negatively charged at neutral pH²⁶, unlike protonated RNA substrates such as 2'-OH or N2H₂, and the SAM

positively charged methyl sulfonium group is less stable than the neutral SAH sulfur, thus making the MePCE methyl-transfer reaction extremely electrostatically favorable and limiting the reverse reaction. After transfer of the methyl group to the G1 γ -phosphate oxygen, both the SAH byproduct and capped-RNA product remain bound in the active site (Figs. 2b and 4b,g), which is facilitated by the extensive protein-RNA interactions discussed above.

To investigate the functional importance of interactions in the triphosphate-binding tunnel, we made individual alanine substitutions at three highly conserved residues ranging from the closest to the methyl-transfer reaction to the solvent-exposed G1 ribose—Y421A ($\alpha 0$), K585A ($\alpha 5'$), and F674A ($\alpha 7$) (Fig. 4c)—and compared their methyltransferase activity to that of wild-type MePCE_{MT} within a single turnover (Fig. 4i). Tyr421, located at the floor of the triphosphate-binding tunnel (Fig. 2b,d), forms a hydrogen bond to and buttresses the γ -phosphate (Fig. 3a,b,f) and is the closest side chain to the methyl group before transfer (Fig. 4h). Y421A was found to be nearly catalytically dead (0.6% activity) (Fig. 4i), thus indicating the critical importance of Tyr421 for formation of the active site and stabilization of the γ -phosphate. Alanine substitution of Lys585, which has hydrogen bonds to the γ - and β -phosphate oxygens (Fig. 3a-e), decreased activity to 32% (Fig. 4i), thus indicating that Lys585 has a major role within the MePCE_{MT}-triphosphate-recognition network. Partial activity is likely to be maintained by the remaining 9

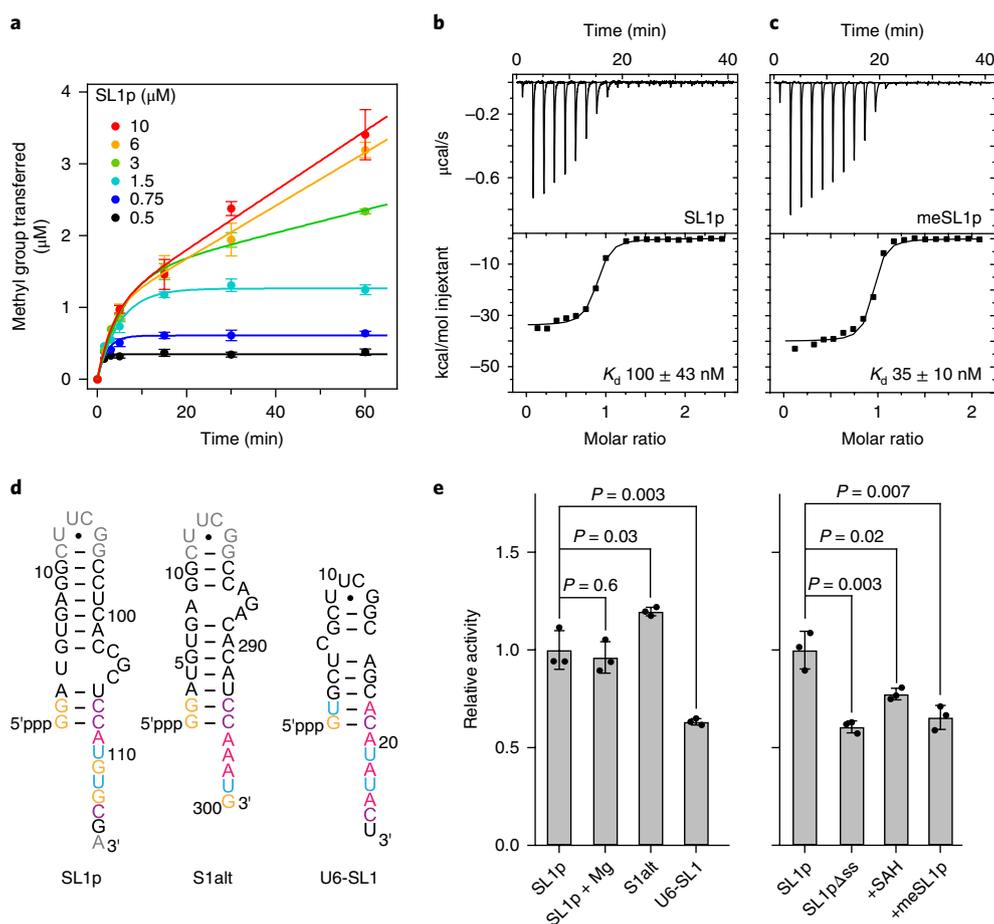


Fig. 5 | MePCE_{MT} multiple-turnover kinetics and binding experiments reveal product and byproduct inhibition and retention. **a**, MePCE_{MT} activity assay showing nonlinear progress curves for varying concentrations of SL1p. Error bars, s.d. for three independent reactions for time points ≥ 5 min. Time points taken at 1.5 min and 3 min were performed once. Data points reflect mean values for time points ≥ 5 min. The solid lines are best fits to equation (1) (Methods). **b,c**, ITC data and plots of MePCE_{MT} binding to the RNA substrates SL1p (**b**) and methylated SL1p (**c**). Three equivalents of SAH were added to MePCE_{MT} in the ITC experiments, and the numbers of replicates are included in Supplementary Table 3. **d**, Sequence and secondary structures of three RNA constructs assayed for MePCE_{MT} binding. **e**, Methyltransferase apparent activity (observed sum of single-turnover and product-inhibited activity) of MePCE_{MT} with different RNA substrates or with inhibitors added under multiple-turnover conditions. The two sets of plots indicate two sets of experiments, each performed with one enzyme stock with an SL1p control. All activities were measured in three independent reactions, and the averages were scaled to the SL1p in the absence of Mg²⁺. Dots indicate values for each of the three independent reactions; bars indicate mean values; error bars indicate s.d. from the three reactions. The single-factor ANOVA1 test was used to compute the *P* values, and the exact *P* values are shown on the graph.

of 11 direct hydrogen bonds to the triphosphate and the retention of most of the left side (helix $\alpha 5'$) of the tunnel. Alanine substitution of Phe674, which would expose the right side (helix $\alpha 7$) and destabilize the ceiling of the tunnel by disrupting the stacking interactions with Arg425 and G1 ribose, decreased activity to 2.5% (Fig. 4i). These results verify the structural and functional importance of the triphosphate-binding-tunnel residues and highlight the importance of residues in the regions that become ordered upon RNA binding.

MePCE binds tightly to capped RNA. Under single-turnover conditions, SL1p was fully monomethylated in less than 15 min (Fig. 1d and Supplementary Fig. 3). In time-dependent activity assays at different SL1p concentrations, we similarly observed quantitative methylation under single-turnover conditions (MePCE_{MT} in molar excess to RNA); however, nonlinear behavior was observed under multiple-turnover conditions (Fig. 5a). These data reflect an initial fast phase corresponding to a quantitative single turnover, followed by a slow phase corresponding to a second turnover event that is limited by product release²⁷. To explain these results, we first investigated the equilibrium dissociation constants (K_d) of

the cofactor SAM and the byproduct SAH to MePCE_{MT} by using isothermal titration calorimetry (ITC) (Supplementary Fig. 9a,b). MePCE_{MT} bound SAM with 13-fold lower affinity than that to SAH, thus suggesting that SAH may be a product inhibitor to MePCE, as seen in some SAM-dependent Rossmann-fold MTases^{28–32}.

Using ITC, we next investigated the binding of MePCE_{MT} to various RNA substrates including SL1p Δ ss (no ssRNA), SL1p Δ 1–3 (no internal loop), and three pairs of unmethylated and methylated constructs: SL1p (Fig. 5b,c), S1alt from the proposed 5' stem of the alternative circular 7SK model, and SL1 of the U6 small nuclear RNA (U6-SL1 and meU6-SL1), the only other predicted substrate of MePCE in humans (Fig. 5d, Supplementary Fig. 9, Table 1 and Supplementary Table 3). To ensure that only binding energies were measured, we performed the experiments in the presence of SAH rather than SAM. SL1p Δ ss had no detectable binding by ITC ($K_d \gg 1 \mu\text{M}$), thus indicating that the 3' ssRNA overhang is required for high-affinity binding. In contrast, no difference in binding affinity was observed between SL1p and SL1p Δ 1–3, a finding consistent with the lack of interactions observed between MePCE_{MT} and the internal asymmetric loop in the crystal structures.

Table 1 | Thermodynamic parameters of MePCE binding to cofactor and RNA substrates

	K_d (nM)	N
Ligand		
SAM	1,580 ± 180	0.8 ± 0.3
SAH	122 ± 35	0.8 ± 0.1
7SK^a		
SL1p	100 ± 43	0.9 ± 0.1
meSL1p	35 ± 10	0.9 ± 0.1
S1alt	800 ± 99	0.6 ± 0.1
meS1alt	146 ± 42	0.7 ± 0.1
SL1pΔss	NB ^b	
SL1pΔ1-3	100 ± 50	0.8 ± 0.1
U6^a		
U6-SL1	30 ± 11	0.8 ± 0.1
meU6-SL1	27 ± 4	0.8 ± 0.1

^aMePCE is in the presence of three equivalents of SAH. ^bNB, no-binding individual replicates were fit in Origin 7 with a one-site fitting equation. The parameters determined from n fits were averaged to determine the mean and s.d., where n is as follows: SAM ($n=2$); SAH ($n=4$); SL1p ($n=6$); meSL1p ($n=3$); S1alt ($n=2$); meS1alt ($n=3$); SL1pΔss ($n=2$); SL1pΔ1-3 ($n=4$); U6-SL1 ($n=4$); meU6-SL1 ($n=2$)

U6-SL1 had a fourfold-higher binding affinity than 7SK SL1p to MePCE_{MT} (Table 1, Supplementary Table 3 and Supplementary Fig. 9). Like SL1p, U6-SL1 has a terminal G-C base pair, and although the penultimate base pair is a U-A instead of a G-C, equivalent hydrogen bonds could be made between protein and base. The first 6 nt of the 3' ssRNA sequence are alternating purine and pyrimidine nucleotides (AUAUACU), as compared with AUGUGCGC in 7SK SL1p (Fig. 5d). In U6, the two A nucleotides are expected to stack similarly to SL1p G111 and G113 (ref. ³³), and the top A nucleotide, like SL1p G111, could accommodate a hydrogen bond to Ser482 (Fig. 3f,k). In contrast, the binding affinity of MePCE_{MT} to S1alt was eightfold lower than that to SL1p. S1alt and SL1p both have two G-C base pairs, but in S1alt, the 3' ssRNA region is not alternating purine-pyrimidine and is 3 nt shorter (AAAUG) (Fig. 5d). Given that SL1p bound MePCE_{MT} with higher affinity and that MePCE is likely to associate with 7SK as it is being transcribed before the 5' end and the region near the 3' end can base-pair⁹, these data suggest that the linear 7SK model is the substrate for MePCE, in agreement with biochemical experiments⁸.

MePCE_{MT} bound to meSL1p with approximately twofold-higher affinity than to SL1p ($\Delta\Delta G_{\text{cap}}$ 0.6 ± 0.3 kcal/mol (mean ± s.d.)) (Table 1, Supplementary Table 3 and Supplementary Fig. 9). The unexpected finding that the product bound with higher affinity than the substrate explains how 7SK is retained by MePCE after capping. MePCE bound to S1alt almost an order of magnitude more weakly than to SL1p, but bound meS1alt with an affinity comparable to that toward uncapped SL1p, with an approximately fivefold higher affinity for meS1alt than S1alt ($\Delta\Delta G_{\text{cap}}$ 1.0 ± 0.2 kcal/mol). These results do not exclude an assembly model in which MePCE binds circular 7SK after capping of linear 7SK.

Although MePCE_{MT} bound tightly to both capped and uncapped U6-SL1 ($\Delta\Delta G_{\text{cap}}$ 0.0 ± 0.2 kcal/mol) with comparable affinity, the affinity was approximately the same as for meSL1p, thus suggesting that MePCE remained bound to U6 small nuclear RNA after capping, in agreement with MePCE pulldown assays^{6,7}.

Catalytic turnover is inhibited by product and byproduct binding. The above results indicate that both capped RNA and SAH byproduct can act as inhibitors. To validate the proposed

product-inhibition mechanism, we performed methyltransferase activity assays to determine whether binding affinities might be correlated with turnover (Fig. 5e and Supplementary Fig. 10). As discussed above, MePCE_{MT} exhibited fast kinetics within one turnover and a slow phase for the subsequent turnovers, which is likely to be limited by product release (Fig. 5a). As expected, addition of Mg²⁺ to the methyltransferase assay had no statistically significant effect on activity. Under multiple turnover conditions, S1alt had 19% higher apparent activity than SL1p. Because meS1alt had lower binding affinity than meSL1p, these results suggest that meS1alt has weaker product inhibition, thus allowing MePCE_{MT} to proceed to multiple turnovers. In contrast, the apparent activity on U6-SL1 was 36% less than that on SL1p. We note that under multiple turnover conditions with U6-SL1, one stoichiometric methylation was completed at approximately 15 min, thus indicating that the decreased activity with U6-SL1 is due not to an abortive methyl-transfer reaction but instead to a decreased rate of product release, in agreement with the higher binding affinity of MePCE for meU6-SL1 than meSL1p (Supplementary Table 3). SL1pΔss had no detectable binding to MePCE_{MT} by ITC but was capped, although with 39% lower activity. This moderately decreased activity probably reflects both weaker protein-RNA binding and drastically decreased product inhibition. As further support for the product-inhibition model, we doped the MTase assays with SAH or meSL1p, which decreased activity by 23% and 34%, respectively.

In summary, the enzymatic activity of MePCE is an extreme case of product inhibition wherein both the byproduct and product have higher affinity toward the MTase domain than that of the corresponding cofactor and RNA substrate. The retention of a monomethylated RNA product also explains why MePCE methylates only a single γ -phosphate oxygen. Together, the structural, binding, and enzymatic data demonstrate that MePCE generally functions as a single-turnover enzyme and explain the dual role of MePCE in 7SK RNP biogenesis in capping the 5' γ -phosphate to protect RNA from degradation and as a constitutive component of the 7SK core RNP.

Comparison to other MTases highlights unique features of MePCE. We compared substrate- and cofactor-binding sites of MePCE to those in structures of two other RNA-capping MTases bound to substrate: VP39 (PDB 1AV6), a well-studied mRNA ribose O2'-cap1 MTase, and Tgs1, an mRNA TMG-capping MTase (PDB 3GDH) (Supplementary Fig. 11). In contrast to the large RNA-binding surface on MePCE, VP39 and Tgs1 have base-specific contacts to only two or one nucleotides, respectively. MePCE has a large positively charged surface area and buried active site with an enclosed cofactor, whereas VP39 and Tgs1 have much smaller positively charged surface areas and relatively shallow active sites with solvent-accessible cofactor. In the MePCE active site, methyl-donor and methyl-acceptor atoms are within vdW radius (Fig. 4f,g and Supplementary Fig. 11g). The donor and acceptor atoms in VP39 and Tgs1 are farther apart, probably because both the ribose 2'-OH group (VP39) and the N2 amino group (Tgs1) require a deprotonation step before methyl transfer (Supplementary Fig. 11h,i).

BCDIN3D^{34,35}, a recently discovered MePCE paralog that methylates the 5'- α -phosphate of tRNA^{His} and pre-microRNA, has substantial sequence homology to MePCE, including helix $\alpha 0$ tyrosine-wheel residues and residues involved in 5'-end and cofactor recognition (Supplementary Fig. 7), thus suggesting that it may recognize its 5'-hairpin-single-stranded substrates in a similar way to MePCE. In contrast, the MTase METTL16, which also modifies 7SK and U6, lacks helix $\alpha 0$ and conserved residues involved in 5'-end recognition by MePCE, in agreement with its internal N⁶-methyladenosine-modification function^{36,37}. This comparison of RNA-capping MTases highlights the unique features of MePCE: a large substrate-recognition surface; enclosed active site; and substrate and cofactor poised in near-transition-state geometry.

Discussion

RNA 5' caps are required for RNA biogenesis, stability, splicing, cellular localization, and protein interactions^{1,2}. However, the molecular mechanisms of the MTases responsible for capping have not been well characterized, primarily because of a lack of structural information on substrate recognition. The MePCE–SAH–me7SK crystal structure is, to our knowledge the first reported structure of an RNA MTase in complex with its capped RNA product at the active site. This structure, together with the MePCE–SAH–7SK crystal structure, reveals a reaction mechanism in which the reactants and products remain in a near-transition-state geometry in the active site for the entire reaction process. Comparison to the substrate-free MePCE–SAH structure shows a disorder-to-order conformational change for the N- and C-terminal regions after binding RNA, thus encapsulating and blocking diffusional exit of the cofactor and positioning the substrate for reaction. The structures also reveal atomic details of an extensive interaction network that recognizes 5'-triphosphate and base-paired and single-stranded nucleotides through both unique and canonical Rossmann-fold methyltransferase α -helices. Notably, the binding affinity of capped 7SK SL1p to MePCE_{MT} with SAH is higher than that of the uncapped RNA. The higher affinity of products over reactants to MePCE explains how MePCE becomes a stable component of the 7SK RNP after catalysis. The structural and mechanistic comparisons of MePCE to other RNA MTases described above further explain why MePCE is inhibited by both product and byproduct, and how it retains product in the active site. Although there are apparently some protein MTases that retain their products (for example, protein L-isoaspartyl methyltransferase³⁰ and protein arginine methyltransferase 3 (PRMT3), through its N-terminal zinc-finger domain³⁸), MePCE is the only known RNA MTase that retains product after catalysis.

MePCE concentrates at the 7SK and U6 promoter regions⁹, and 5' capping is likely to occur cotranscriptionally after the specific helix-ssRNA region is transcribed. Our binding analysis and activity assays indicate that U6-SL1 interacts with MePCE with higher affinity than does 7SK and is capped by MePCE in a similar product-inhibited manner. These observations validate U6 as a MePCE substrate and explain the association of MePCE with U6 in vivo (through pulldown assays), at least before formation of the U4/U5/U6 tri-small nuclear RNP^{6,7,39}. Given that an RNA construct corresponding to the 5' end in the circular 7SK model bound MePCE_{MT} with approximately eightfold lower affinity than a construct corresponding to the linear 7SK model, and in vivo the circular 7SK model cannot form until transcription of full-length 7SK is nearly complete, we propose that the linear 7SK model is the substrate for MePCE capping. Larp7, which has strand melting and annealing properties⁴⁰ common to LARPs and La protein^{41,42}, may subsequently remodel 7SK upon assembly, thereby altering the MePCE–RNA interface and promoting the stable association between MePCE and Larp7. The retention of MePCE on 7SK as a core component of the 7SK RNP highlights the unique nature of 7SK: it is capped at the 5' end by a specialized MTase that remains bound after catalysis, and the same MTase is then joined with a 3'-end-associating La-related protein, thus forming a stable ternary RNP. This work provides a structural and mechanistic basis for the role of MePCE in the first steps of 7SK RNP assembly.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41589-018-0188-z>.

Received: 13 August 2018; Accepted: 6 November 2018;
Published online: 17 December 2018

References

1. Byszewska, M., Śmietański, M., Purta, E. & Bujnicki, J. M. RNA methyltransferases involved in 5' cap biosynthesis. *RNA Biol.* **11**, 1597–1607 (2014).
2. Schapira, M. Structural chemistry of human RNA methyltransferases. *ACS Chem. Biol.* **11**, 575–582 (2016).
3. Singh, R. & Reddy, R. Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proc. Natl Acad. Sci. USA* **86**, 8280–8283 (1989).
4. Shumyatsky, G. P., Tillib, S. V. & Kramerov, D. A. B2 RNA and 7SK RNA, RNA polymerase III transcripts, have a cap-like structure at their 5' end. *Nucleic Acids Res.* **18**, 6347–6351 (1990).
5. Shimba, S., Buckley, B., Reddy, R., Kiss, T. & Filipowicz, W. Cap structure of U3 small nucleolar RNA in animal and plant cells is different. gamma-monomethyl phosphate cap structure in plant RNA. *J. Biol. Chem.* **267**, 13772–13777 (1992).
6. Jeronimo, C. et al. Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol. Cell* **27**, 262–274 (2007).
7. Muniz, L., Egloff, S. & Kiss, T. RNA elements directing in vivo assembly of the 7SK/MePCE/Larp7 transcriptional regulatory snRNP. *Nucleic Acids Res.* **41**, 4686–4698 (2013).
8. Brogie, J. E. & Price, D. H. Reconstitution of a functional 7SK snRNP. *Nucleic Acids Res.* **45**, 6864–6880 (2017).
9. Xue, Y., Yang, Z., Chen, R. & Zhou, Q. A capping-independent function of MePCE in stabilizing 7SK snRNA and facilitating the assembly of 7SK snRNP. *Nucleic Acids Res.* **38**, 360–369 (2010).
10. Yang, Z., Zhu, Q., Luo, K. & Zhou, Q. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* **414**, 317–322 (2001).
11. Nguyen, V. T., Kiss, T., Michels, A. A. & Bensaude, O. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* **414**, 322–325 (2001).
12. Kohoutek, J. P-TEFb: the final frontier. *Cell. Div.* **4**, 19 (2009).
13. Krueger, B. J. et al. LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res.* **36**, 2219–2229 (2008).
14. Markert, A. et al. The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. *EMBO Rep.* **9**, 569–575 (2008).
15. He, N. et al. A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol. Cell* **29**, 588–599 (2008).
16. Michels, A. A. et al. Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J.* **23**, 2608–2619 (2004).
17. Muniz, L., Egloff, S., Ughy, B., Jády, B. E. & Kiss, T. Controlling cellular P-TEFb activity by the HIV-1 transcriptional transactivator Tat. *PLoS Pathog.* **6**, e1001152 (2010).
18. Wassarman, D. A. & Steitz, J. A. Structural analyses of the 7SK ribonucleoprotein (RNP), the most abundant human small RNP of unknown function. *Mol. Cell. Biol.* **11**, 3432–3445 (1991).
19. Marz, M. et al. Evolution of 7SK RNA and its protein partners in metazoa. *Mol. Biol. Evol.* **26**, 2821–2830 (2009).
20. Yazbeck, A. M., Tout, K. R. & Stadler, P. F. Detailed secondary structure models of invertebrate 7SK RNAs. *RNA Biol.* **15**, 158–164 (2018).
21. Uchikawa, E. et al. Structural insight into the mechanism of stabilization of the 7SK small nuclear RNA by LARP7. *Nucleic Acids Res.* **43**, 3373–3388 (2015).
22. Eichhorn, C. D., Chug, R. & Feigon, J. hLARP7 C-terminal domain contains an xRRM that binds the 3' hairpin of 7SK RNA. *Nucleic Acids Res.* **44**, 9977–9989 (2016).
23. Eichhorn, C. D., Yang, Y., Repeta, L. & Feigon, J. Structural basis for recognition of human 7SK long noncoding RNA by the La-related protein Larp7. *Proc. Natl Acad. Sci. USA* **115**, E6457–E6466 (2018).
24. Singh, R., Gupta, S. & Reddy, R. Capping of mammalian U6 small nuclear RNA in vitro is directed by a conserved stem-loop and AUUAUC sequence: conversion of a noncapped RNA into a capped RNA. *Mol. Cell. Biol.* **10**, 939–946 (1990).
25. Cosgrove, M. S., Ding, Y., Rennie, W. A., Lane, M. J. & Hanes, S. D. The Bin3 RNA methyltransferase targets 7SK RNA to control transcription and translation. *Wiley Interdiscip. Rev. RNA* **3**, 633–647 (2012).
26. Cheng, H., Sukal, S., Callender, R. & Leyh, T. S. γ -phosphate protonation and pH-dependent unfolding of the Ras.GTP.Mg²⁺ complex: a vibrational spectroscopy study. *J. Biol. Chem.* **276**, 9931–9935 (2001).
27. Copeland, R. A. *Enzymes: a Practical Introduction to Structure, Mechanism, and Data Analysis*. 188–265 (John Wiley & Sons: New York, 2000).
28. Husain, N. et al. Structural basis for the methylation of A1408 in 16S rRNA by a panaminoglycoside resistance methyltransferase NpmA from a clinical isolate and analysis of the NpmA interactions with the 30S ribosomal subunit. *Nucleic Acids Res.* **39**, 1903–1918 (2011).
29. Shi, Y. Q. & Rando, R. R. Kinetic mechanism of isoprenylated protein methyltransferase. *J. Biol. Chem.* **267**, 9547–9551 (1992).

30. Johnson, B. A. & Aswad, D. W. Kinetic properties of bovine brain protein L-isoaspartyl methyltransferase determined using a synthetic isoaspartyl peptide substrate. *Neurochem. Res.* **18**, 87–94 (1993).
31. Swiercz, R., Person, M. D. & Bedford, M. T. Ribosomal protein S2 is a substrate for mammalian PRMT3 (protein arginine methyltransferase 3). *Biochem. J.* **386**, 85–91 (2005).
32. Clarke, S. & Banfield, K. in *Homocysteine in Health and Disease* (eds. Carmel, R. & Jacobsen, D. W.) 63–78 (Cambridge University Press, Cambridge, 2001).
33. Turner, D. H. & Mathews, D. H. NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucleic Acids Res.* **38**, D280–D282 (2010).
34. Martinez, A. et al. Human BCDIN3D monomethylates cytoplasmic histidine transfer RNA. *Nucleic Acids Res.* **45**, 5423–5436 (2017).
35. Xhemalce, B., Robson, S. C. & Kouzarides, T. Human RNA methyltransferase BCDIN3D regulates microRNA processing. *Cell* **151**, 278–288 (2012).
36. Warda, A. S. et al. Human METTL16 is a N⁶-methyladenosine (m⁶A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. *EMBO Rep.* **18**, 2004–2014 (2017).
37. Pendleton, K. E. et al. The U6 snRNA m⁶A methyltransferase METTL16 regulates SAM synthetase intron retention. *Cell* **169**, 824–835.e14 (2017).
38. Choi, S., Jung, C.-R., Kim, J.-Y. & Im, D.-S. PRMT3 inhibits ubiquitination of ribosomal protein S2 and together forms an active enzyme complex. *Biochim. Biophys. Acta* **1780**, 1062–1069 (2008).
39. Didychuk, A. L., Butcher, S. E. & Brow, D. A. The life of U6 small nuclear RNA, from cradle to grave. *RNA* **24**, 437–460 (2018).
40. Hussain, R. H., Zawawi, M. & Bayfield, M. A. Conservation of RNA chaperone activity of the human La-related proteins 4, 6 and 7. *Nucleic Acids Res.* **41**, 8715–8725 (2013).
41. Bayfield, M. A., Yang, R. & Maraia, R. J. Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim. Biophys. Acta* **1799**, 365–378 (2010).
42. Maraia, R. J., Mattijssen, S., Cruz-Gallardo, I. & Conte, M. R. The La and related RNA-binding proteins (LARPs): structures, functions, and evolving perspectives. *Wiley Interdiscip. Rev. RNA* **8**, e1430 (2017).

Acknowledgements

This work was supported by NIH grant GM107567 to J.F. and American Cancer Society Postdoctoral Fellowship 126777-PF-14-179-01-DMC to C.D.E. We acknowledge NMR equipment grant NIH S10OD016336 and DOE grant DE-FC0302ER63421 for partial support of NMR and X-ray core facilities. The authors thank M. Capel, K. Rajashankar, N. Sukumar, F. Murphy, I. Kourinov, and J. Schuermann of Northeastern Collaborative Access Team (NE-CAT) beamline ID-24 at the Advanced Photon Source (APS) of Argonne National Laboratory, which are supported by NIH grants P41 RR015301 and P41 GM103403. Use of the APS is supported by DOE under contract DE-AC02-06CH11357. We thank Y. Chen for the help with mass spectrometry data collection and analysis in the UCLA MIC mass spectrometry facility, supported in part by NIH instrumentation grant 1S10OD016387; M. Collazo for help with crystal optimization; and K. Jain, J. Lowenson, and S. Clarke for helpful discussions.

Author contributions

Y.Y. and C.D.E. designed and performed experiments, analyzed data, and wrote the paper; Y.W. prepared RNA samples; D.C. helped with X-ray data collection and processing; and J.F. supervised all aspects of the work, analyzed data, and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41589-018-0188-z>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to J.F.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2018

Methods

Protein expression and purification. The pET28a-LIC recombinant plasmid with gene encoding the hexahistidine-tagged human MePCE (FLJ20257) MTase domain (residues 400–689) was a gift from C. Arrowsmith (Addgene plasmid 25304) and was transformed into *Escherichia coli* BL21-Gold(DE3) (Agilent Technologies) for protein expression. Protein point substitutions were carried out with a Q5 site-directed mutagenesis kit (NEB) and verified by DNA sequencing. Bacterial cultures were grown in minimal medium at 37 °C to an OD₆₀₀ of 0.6, then transferred to 18 °C for 1 h before induction with 0.5 mM IPTG for 18 h. Cells were pelleted, resuspended with resuspension buffer (buffer R, 20 mM HEPES, pH 8, 1 M NaCl, 1 mM TCEP, 15 mM imidazole, 5% glycerol, and 1 mM PMSF) supplemented with lysozyme, then sonicated. Cell lysate was clarified by centrifugation and filtration, and the hexahistidine-tagged proteins were purified with a Ni–Sephareose affinity column (HisTrap HP; GE Healthcare) and further purified by size-exclusion chromatography (SEC; HiLoad 26/600 Superdex 75; GE Healthcare) in crystallization buffer (buffer C, 20 mM HEPES, pH 7.5, 250 mM KCl, and 1 mM TCEP) or binding buffer (buffer B, 20 mM HEPES, pH 7.5, 150 mM KCl, and 1 mM TCEP). The concentration of protein peak fractions was measured, and the protein was stabilized by addition of a two- or threefold molar-excess of *S*-(5′-adenosyl)-*L*-homocysteine (SAH) (Sigma A9384) or *S*-(5′-adenosyl)-*L*-methionine *p*-toluenesulfonate salt (SAM) (Sigma A2408; ≥80% purity) cofactors.

In vitro transcription of RNA. RNA samples were prepared by in vitro transcription as described previously²² with chemically synthesized DNA templates (Integrated DNA Technologies). Briefly, the transcription reaction was incubated for 4–6 h at 37 °C and was followed by purification by 15–20% denaturing PAGE, excision of RNA and elution with an Elutrap device (GE Waters). Collected RNA eluents were buffer exchanged into sterilized nanopure water supplied with counterions by the addition of high-salt buffer (1.5 M KCl), and briefly buffer exchanged into nanopure water before being annealed by heating to 95 °C for 5 min and snap cooling on ice for 1 h. RNAs were then desalted into sterilized nanopure water or buffer C/B and concentrated with Amicon ultracentrifugal filters (3-kDa MWCO, Millipore). Stock solutions of RNA were stored in a –20 °C freezer.

Methylated-RNA sample preparation. MePCE_{MT} RNA, and SAM were mixed in molar ratios of 1:1.3 or 0.5:1.3 and incubated at 37 °C for >24 h in buffer B. RNA was purified by phenol–chloroform extraction followed by ethanol precipitation and buffer exchanged into nanopure water. Sample homogeneity was verified by mass spectrometry (Supplementary Figs. 3 and 12–14).

Crystallization and data collection. All crystallization trials were performed with the hanging-drop vapor-diffusion method at 4 °C with MePCE_{MT} in complex with 7SK SL1p RNA and SAM or SAH at a concentration of 8–10 mg/mL. The molar ratio of cofactor-bound MePCE_{MT} and SL1p RNA was determined with electrophoretic mobility shift assays, in which the RNA reached stoichiometric saturation with protein. In our hands, the SL1p RNA was saturated by a 1.5-fold molar excess of MePCE_{MT} in the presence of SAM or SAH. For assembly of the MePCE_{MT}–SL1p complex, a dilute protein solution from the SEC elution in buffer C with added SAM or SAH was combined with RNA in buffer C and concentrated with 3-kDa MWCO Amicon ultracentrifugal filters (Millipore) at 4 °C. The final molar concentration of MePCE_{MT} ($\epsilon = 51,715 \text{ l mol}^{-1} \text{ cm}^{-1}$ at A_{280}) was 0.19–0.24 mM, and that of SL1p RNA ($\epsilon = 317,600 \text{ l mol}^{-1} \text{ cm}^{-1}$ at A_{260}) was 0.13–0.16 mM.

Before crystallization, the protein–RNA solutions were mixed with the well-solution buffers at a 2:1 (vol/vol) ratio. For the MePCE–SAH–7SK structure, single crystals grew in approximately 1 week with a well solution of 0.1 M phosphate/citrate, pH 4.2, 0.214 M lithium sulfate and 17.25% (wt/vol) PEG 1000. The crystals were transferred in a cryoprotectant containing 50% glycerol and flash frozen in liquid nitrogen. The structure was solved by molecular replacement with Phaser⁴³ with the MePCE–SAH structure (PDB 5UNA, chain E) as a search model. RNA density for the 5′-triphosphate and the first three bases was apparent in the electron density map. RNA was built into the electron density with Coot⁴⁴ and iteratively refined to 2.0-Å resolution with Phenix⁴⁵.

For the MePCE–SAH–me7SK structure, the sample was prepared as described above with SAM instead of SAH. Single crystals grew in approximately 1 week and continued to grow over 2 months in a well solution of 0.1 M phosphate/citrate, pH 4.2, 0.2 M lithium sulfate, and 25% (wt/vol) PEG 1000. The crystals were transferred to an undiluted drop of well solution supplemented with 1.2 mM MgCl₂, soaked for 2 h and flash frozen in liquid nitrogen. The structure was solved by molecular replacement with Phaser with the refined MePCE–SAH–7SK model. The additional methyl group on the γ -phosphate was apparent from both $2F_o - F_c$ and $F_o - F_c$ maps (Supplementary Fig. 4). The γ -phosphate monomethylated GTP geometry was built with phenix.elbow⁴⁶, and the final structure was refined to 2.1-Å resolution with Phenix. Independent soaking experiments were also performed with both MePCE–SAH–me7SK and MePCE–SAH–7SK crystals in the following conditions: 1.2 mM MgCl₂, 1.2 mM SmCl₃, 2.4 mM MgCl₂, and 2.4 mM SmCl₃. Two soaking times—overnight or 2 h—were used for all conditions. MePCE–SAH–me7SK crystal from the 2-h, 1.2 mM MgCl₂ condition was used for final refinement and deposition.

All data were collected on beamline 24-ID-C (NE-CAT) at the Advanced Photon Source at Argonne National Laboratory (APS-ANL, Argonne, IL, USA) at a wavelength of 0.9792 Å and temperature of 100 K, with a Dectris Pilatus 6M detector. Data were indexed and integrated with XDS and scaled with XSCALE⁴⁷. Final iterative rounds of model building and refinement were performed with Coot and Phenix with TLS refinement⁴⁸. Composite refine omit maps⁴⁹ were generated with Phenix, and PyMOL 2.0.6 was used to generate structural figures. Crystallization data collection and refinement statistics are presented in Supplementary Table 1. The electrostatic potential of the MePCE–SAH–7SK, VP39, and Tgs1 structures was calculated with Adaptive Poisson-Boltzmann Solver (APBS) through the PDB2PQR⁵⁰ and APBS⁵¹ webserver.

NMR spectroscopy. Capped RNA samples for ¹³C NMR experiments were generated by incubating freshly purified MePCE_{MT}, *S*-methyl-¹³C SAM (99 atom % ¹³C; Sigma Aldrich/ISOTEC), and RNA samples in a molar ratio of 1:3:1 in buffer C at 37 °C for 1 h. The reactions were quenched, and RNA was separated by phenol–chloroform extraction. The RNA samples were further purified by ethanol precipitation, air dried, and dissolved in NMR buffer. The uncapped and capped RNA NMR samples contained 50–100 μM RNA, in a buffer of 10 mM sodium phosphate, pH 6.2, 60 mM KCl, 5% (vol/vol) D₂O, and 0.6 mM Na₃. All NMR spectra were collected at 15 °C with Bruker DRX 500 MHz and Avance 800 MHz spectrometers equipped with cryogenically cooled probes. Data were collected with TopSpin (Bruker), processed with NMRPipe⁵², and analyzed in NMRFAM-Sparky⁵³.

1D ¹H flipback NMR spectra⁵⁴ were used to quantify (*S,S*) and (*R,S*) diastereomers of SAM. SAM salts were dissolved in the RNA NMR buffer. The peaks at 3.01 p.p.m. and 2.97 p.p.m. were assigned to methyl protons from (*S,S*) and (*R,S*) diastereomers, respectively, on the basis of the decrease in relative intensity of the (*S,S*) peak after a 7-d incubation and previously reported chemical-shift values⁵⁵. SAM *p*-toluenesulfonate salt and SAM iodide salt (Sigma A4377) were found to contain 76.63% and 73.79% (*S,S*) diastereomer, respectively. All SAM concentrations in the methyltransferase and ITC experiments refer to the total concentration of SAM, on the basis of the A_{260} . However, the SAM concentrations were adjusted to the active (*S,S*) diastereomer in calculation of molar methyl transfer.

Mass spectrometry analysis of RNA and purified MePCE_{MT} protein. The uncapped and capped RNA samples were desalted with ammonium acetate precipitation to replace the Na⁺ on the RNA backbone⁵⁶. The desalted RNA samples were dissolved in ultrapure water to 20 μM and analyzed with the direct infusion method on a Q Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific) in negative mode with 70,000 resolving power at 200 *m/z*. 100 scans were acquired to improve the signal-to-noise ratio.

SAM/SAH standard mixture and freshly purified MePCE_{MT} protein were analyzed on a ZORBAX C3 column, 3.0 \times 150 mm, with bead size 3.5 μm (Agilent Technologies) with Agilent 6530 Q-TOF ESI with a 1260 Infinity LC. Buffers A and B were 0.1% formic acid in water and 0.1% formic acid in acetonitrile, respectively. A linear gradient from 1% to 99% B was applied for 8 min, and eluates were analyzed with an ESI-TOF mass spectrometer in positive mode.

Methyltransferase assay. In vitro RNA methyltransferase assays were adapted from established protocols^{57,58}. All experiments were performed in a reaction buffer of 50 mM HEPES, pH 7.5, 150 mM KCl, and 1 mM TCEP at 37 °C with 1 μM MePCE_{MT} and 20 μM SAM. SAM is a mixture of 20:1 molar ratio of *S*-(5′-adenosyl)-*L*-methionine *p*-toluenesulfonate salt (SAM) (Sigma A2408; ≥80% purity) to *S*-(5′-adenosyl)-*L*-[methyl-³H]methionine ([me³H]SAM) (stock solution of 6.7 μM (81.9 Ci/mmol) in 10 mM H₂SO₄/EtOH (9:1, vol/vol); PerkinElmer Life Sciences). Reaction mixtures of 25 μL without RNA substrates in thin-walled PCR tubes were incubated at 37 °C for 5 min, and reactions were initiated by addition of 5 μL of RNA stock solutions (6 \times concentrations in the range of 3–60 μM). Reactions were quenched by prompt heating to 98 °C and were maintained for 2 min in a preheated thermocycler at different time points. 25 μL of the 30 μL reaction mixtures was spotted on Amersham Hybond-N⁺ membrane (GE Healthcare), air dried for 30 min, washed twice with 100 mL of 50 mM NaHCO₃, pH 9, for 45 min and air dried for 90 min. The membranes were placed in vials and soaked with 5 mL of Safety-Solve scintillation solution (Research Products International 111177) overnight and counted for three cycles of 5 min with a Beckman scintillation counter. Product-RNA concentration was quantified by conversion of c.p.m. counts to molar concentrations with 58% counting efficiency for the scintillation counter, SAM-to-[me³H]SAM molar ratio, and (*S,S*) diastereomer purity of 76.63%, as determined from NMR.

The progress curves with varying RNA concentrations were fit to a two-step kinetics equation as follows:

$$[\text{meSL1p}] = A[1 - e^{-kt}] + Vt \quad (1)$$

where A is the burst amplitude, k is the burst rate, and V is the reaction speed for the linear slow phase. Nonlinear regression analysis was performed in Igor Pro 8 (WaveMetrics). For comparing protein substitutions with WT MePCE_{MT}, 10 μM of

SL1p was used, and the relative activities within one turnover were calculated from the 2-min time points. For comparing different RNA substrates and the effects of the product inhibitors under multiple-turnover conditions, all RNA concentrations were 10 μM and the SAH/meSL1p inhibitor concentrations were 5 μM . The 30-min time points were used for calculating multiple-turnover relative activities. To compute P values between different groups of datasets, the single-factor ANOVA1 test associated with Igor Pro 8 (WaveMetrics) was used. The between-group number of degrees of freedom was $n - 1$, where n is the number of groups.

Isothermal titration calorimetry. The binding dissociation coefficients (K_D) for binding of MePCE_{MT} to cofactors SAM and SAH and substrate RNA constructs were determined with a MicroCal 200 ITC instrument (GE Health Sciences). All samples were in buffer B, and experiments were performed at 298 K. Buffer B was used for the final size-exclusion-chromatography purification step, and MePCE_{MT} was used for ITC experiments within 48 h of purification without further concentration or buffer exchange. MePCE_{MT} was added to the cell in concentrations of 8–9 μM and 100–250 μM titrant (SAM, SAH, or RNA). SAM and SAH were dissolved directly into buffer B, and the concentration was determined by UV spectroscopy with the extinction coefficient $\epsilon = 15,400 \text{ l mol}^{-1} \text{ cm}^{-1}$ at A_{260} . RNA was extensively buffer-exchanged into buffer B. Calorimetric data were fit with ORIGIN 7 (MicroCal). The binding parameters stoichiometry (N), entropy (ΔS), enthalpy (ΔH), and association constant (K) were kept as floating variables during each fit. Experiments were performed two to six times, each experiment was fit individually, and binding parameters were averaged. The thermodynamic parameters and number of replicates are listed in Supplementary Table 3.

Multiple sequence alignment. Sequences from the MTase domains of human, mouse, *Danio rerio*, *Drosophila melanogaster*, *Arabidopsis thaliana*, pea, and *Schizosaccharomyces pombe* MePCE; and human BCDIN3 and Mettl16 were aligned with the NCBI constraint-based multiple alignment tool (COBALT)⁵⁹ and analyzed with ESPrpt 3.0 (ref. ⁶⁰).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Atomic coordinates and structure factors have been deposited in the Protein Data Bank under the following accession codes: PDB 6DCB (MePCE–SAH–7SK) and PDB 6DCC (MePCE–SAH–me7SK). All other data generated or analyzed in this study are included in the published article (and its supplementary information files) or are available from the corresponding author upon reasonable request.

References

43. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

44. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
45. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 (2010).
46. Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D. Biol. Crystallogr.* **65**, 1074–1080 (2009).
47. Kabsch, W. Xds. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132 (2010).
48. Painter, J. & Merritt, E. A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D. Biol. Crystallogr.* **62**, 439–450 (2006).
49. Hodel, A., Kim, S.-H. & Brünger, A. T. Model bias in macromolecular crystal structures. *Acta Crystallogr. A.* **48**, 851–858 (1992).
50. Dolinsky, T. J. et al. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **35**, W522–W525 (2007).
51. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA* **98**, 10037–10041 (2001).
52. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
53. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).
54. Grzesiek, S. & Bax, A. The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc.* **115**, 12593–12594 (1993).
55. Vinci, C. R. & Clarke, S. G. Recognition of age-damaged (R,S)-adenosyl-L-methionine by two methyltransferases in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* **282**, 8604–8612 (2007).
56. Shah, S. & Friedman, S. H. An ESI-MS method for characterization of native and modified oligonucleotides used for RNA interference and other biological applications. *Nat. Protoc.* **3**, 351–356 (2008).
57. Jain, K., Jin, C. Y. & Clarke, S. G. Epigenetic control via allosteric regulation of mammalian protein arginine methyltransferases. *Proc. Natl Acad. Sci. USA* **114**, 10101–10106 (2017).
58. Smietanski, M. et al. Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. *Nat. Commun.* **5**, 3004 (2014).
59. Papadopoulos, J. S. & Agarwala, R. COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* **23**, 1073–1079 (2007).
60. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a | Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

Crystallographic data were acquired with software associated with each beamline. NMR data were acquired with TopSpin from Bruker. Mass spec data were collected with ThermoFisher Exactive Series Instrument Control Software. ITC data were collected using MicroCal software provided with the instrument.

Data analysis

Phenix v.1.13.2998, Coot v.0.8.9.1 EL, XDS, PyMOL v.2.0.6, PDB2PQR v.2.0.0, APBS, NMRPipe v.2009.015.15.35, NMRFAM-Sparky v.1.2, Igor Pro v.8.00

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data availability statement is included in the manuscript.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	A sample size of three independent experiments are chosen for comparing all relative enzymatic activities. Two to six independent experiments with good agreement within each group are performed for binding experiments. No sample size calculation was performed. The sample sizes were chosen by commonly used sizes for these type of experiments.
Data exclusions	No data were excluded from the analyses.
Replication	All replicates were performed in independent measurements and were successful. Additional control kinetic and ITC experiments with wild-type protein and RNA from three or more batches of protein purified show good reproducibility.
Randomization	The order of protein substitutes used is randomized and shows no correlation between order of usage and resulting relative activities.
Blinding	Not applicable to the study as all biochemical reagents are clearly labeled.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging