Published in final edited form as:

Wiley Interdiscip Rev RNA. 2010 ; 1(1): 152–172. doi:10.1002/wrna.19.

Enzymology of RNA cap synthesis

Agnidipta Ghosh and Christopher D. Lima*

Structural Biology Program, Sloan-Kettering Institute, New York, NY 10065, USA

Abstract

The 5' guanine-N7 methyl cap is unique to cellular and viral messenger RNA (mRNA) and is the first co-transcriptional modification of mRNA. The mRNA cap plays a pivotal role in mRNA biogenesis and stability, and is essential for efficient splicing, mRNA export, and translation. Capping occurs by a series of three enzymatic reactions that results in formation of N7-methyl guanosine linked through a 5'-5' inverted triphosphate bridge to the first nucleotide of a nascent transcript. Capping of cellular mRNA occurs co-transcriptionally and *in vivo* requires that the capping apparatus be physically associated with the RNA polymerase II elongation complex. Certain capped mRNAs undergo further methylation to generate distinct cap structures. Although mRNA capping is conserved among viruses and eukaryotes, some viruses have adopted strategies for capping mRNA that are distinct from the cellular mRNA capping pathway.

Cap formation is the first co-transcriptional 5' end modification of nascent messenger RNA (mRNA).¹ The 5' cap is a characteristic signature of eukaryal and viral mRNAs, and is absent in bacterial and archaeal transcripts. The 5' cap is also a defining feature of transcripts that are produced by RNA polymerase II. Eukaryal and viral mRNA are capped by N7-methyl guanosine (m7G) that is linked through an inverted 5'-5' triphosphate bridge to the initiating nucleoside of a nascent transcript (reviewed in Refs 2–4). A large body of evidence suggests that the 5' cap structure is necessary for efficient pre-mRNA splicing, export, stability, and translation.^{5–7}

Capping by m7G occurs in sequential steps and is catalyzed by three enzymatic activities: (1) hydrolysis of 5' γ -phosphate of a nascent pre-mRNA, (2) transfer of a guanine monophosphate nucleoside to the 5' diphosphate mRNA end, and (3) methylation of the guanine N7 position. Early research on viral capping in *reovirus* and *vaccinia* virus defined the first cap structures and the enzymes required for cap formation.^{1,8,9} Subsequently, cellular pathways for m7G cap formation were identified and characterized in several eukaryotic organisms (reviewed in Ref 10). While the steps required for cap formation are conserved among viruses and eukaryotes, significant variations exist with respect to the genetic organization of capping enzymes as well as the structure and catalytic mechanism of cap-forming enzymes in different taxonomical groups.³

Although the mRNA cap generally refers to a 5' m7G linked to the first nucleoside of a nascent transcript, many other distinct caps exist throughout biology. In many instances, m7G capped mRNA can be further modified at nucleoside 2'-OH positions to generate hypermethylated forms of the cap. Furthermore, some small nuclear and nucleolar RNAs required for pre-mRNA splicing (e.g., U1, U2, U4, and U5), pre-rRNA processing (U3 and U8), and telomere addition (telomerase RNA) are capped by a trimethylguanosine (TMG) cap, while a handful of other small RNAs (e.g., mammalian U6 and 7SK, mouse B2, and

plant U3) are capped by a γ -methylphosphate at the 5' end of the transcript. In this review, we discuss the enzymology of 5' cap synthesis with an emphasis on different strategies, both structural and mechanistic, utilized in cellular and viral capping reactions.

CAP STRUCTURES

The 5' guanine-N7 cap (m7G) or cap 0 is formed in three sequential reactions catalyzed by enzymes called RNA triphosphatase, RNA guanylyltransferase, and RNA guanine-N7 methyltransferase²⁻⁴ (Figure 1). Cap 0 can undergo further methylation on the ribose 2'-hydroxyl (2'-O) of the first nucleoside in a reaction catalyzed by RNA 2'-O-ribose methyltransferase to produce the cap 1 structure. Cap1 structures that contain a 5' terminal 2'-O methylated adenosine can undergo further methylation at the N6 position.¹¹ Further methylation of cap 1 at the ribose 2'-O position of the second nucleoside by similar enzymes results in formation of the cap 2 structure (Figure 1). Although the cap 0 and cap 1 modification of a nascent mRNA occurs in the nucleus, cap 2 modification occurs in the cytoplasm. Additional methylation of the cap can occur, and in trypanosomatids (*Leishmania* and *Trypanosoma*) the first four nucleosides of the nascent transcript undergo hypermethylation to generate the cap 4 structure.¹² In cap 4, the first nucleoside (adenine) of the transcript undergoes further methylation at the N6 and ribose 2'-O positions in reactions catalyzed by an as yet uncharacterized enzyme and a ribose 2'-O specific methyltransferase (TbMTr1), respectively, to generate 6,6,2 trimethyl adenine. The following three nucleosides (uracils) are subsequently methylated by two additional ribose 2'-O specific methyltransferases (TbMTr2 and TbMTr3) at ribose 2'-O positions to complete cap 4 synthesis¹³ (Figure 1).

Certain small nuclear, nucleolar, and telomerase RNAs are capped by trimethylguanosine or TMG cap that is derived from the cap 0 structure. In TMG caps, cap 0 is dimethylated at the N2 guanine position in a reaction catalyzed by cap TMG synthase to form a 2,2,7 trimethylguanosine cap or TMG cap^{14,15} (Figure 1). Trimethylguanosine synthase 1 (TGS1) has been identified as the *bona fide* dimethyltransferase for TMG cap synthesis.¹⁵ A subset of small RNAs, including mammalian U6 and 7SK, mouse B2 and plant U3, is capped with a distinct γ -methyl phosphate cap. The γ -methyl cap is catalyzed by γ -methyltransferase, which transfers a methyl group from S-adenosylmethionine (AdoMet) to a γ -phosphate oxygen at the unprocessed 5' end of the primary transcript^{16,17} (Figure 1).

WHY HAVE CAPS?

Cap modification is indispensable for growth of eukaryotic cells because it plays essential roles in regulating mRNA stability, mRNA maturation and export, and translation initiation (reviewed in Ref 18). The cap contributes to mRNA stability by physically protecting the mRNA from 5' $\rightarrow$ 3' exoribonucleases. Importantly, the cap must be removed by enzymes such as Dcp2 to generate a 5' monophosphate terminated RNA, which is a substrate for 5' $\rightarrow$ 3' exoribonucleases such as Xrn1/Rat1 (reviewed in Ref 19).

The 5' cap structure is required for efficient mRNA splicing. Splicing is mediated by a nuclear cap-binding complex that specifically interacts with the m7G cap to promote efficient interactions between snRNPs and 5' splice sites.²⁰ Interaction between the nuclear cap-binding complex and nascent mRNA is presumed to be one of the earliest steps in spliceosome assembly in which mRNA is packaged into a ribonucleo-protein particle (mRNP) for export to the cytoplasm.²¹ The nuclear cap-binding complex is essential for the export of spliceosomal components (U snRNAs) from the nucleus.²²

The 5' cap structure is required for translation²³ because it serves as a signal to direct the translational machinery to the 5' end of the protein coding mRNAs²⁴ and because the rate-

limiting step in establishing translational reading frame is recruitment of the 40S ribosomal subunit to the 5' end of mRNA. Eukaryotic translation initiation factor 4F (eIF4F) is a key component of the translational apparatus that directly binds the 5' m7G cap structure.¹⁸

ENZYMOLOGY AND THE CLASSICAL PATHWAY OF CAP SYNTHESIS

The three reactions required for m7G cap or cap 0 synthesis are catalyzed sequentially by the enzymatic activities of RNA triphosphatase, RNA guanylyltransferase and RNA guanine-N7 methyltransferase (Figures 1 and 2). RNA triphosphatases are the first enzymes to act on nascent pre-mRNAs ($p_{\gamma}p_{\beta}p_{\alpha}$ RNA) in reactions that remove the 5' γ -phosphate from the nascent pre-mRNA to generate a 5' diphosphate end ($p_{\beta}p_{\alpha}$ RNA). In the second reaction, the RNA guanylyltransferases cap the 5' diphosphate mRNA with GMP to form 5' guanylated-RNAs ($Gp_{\alpha}p_{\beta}p_{\alpha}$ RNA). In the third reaction, RNA guanine-N7 methyltransferases transfer a methyl group from the *S*-adenosylmethionine (AdoMet) methyl donor to the N7 position of the terminal guanine base ($m7Gp_{\alpha}p_{\beta}p_{\alpha}$ RNA) to complete the synthesis of cap 0.

RNA Triphosphatases

RNA triphosphatases differ with respect to structure and catalytic mechanism across evolution and can be grouped into two distinct families: (1) metal-dependent RNA triphosphatases of lower eukaryotes such as fungi and protozoa, and (2) metal-independent RNA triphosphatases of nematodes, metazoa, and plants (reviewed in Refs 25,26).

The metal-dependent RNA triphosphatase family is exemplified by the RNA triphosphatase Cet1 from budding yeast. This enzyme hydrolyzes the terminal phosphate from pppRNA.²⁷ The crystal structure of Cet1 revealed a homodimeric architecture with two equivalent active sites, each located in the center of a topologically closed eight-stranded anti-parallel β barrel (the triphosphate tunnel) (Figures 2 and 3(a)) that is the signature for 'triphosphate tunnel metalloenzyme' (TTM) superfamily.^{28,29} The Cet1 active site is composed of several basic residues that are presumed to coordinate the triphosphate moiety of pppRNA and several acidic residues that coordinate two metal ions that in turn interact with the pppRNA triphosphate moiety.^{30–32} In the Cet1 structure, a sulfate ion was observed in the active site that was coordinated by several basic residues and a single manganese ion that was in turn coordinated by several acidic residues²⁹ (Figure 3(a)). The position of the sulfate ion was proposed to reflect the position of the γ -phosphate of pppRNA. Over the past decade, other Cet1-like enzymes have been identified and characterized in protozoa.^{33,34} These triphosphatases were also shown to possess metal-dependent phosphohydrolase activity, and their primary sequences show similarities within the eight β strands that compose the Cet1 triphosphate tunnel. Structure-function analyses have also revealed that most DNA virus-encoded RNA triphosphatases are metal-dependent phosphohydrolases that share biochemical and structural similarities with Cet1³⁵ (Figure 3(b)).

RNA triphosphatases of metazoan and plant capping enzymes are metal-independent phosphohydrolases that contain a canonical HCxxxxR(S/T) motif in their phosphate binding or P-loop, a signature motif of the cysteine phosphatase superfamily.³⁶ Like their protein phosphatase cousins, these RNA triphosphatases catalyze metal-independent phosphoryl-transfer in a two-step reaction. In step 1, the conserved cysteine of the P-loop attacks the 5' γ -phosphate of the nascent transcript to form a covalent protein-cysteinyll-*S*-phosphate intermediate and release the 5' ppRNA product (Figure 2). In step 2, the covalent phosphoenzyme is hydrolyzed to liberate inorganic phosphate. The structure of a metal-independent RNA triphosphatase family member was first revealed by analysis of the RNA triphosphatase domain of mouse capping enzyme Mce1³⁷ (Figure 3(c)) whose active site

was located in a deep and positively charged pocket that appeared to be capable of accepting 5' pppRNA but not 5' ppRNA as a substrate.

RNA Guanylyltransferases

Cellular and viral RNA guanylyltransferases transfer GMP to the diphosphate end of a nascent transcript in a reversible two-step ping-pong reaction (Figure 2) that involves (1) binding of GTP and magnesium to catalyze formation of a covalent enzyme-(lysyl-*N*)-GMP intermediate and release of pyrophosphate, and (2) binding of the ppRNA substrate to catalyze transfer of the GMP to the diphosphate end (reviewed in Ref 38). Structural studies highlighted that RNA guanylyltransferases have a bipartite domain architecture that includes an *N*-terminal nucleotidyltransferase (NTase) domain and a C-terminal OB-fold domain^{39,40} (Figure 4). Primary sequence similarity and structural studies showed that RNA guanylyltransferases share similarities with adenosine triphosphate (ATP)-dependent and nicotinic adenine dinucleotide (NAD)⁺-dependent DNA and RNA ligases, enzymes that comprise a branch of the covalent nucleotidyltransferase superfamily.^{25,38} The NTase domain includes a nucleotide-binding pocket that is formed by five signature sequence motifs (I, III, IIIa, IV, and V) that define the polynucleotide ligase/RNA capping enzyme superfamily.⁴¹ Motif I (KxDG) contains the lysine nucleophile that forms the covalent (lysyl)-*N*-GMP intermediate (Figure 4).

Crystal structures of the *Chlorella* virus RNA guanylyltransferase PbCV-1 were determined in two configurations. In the 'open' configuration, the enzyme was observed bound to GTP, whereas in the 'closed' configuration the enzyme was observed as an enzyme-guanylate intermediate⁴⁰ (Figure 4(a)). In later studies, the first structure of a cellular RNA guanylyltransferase (Cgt1) was from *Candida albicans*. In this work, the enzyme was observed in two configurations, one 'open' GTP-bound configuration and one 'open' enzyme-guanylate configuration (Figure 4(b)).³⁹ Based on these structures, and structural studies of other members of the covalent nucleotidyltransferase superfamily, a general mechanism for phosphoryl-transfer has emerged.^{25,38,42} Magnesium and GTP bind to the 'open' form of RNA guanylyltransferase (Figure 4(a) and 4(b)) inducing domain closure which facilitates interactions between GTP and residues within the OB-fold domain that are required for formation of the enzyme-(lysyl-*N*)-GMP adduct (Figure 4(b)). Cleavage of GTP releases interactions between the OB-fold domain and GTP β - γ phosphates allowing the enzyme to 'open' to release pyrophosphate. In most instances, structural data suggest that the covalent nucleotide adduct undergoes a conformational change from *syn* to *anti* to properly position the nucleotide for attack by the incoming diphosphate-terminated RNA substrate (Figure 2). The enzyme must undergo another round of closure to catalyze GMP transfer to the ppRNA substrate and opening to release the capped RNA.³⁸

RNA Guanine-N7 Methyltransferases

Cellular and viral RNA guanine-N7 methyltransferases are *S*-adenosylmethionine (AdoMet) dependent enzymes that catalyze methyl transfer from the AdoMet methyl donor to the N7 guanine acceptor to complete cap 0 synthesis (m7GpppRNA) (Figure 2). Crystal structures of a cellular N7-guanine methyltransferase from the intracellular parasite *Encephalitozoon cuniculi*⁴³ were determined alone or as individual structures in complex with the methyl donor AdoMet, the product AdoHcy, or a cap analog m7GpppG (Figure 5(b)). These studies highlighted that guanine-N7 methyltransferases have two ligand binding sites located in a deep inter-segment cleft, one for the methyl donor AdoMet and the other for the guanosine methyl acceptor. Superposition of these structures and ligand complexes suggested an in-line mechanism of methyl transfer, albeit without direct contacts between the enzyme and either the N7 atom of guanine (the attacking nucleophile), the methyl carbon of AdoMet, or the sulfur of AdoMet/AdoHcy (the leaving group). The structures suggest that catalysis of

cap N7 methylation is accomplished by optimizing proximity and orientation of the substrates, assisted by a favorable electrostatic environment.⁴³ Subsequent structural studies combined with mutational analysis of Ecm1 and the budding yeast guanine-N7 methyltransferase Abd1 suggest these mechanisms are conserved across evolution.^{43,44} Also of note is the recent unpublished crystal structure of human guanine-N7 methyltransferase Hcm1 (PDB code 3EPP) that is structurally similar to Ecm1 (Figure 5(c)).

Earlier studies on the *vaccinia* guanine-N7 methyltransferase revealed that it was a heterodimeric protein consisting of a catalytic subunit (D1) and a regulatory subunit (D12). Although the D1 subunit displayed low intrinsic activity alone, association with the D12 subunit was required for full activity.^{45–49} This phenomenon was explained in part by a recent crystal structure of the D1 subunit in complex with the regulatory D12 subunit⁵⁰ that showed that D12 binds to a D1 surface distal from its active sites suggesting that it stabilizes D1 through an allosteric mechanism (Figure 5(a)), a result confirmed by a subsequent mutational analysis.⁵¹ These studies also highlighted a role for the nine *N*-terminal amino acids of the D1 subunit because these residues fold over the substrate binding pockets⁵⁰ (Figure 5(a)).

HIGHER-ORDER ARCHITECTURE OF THE CAPPING APPARATUS

While activities required for cap 0 synthesis are conserved, the way in which the enzymes are organized differs across eukaryal and viral evolution. In fungi, the three enzymatic activities are encoded by separate genes, one for the metal-dependent triphosphatase, one for the guanylyltransferase, and one for the guanine-N7 methyltransferase (Figure 6).^{52,53} In metazoa and plants, capping activities are encoded in two polypeptides, one for the guanine-N7 methyltransferase and one that contains a bifunctional enzyme with an *N*-terminal metal-independent RNA triphosphatase domain fused to a C-terminal RNA guanylyltransferase domain^{54,55} (Figure 6).

Fungi

In addition to variations in genetic organization as noted earlier, differences also exist with respect to interactions between capping enzymes. In budding and fission yeast, the RNA triphosphatase is a homodimer. The *Saccharomyces cerevisiae* Cet1 structure revealed that an extended conformation of an *N*-terminal Cet1 polypeptide stretched across the Cet1 dimer interface to form contacts to both Cet1 protomers²⁹ (Figure 3(a)). The functional importance of these contacts was highlighted by a Cet1 construct that contained an *N*-terminal deletion up to amino acid 275, which resulted in a catalytically active, albeit monomeric form of Cet1 that could not complement the essential functions of Cet1 *in vivo* unless fused covalently to a guanylyltransferase.⁵⁶

Interactions between the triphosphatase and guanylyltransferase are also important for function in budding yeast.⁵⁷ Sedimentation analysis suggested that Cet1 and Ceg1 interacted in a 2:1 complex^{56–58} and interactions between Cet1 and Ceg1 were dependent on a 34 amino acid segment of Cet1 that contained a conserved WAQKW motif^{56,59,60} and the OB domain of Ceg1.^{39,61} Structure/function analysis suggested this was also true for the *Candida albicans* guanylyltransferase Cgt1 because mutations in its OB domain led to defects in Cet1 binding *in vitro* and to defects in complementing the essential functions of the guanylyltransferase *in vivo*.³⁹ Interactions between Cet1 and Ceg1 likely serve two essential functions *in vivo*: (1) to stimulate or stabilize guanylyltransferase activity,^{57,58} and (2) to recruit Cet1 to the RNA polymerase II elongation complex via interactions with Ceg1 (see below).^{58,62,63}

Recent biochemical and structural studies showed the basis for interactions between Cet1 and Ceg1 and that this complex co-exists *in vitro* as a 2:1 Cet1-Ceg1 heterotrimer and a 2:2 Cet1-Ceg1 heterotetramer⁶⁴ (Figure 7(a)). The crystal structure showed a heterotetrameric complex composed of one Cet1 homodimer and two Ceg1 molecules that associate via interactions between the Ceg1 OB domain and an extended Cet1 WAQKW amino acid motif (Figure 7(a)). The Cet1 WAQKW motif and triphosphatase domain motif are linked by a flexible tether, as evidenced by its susceptibility to proteolysis.⁵⁶ The importance of the Cet1 WAQKW motif for interactions with Ceg1 was illustrated in previous mutational studies which showed that it was required for interactions with Ceg1,^{59,60} whereas the importance of amino acid side chains in the Ceg1 OB domain for this interaction were revealed by mutational and functional studies of the guanylyltransferases of *C. albicans*³⁹ and *S. cerevisiae*.⁶⁴ Capping enzymes of *Schizosaccharomyces pombe* share sequence similarity to those of other budding yeasts including elements in the triphosphatase that are known to be important for homodimerization⁶⁵; however, no interactions have been detected between the triphosphatase and guanylyltransferase.⁶⁶

Protozoa

Similar to fungi, several protozoan parasites such as *Encephalitozoan cuniculi*, *Plasmodium falciparum*, and *Trypanosoma* species encode capping machinery in three separate polypeptide chains. In these organisms, the RNA triphosphatase appears similar to the metal-dependent Cet1 triphosphatase.²⁹ In *Leishmania*, capping enzymes are synthesized in two parts with the Cet1-like RNA triphosphatase activity encoded in a single polypeptide, whereas the RNA guanylyltransferase and guanine-N7 methyltransferase activities are encoded in a bifunctional polypeptide²⁶ (Figure 6).

Metazoa and Plants

Metazoans and plants include a bifunctional capping enzyme in which the triphosphatase and guanylyltransferase catalytic domains are contained within a single polypeptide.²⁶ The linker connecting the two domains is probably flexible because it is composed by acidic residues that are predicted to be unstructured. For instance, in the mouse capping enzyme for which the structure of the triphosphatase domain is known, the linker includes 21 amino acids, 70% of which are acidic in nature. A flexible linker between triphosphatase and guanylyltransferase domains is also observed in the *S. cerevisiae* Cet1-Ceg1 complex,⁶⁴ and is predicted to be important for guanylyltransferase function because it would limit steric impediments when the guanylyltransferase undergoes large conformational changes during capping (see the Structures and Mechanisms of the Classical Cap-forming Enzymes section above).

Viruses

Eukaryotic viruses that replicate in the cytoplasm encode their own capping machineries. Similar to fungi, in *Chlorella* virus the triphosphatase, guanylyltransferase, and guanine-N7 methyltransferase are encoded in separate polypeptides.⁴⁰ In contrast, *baculovirus* encodes a bifunctional capping enzyme in which an *N*-terminal Cet1-like metal-dependent RNA triphosphatase domain is fused by a linker to the guanylyltransferase (Figure 6). The viral genome also encodes a cysteine phosphatase type RNA triphosphatase.^{67,68} Similar to *baculovirus*, a member of the *iridovirus* family (Infectious Spleen and Kidney Necrosis virus, ISKNV) encodes a bifunctional capping enzyme. In this virus, the *N*-terminal triphosphatase domain belongs to the family of metal-independent cysteine phosphatases^{69–73} (Figure 6). Interestingly, these viruses do not appear to encode RNA guanine-N7 methyltransferases suggesting that they rely on the host cell for cap methyltransferase activity. *Poxvirus*, *African swine fever virus*, and *mimivirus* each encode

mRNA capping activities in a single trifunctional polypeptide with an *N*-terminal metal-dependent RNA triphosphatase domain, a central RNA guanylyltransferase domain and a C-terminal RNA methyltransferase domain^{26,35} (Figure 6).

Several double-stranded RNA viruses synthesize the cap 1 structure. In some of the *Reoviridae* family members, each of the four capping activities for cap 1 synthesis are contained in a single multifunctional polypeptide chain^{74–76} and the 5' end of transcripts are capped with cap 1 in four steps prior to viral RNA export to the host cell cytoplasm. Structural studies on the *reovirus* core showed that all essential enzymatic activities required for cap 1 synthesis reside in the $\lambda 2$ polypeptide of the viral core⁷⁶ including a guanylyltransferase, and two C-terminal $\lambda 2$ methyltransferase domains with structural similarity to *vaccinia* virus VP39 (see below). Although the *reovirus* guanylyltransferase bears no resemblance to other cellular guanylyltransferases, it does contain a lysine residue that is known to form a lysyl-*N*-GMP adduct. Based on the respective positions of the active sites within the context of the virus core, these domains appeared poised to cap the RNA as it is extruded from the virus.⁷⁶ *Blue-tongue* virus also encodes four enzymatic activities for cap 1 modification in a single polypeptide chain, VP4. Structural studies of VP4 showed a multidomain architecture that includes an *N*-terminal kinase-like domain predicted to be important for protein–protein interactions, central N7-methyltransferase and 2'-O methyltransferase domains, and a C-terminal domain that was proposed to contain both triphosphatase and guanylyltransferase activities despite its lack of structural similarity to any characterized RNA triphosphatase or guanylyltransferase enzymes⁷⁷ (Figure 7(b)).

COORDINATING CAP FORMATION WITH RNA SYNTHESIS

In eukaryotes, capping occurs co-transcriptionally and is facilitated by direct recruitment of the capping enzymes to the transcription apparatus where it modifies the nascent mRNA when it reaches a length of approximately 30 nucleotides.^{62,63,78,79} In a conceptually similar mechanism, *vaccinia* virus mRNA capping enzyme interacts directly with the *vaccinia* RNA polymerase to facilitate co-transcriptional capping of the nascent mRNA when it reaches a length of at least 31 nucleotides.^{80,81} In eukaryotes, capping is facilitated by direct recruitment of the capping apparatus to RNAP II via interactions with the phosphorylated C-terminal domain (CTD) of the largest RNAP II subunit Rpb1 (reviewed in Ref 82). The CTD is composed of tandem heptapeptide repeats with the consensus sequence Y₁S₂P₃T₄S₅P₆S₇.^{83–85} The CTD undergoes successive rounds of phosphorylation and dephosphorylation at serine residues 2, 5, and 7 in coordination with the transcriptional cycle. The CTD is essential,^{63,86,87} and serves as a landing platform for factors that regulate initiation, elongation, and termination steps of RNAP II transcription.⁸² During initiation and early elongation, the CTD is enriched in Ser5 and Ser7 phosphorylation and as elongation proceeds, Ser5 and Ser7 phosphorylation decreases, whereas Ser2 phosphorylation increases and persists until RNAP II reaches the 3' end of the transcription unit. Although Ser7 phosphorylation occurs early in the transcription cycle, this position in the CTD heptad array is the most degenerate and is not required for viability in yeast or mammalian cells.^{88,89}

Recruitment of capping activities to RNAP II requires direct interactions between components of the capping apparatus and phosphorylated RNAP II CTD^{39,52,63,90–93} (Figure 8). Although recruitment of the capping apparatus to RNAP II is conserved across evolution, a wide array of mechanisms has evolved to achieve this end. In many organisms examined thus far, the guanylyltransferase and methyltransferase interact directly with the phosphorylated RNAP II CTD. In *S. cerevisiae* and *C. albicans*, interactions between the triphosphatase and guanylyltransferase OB domain (see above) are required for cell viability and presumably for recruitment of the triphosphatase to RNAP II under normal expression

conditions.^{39,58,60,61} However, alternative mechanisms exist for recruitment of Cet1 to the transcription apparatus because a form of *S. cerevisiae* Cet1 that no longer interacted with Ceg1 was still observed at 5' ends of genes⁶⁰ and because the *C. albicans* triphosphatase was shown to directly interact with the phosphorylated RNAP II CTD.⁹⁴

In fission yeast, the *S. pombe* triphosphatase Pct1 and guanylyltransferase Pce1 do not interact with each other and both bind phosphorylated RNAP II CTD independently^{66,94} (Figure 8). *In vitro* analyses showed that both enzymes preferred to bind a CTD substrate that was phosphorylated at Ser5 or doubly phosphorylated at Ser2 and Ser5. In mammals and plants, the triphosphatase is fused to the guanylyltransferase, thus the triphosphatase is recruited to RNAP II through its covalent association with the guanylyltransferase, which interacts directly with the phosphorylated RNAP II CTD^{35,54,93} (Figure 8). Although the Mce1 mouse guanylyltransferase domain binds either Ser2 or Ser5 phosphorylated CTD, its guanylyltransferase activity is only stimulated by binding to CTD that is phosphorylated at Ser5.⁹⁰

The structural basis for recognition of the phosphorylated RNAP II CTD by the guanylyltransferase was exemplified by a complex between *C. albicans* guanylyltransferase Cgt1 and a CTD peptide that contained phosphorylated Ser5 residues (Figure 7(c)).³⁹ The structure showed that the CTD adopted an extended conformation in interactions that were localized to the *N*-terminal Cgt1 NTase domain. It is of some interest that Cgt1 residues involved in CTD binding are not strictly conserved among guanylyltransferase family members, suggesting that other guanylyltransferases may recognize the phosphorylated CTD in a different manner.

EVOLUTION OF THE CAPPING APPARATUS

Cap modification is unique to cellular and viral mRNAs and the emergence of the cap in eukaryotes and viruses coincides with (1) the loss of prokaryotic Shine-Dalgarno sequence that directs ribosomes to mRNA, (2) the appearance of 5' exoribonucleases in eukaryotes, and (3) the emergence of three eukaryotic RNA polymerases with distinct functionalities.

Capping enzymes have apparently evolved from ancestral protein families. Metal-dependent RNA triphosphatases of lower eukaryotes (protozoa and fungi) share biochemical and structural similarities with viral RNA triphosphatases.^{29,35} These enzymes form a large enzyme superfamily, now known as 'triphosphatase tunnel metalloenzymes' (TTM)²⁸ that include TTM family members in archaea (*Pyro-coccus furiosus* protein of unknown biochemical function), bacteria (*Vibrio parahaemolyticus* CyaB-like adenylated cyclase), and mammalian thiamine triphosphatase.⁹⁵ Recently, another Cet1-like tunnel fold was exemplified for a membrane associated polyphosphate polymerase, VTC.⁹⁶ Therefore, the tunnel fold first described for Cet1²⁹ is the prototype for a large enzyme family encompassing three domains of life, with functions not limited to 5' RNA processing. In contrast, the metal-independent RNA triphosphatases of metazoa and plants are restricted to eukaryotes and belong to the cysteine phosphatase superfamily, which includes protein tyrosine phosphatases, dual specificity phosphatases, and phosphoinositide phosphatases.³⁶ It is conceivable that metal-independent RNA triphosphatases evolved from a common ancestor, altering various sequence and structural elements to engage the pre-mRNA substrate.

RNA guanylyltransferases are structurally and mechanistically related to ATP and NAD⁺-dependent DNA ligases and ATP-dependent RNA ligases, proteins that comprise a superfamily of covalent nucleotidyltransferases that involves formation of a covalent enzyme-(lysyl-*N*)-nucleoside monophosphate adduct.^{25,38} Crystal structures of DNA

ligases, RNA ligases, and RNA guanylyltransferase highlight a common architecture with respect to the active site and nucleotidyltransferase domain and account for determinants that dictate nucleoside specificities. Although ATP-dependent DNA and RNA ligases are found in all three kingdoms of life, RNA guanylyltransferases are currently limited to eukarya and viruses that replicate in the cytoplasm. Thus, RNA guanylyltransferases apparently evolved from an ancestral ligase during the emergence of eukarya, and through mutations acquired specificity for GTP and single-stranded ppRNA.

The guanine-N7 methyltransferase contains structural and sequence motifs characteristic of the class I methyltransferase family, a large family of AdoMet-dependent methyltransferases that are found in all three kingdoms of life.⁹⁷ Because of the widespread distribution of these enzymes, it is likely that RNA guanine-N7 methyltransferases evolved from a common ancestor through acquisition of mutations that generated specificity for the GpppRNA substrate.

UNCONVENTIONAL PATHWAYS OF CAP FORMATION

Capping with GDP

In *rabdoviruses*, RNA-dependent RNA polymerase L protein (exemplified in *vesicular stomatitis virus*) catalyzes mRNA capping and methylation, but it differs mechanistically from cellular capping enzymes.⁹⁸ *Rhabdovirus* L protein utilizes a 5'-triphosphate-ended oligo-RNA as a substrate to form a covalent enzyme-RNA complex that is achieved through a phosphoamide bond between the 5' α -phosphate of the viral RNA and a histidine in the L protein⁹⁹ (Figure 9(a)). The L protein of VSV is also a GTPase *in vitro*, the product of which, namely, GDP, acts as a nucleophile to generate a 5' capped GpppRNA.

Cap Snatching

Synthesis of mRNA in *influenza virus* is initiated by stealing cellular caps¹⁰⁰ (Figure 9(a)), and the viral RNA-dependent RNA polymerase is responsible for this unique 'cap snatching' mechanism. The viral polymerase is a heterotrimer composed of three subunits: PA, PB1 and PB2. Although polymerase activity is localized to the PB1 subunit,¹⁰¹ the PB2 subunit recognizes and binds cellular 5' capped transcripts in a manner that is stimulated by interactions between the 5' end of the viral RNA template and PB1.¹⁰² A series of recent crystal structures revealed that the interface between PA1 and PB2 is important for viral polymerase function¹⁰³ and identified the cap-binding site in PB2.¹⁰⁴ Upon binding capped cellular transcripts, the PA N-terminal domain catalyzes endonucleolytic cleavage between nucleotides 10 and 13 to generate a short capped RNA oligoribonucleotide^{100,105–107}. The capped oligoribonucleotide products of this reaction are then used to prime synthesis of viral mRNAs, which are polyadenylated at their 3' ends by the polymerase.

Capping with m⁷GMP

In *alphavirus*, the nonstructural protein nsP1 is responsible for guanylation of nascent transcripts. nsP1 also catalyzes AdoMet-dependent methyltransfer to the N7 position of the guanine cap¹⁰⁸ (Figure 9(c)). Biochemical studies showed that nsP1 forms a covalent enzyme-GMP intermediate, however unlike cellular guanylyltransferases, nsP1 was covalently associated with N7-methyl GMP¹⁰⁸ suggesting that *alphavirus* nsP1 employs a unique mRNA capping activity where methylation of GMP occurs prior to formation of 5'-5' triphosphate bridge between GMP and the nascent transcript.

CAP-SPECIFIC RIBOSE 2'-O METHYLTRANSFERASES

DNA Viruses

Vaccinia virus VP39 protein participates in maturation of both 5' and 3' ends of viral mRNA.^{109,110} VP39 catalyzes ribose 2'-O methyltransferase activity of cap 0,¹¹¹ and later biochemical studies suggested that two distinct VP39 surfaces were responsible for recognition of cap 0 and the trailing RNA.¹¹² A suite of structural studies showed the architecture of VP39 and determinants required for cap 0 recognition, AdoMet binding and 2'-O methylation,^{113–116} which included (1) specific recognition of the m7G portion of cap 0 via stacking interactions with aromatic residues and hydrogen bonding interactions with the base, (2) nonsequence specific interactions with the trailing RNA chain, and (3) positioning of the AdoMet donor and activation of the 2'-OH as a methyl acceptor by suppressing its pKa through proximity to a positively charged amino acid side chain and the AdoMet donor.

ssRNA Viruses

The genus *flavivirus* includes Dengue and West Nile viruses. The *flavivirus* genome is a single-stranded RNA that is capped at its 5' end by cap 1.¹¹⁷ This RNA encodes a polyprotein that is processed by viral and cellular proteases into three structural proteins and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5). The NS5 protein encodes RNA-dependent RNA polymerase activity and its N-terminal domain harbors AdoMet-dependent N7-methyltransferase and 2'-O methyltransferase activities, both of which are necessary for synthesis of the RNA cap 1 structure.^{118,119} Methylation of N7-guanine precedes ribose 2'-O methylation.¹²⁰ The domain contains a single AdoMet binding site, exemplified by a crystal structure,¹²¹ and, thus, the dual methyltransferase activity apparently requires repositioning of the substrate for the two successive methyl transfer reactions. Similar to *flaviviruses*, *rhabdoviruses* synthesize RNA cap1 structure, and in *vesicular stomatitis* virus (VSV) two methyltransferase activities share a single AdoMet binding site; however, in this virus 2'-O methyltransferase activity precedes and stimulates guanine-N7 methyltransferase activity.^{2,122}

dsRNA Viruses

Double-stranded RNA viruses such as *reovirus* and *bluetongue* virus belong to the *reoviridae* family. Structural studies of VP4 showed a multidomain architecture that included two methyltransferase domains, one that shares similarity to cellular guanine-N7 methyltransferases and one that shares similarity to the *vaccinia* virus VP39 2'-O methyltransferase⁷⁷ (Figure 7(b)). Structural studies on the *reovirus* core⁷⁶ revealed two C-terminal λ 2 methyltransferase domains; however in this case, both methyltransferase domains shared structural similarity to VP39 2'-O methyltransferase suggesting that one of these putative 2'-O methyltransferases is capable of catalyzing guanine-N7 methyltransferase activity.

Trypanosomes: Cap 4 Synthesis

Family members of *Kinetoplastidae* include *Trypanosoma* and *Leishmania* species that synthesize cap 4, a hypermethylated form of the cap in spliced leader (SL) sequences¹² (Figure 1). Cap 4 in *Kinetoplastidae* is critical for pre-mRNA trans-splicing and translational efficiency.^{123–125} In *T. brucei*, the genes encoding AdoMet-dependent ribose 2'-O methyltransferases have been identified.¹²⁶ The 2'-O methyltransferase TbMtr1 methylates the first transcribed nucleoside, TbMtr2 is responsible for methylation of the second nucleoside, while TbMtr3 has been implicated in methylation of nucleosides at positions 3

and 4. Each of these methyltransferases share sequence similarity to *vaccinia* VP39 and are presumed to function in an analogous manner.

TRIMETHYLGUANINE SYNTHASE

The m7G cap of small nuclear and small nucleolar RNAs are hypermethylated by the enzyme trimethylguanine synthase 1 (TGS1) in *S. cerevisiae*.^{15,127} TGS1 is conserved in eukaryotes and catalyzes two successive AdoMet-dependent methyl transfer reactions to the N2 position of the m7G cap to generate the trimethyl cap (TMG cap)¹⁴ (Figure 1). TGS1 is nonessential in *S. cerevisiae*, but *tsg1* deletion elicits a cold sensitivity growth phenotype and defects in trimethylguanosine cap formation on certain snRNAs and snoRNAs.¹⁵ TGS1 has also been reported to hypermethylate N7-guanosine cap of telomerase RNA TLC1¹²⁸ and consistent with this hypothesis, *tsg1* deletion affects telomere length and results in premature aging. A recent structural study of human TGS1 showed that its catalytic core has a bipartite domain organization that shares similarity to other tRNA or rRNA methyltransferases that modify exocyclic positions on the nucleoside base.^{129,130} In the latter study, TGS1 was bound to cap 0 and AdoHcy. The cap 0 exocyclic N2 position was pointing toward the AdoHcy binding site suggesting that it would be in a suitable configuration to accept methyl groups if AdoMet were present.

GAMMA-METHYLTRANSFERASE

Gamma-methyltransferases catalyze transfer of methyl group from AdoMet to a γ -phosphate oxygen at the unprocessed end of some small nuclear RNAs (mammalian U6, 7SK, mouse B2 and plant U3,^{16,17} Figure 1). γ -methyltransferases are unable to methylate 5'-diphosphate end of a transcript. Thus, γ -methylation ensures that any transcripts that have been processed by RNA triphosphatase are not protected by β -methylation. More recent studies identified the polypeptide BCDIN3 as the γ -methyltransferase for 7SK snRNA.¹³¹ Although it is clear that BCDIN3 contains an AdoMet binding domain and that its activities are responsible for γ -methylation of 7SK sRNA, the precise mechanism of methyl transfer to the 5' triphosphate terminated RNA remains unknown. Furthermore, it remains unclear whether BCDIN3 is responsible for all γ -methylation or if other enzymes with this activity await discovery.¹³²

CONCLUSION

Since discovery of RNA caps more than 30 years ago, much effort has been devoted to understanding the enzymes and mechanisms that underlie its formation. Much progress has been made in the past decade to unravel the structural and mechanistic basis for individual capping reactions, especially with respect to the cellular capping apparatus. Although much insight has been gained by these studies, the precise manner by which the nascent mRNA is recognized remains unclear in most instances because structural studies have relied on biochemical and mutational studies, or complexes with cap analogs in an attempt to infer capping enzyme interactions with the pre-mRNA substrate.

Cap structures originating from viral or cellular enzymes are often similar, although the physical organization of genes, subunit composition, structure, and catalytic mechanisms differ.^{133,134} Although most viral, protozoan, and fungal RNA triphosphatases belong to the 'triphosphate tunnel metalloenzyme' (TTM) superfamily, metazoan RNA triphosphatases belong to a unique clade of metal-independent cysteine phosphatases. Therefore, mechanism-based inhibitors of the TTM family should be selective against viruses, parasitic fungi, and protozoa. Cap methylation also may afford another opportunity to develop anti-infective compounds because: (1) high cellular AdoHcy levels are capable of inhibiting viral replication,^{135,136} (2) the AdoMet analog and natural antibiotic, sinefungin, is a cap

methylation inhibitor that inhibits proliferation of viruses, fungi and parasitic protozoans,^{44,134} and (3) sinefungin selectively inhibits fungal cap methyltransferases in comparison with the human enzyme.¹³⁷ In contrast to the canonical cellular capping pathway, many RNA viruses evolved distinct capping strategies, which may afford additional opportunities to develop selective anti-viral strategies.

Much work remains to fully understand mechanisms that coordinate capping with transcription. Although capping proceeds in a sequential fashion with one activity preceding the other, precise details have yet to emerge to suggest how RNA progressively transits from one active site to the next. Although this may be a random process facilitated by co-localization of the respective capping activities, it would seem more attractive to present the nascent RNA to the triphosphatase first, to the guanylyltransferase second and to the guanine-N7 methyltransferase last. In a final note, it is clear that the cellular capping apparatus is recruited to early RNAP II elongation complexes in a manner that is facilitated by phosphorylation of the CTD, but we do not fully understand if this interaction is sufficient for directing the capping apparatus to the transcription complex to facilitate capping at the very earliest stages of transcription. Furthermore, it remains unclear how the capping apparatus is regulated by other transcription elongation factors such as Spt5.

References

1. Martin SA, Paoletti E, Moss B. Purification of mRNA guanylyltransferase and mRNA (guanine-7-) methyl-transferase from vaccinia virions. *J Biol Chem.* 1975; 250:9322–9329. [PubMed: 1194286]
2. Banerjee AK. 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiol Rev.* 1980; 44:175–205. [PubMed: 6247631]
3. Shuman S. Structure, mechanism, and evolution of the mRNA capping apparatus. *Prog Nucleic Acid Res Mol Biol.* 2001; 66:1–40. [PubMed: 11051760]
4. Shatkin AJ. Capping of eucaryotic mRNAs. *Cell.* 1976; 9:645–653. [PubMed: 1017010]
5. Fresco LD, Buratowski S. Conditional mutants of the yeast mRNA capping enzyme show that the cap enhances, but is not required for, mRNA splicing. *RNA.* 1996; 2:584–596. [PubMed: 8718687]
6. Schwer B, Mao X, Shuman S. Accelerated mRNA decay in conditional mutants of yeast mRNA capping enzyme. *Nucleic Acids Res.* 1998; 26:2050–2057. [PubMed: 9547258]
7. Schwer B, Shuman S. Conditional inactivation of mRNA capping enzyme affects yeast pre-mRNA splicing in vivo. *RNA.* 1996; 2:574–583. [PubMed: 8718686]
8. Furuichi Y, Muthukrishnan S, Tomasz J, Shatkin AJ. Mechanism of formation of reovirus mRNA 5'-terminal blocked and methylated sequence, m7GpppGmpC. *J Biol Chem.* 1976; 251:5043–5053. [PubMed: 821947]
9. Moss B, Gershowitz A, Wei CM, Boone R. Formation of the guanylated and methylated 5'-terminus of vaccinia virus mRNA. *Virology.* 1976; 72:341–351. (also see erratum 75:260). [PubMed: 948865]
10. Shuman S. Capping enzyme in eukaryotic mRNA synthesis. *Prog Nucleic Acid Res Mol Biol.* 1995; 50:101–129. [PubMed: 7754031]
11. Wei C, Gershowitz A, Moss B. N6, O2'-dimethyladenosine a novel methylated ribonucleoside next to the 5' terminal of animal cell and virus mRNAs. *Nature.* 1975; 257:251–253. [PubMed: 1161029]
12. Bangs JD, Crain PF, Hashizume T, McCloskey JA, Boothroyd JC. Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. *J Biol Chem.* 1992; 267:9805–9815. [PubMed: 1349605]
13. Zamudio JR, Mittra B, Campbell DA, Sturm NR. Hypermethylated cap 4 maximizes *Trypanosoma brucei* translation. *Mol Microbiol.* 2009; 72:1100–1110. [PubMed: 19504740]
14. Hausmann S, Shuman S. Specificity and mechanism of RNA cap guanine-N2 methyltransferase (Tgs1). *J Biol Chem.* 2005; 280:4021–4024. [PubMed: 15590684]

15. Mouaikel J, Verheggen C, Bertrand E, Tazi J, Bordonne R. Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol Cell*. 2002; 9:891–901. [PubMed: 11983179]
16. Gupta S, Busch RK, Singh R, Reddy R. Characterization of U6 small nuclear RNA cap-specific antibodies. Identification of gamma-monomethyl-GTP cap structure in 7SK and several other human small RNAs. *J Biol Chem*. 1990; 265:19137–19142. [PubMed: 2229067]
17. Singh R, Reddy R. Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proc Natl Acad Sci U S A*. 1989; 86:8280–8283. [PubMed: 2813391]
18. Varani G. A cap for all occasions. *Structure*. 1997; 5:855–858. [PubMed: 9261078]
19. Liu H, Kiledjian M. Decapping the message: a beginning or an end. *Biochem Soc Trans*. 2006; 34:35–38. [PubMed: 16246173]
20. Lewis JD, Gorlich D, Mattaj IW. A yeast cap binding protein complex (yCBC) acts at an early step in pre-mRNA splicing. *Nucleic Acids Res*. 1996; 24:3332–3336. [PubMed: 8811086]
21. Lewis JD, Izaurralde E, Jarmolowski A, McGuigan C, Mattaj IW. A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes Dev*. 1996; 10:1683–1698. [PubMed: 8682298]
22. Izaurralde E, Lewis J, Gamberi C, Jarmolowski A, McGuigan C, et al. A cap-binding protein complex mediating U snRNA export. *Nature*. 1995; 376:709–712. [PubMed: 7651522]
23. Muthukrishnan S, Both GW, Furuichi Y, Shatkin AJ. 5'-Terminal 7-methylguanosine in eukaryotic mRNA is required for translation. *Nature*. 1975; 255:33–37. [PubMed: 165427]
24. Gingras AC, Raught B, Sonenberg N. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu Rev Biochem*. 1999; 68:913–963. [PubMed: 10872469]
25. Gu M, Lima CD. Processing the message: structural insights into capping and decapping mRNA. *Curr Opin Struct Biol*. 2005; 15:99–106. [PubMed: 15718140]
26. Shuman S. What messenger RNA capping tells us about eukaryotic evolution. *Nat Rev Mol Cell Biol*. 2002; 3:619–625. [PubMed: 12154373]
27. Ho CK, Pei Y, Shuman S. Yeast and viral RNA 5' triphosphatases comprise a new nucleoside triphosphatase family. *J Biol Chem*. 1998; 273:34151–34156. [PubMed: 9852075]
28. Gong C, Smith P, Shuman S. Structure-function analysis of Plasmodium RNA triphosphatase and description of a triphosphate tunnel metalloenzyme superfamily that includes Cet1-like RNA triphosphatases and CYTH proteins. *RNA*. 2006; 12:1468–1474. [PubMed: 16809816]
29. Lima CD, Wang LK, Shuman S. Structure and mechanism of yeast RNA triphosphatase: an essential component of the mRNA capping apparatus. *Cell*. 1999; 99:533–543. [PubMed: 10589681]
30. Bisaillon M, Shuman S. Structure-function analysis of the active site tunnel of yeast RNA triphosphatase. *J Biol Chem*. 2001; 276:17261–17266. [PubMed: 11279161]
31. Bisaillon M, Shuman S. Functional groups required for the stability of yeast RNA triphosphatase in vitro and in vivo. *J Biol Chem*. 2001; 276:30514–30520. [PubMed: 11395522]
32. Gong C, Martins A, Shuman S. Structure-function analysis of Trypanosoma brucei RNA triphosphatase and evidence for a two-metal mechanism. *J Biol Chem*. 2003; 278:50843–50852. [PubMed: 14525979]
33. Hausmann S, Vivares CP, Shuman S. Characterization of the mRNA capping apparatus of the microsporidian parasite Encephalitozoon cuniculi. *J Biol Chem*. 2002; 277:96–103. [PubMed: 11687593]
34. Katinka MD, Duprat S, Cornillot E, Metenier G, Thomarat F, et al. Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi. *Nature*. 2001; 414:450–453. [PubMed: 11719806]
35. Benarroch D, Smith P, Shuman S. Characterization of a trifunctional mimivirus mRNA capping enzyme and crystal structure of the RNA triphosphatase domain. *Structure*. 2008; 16:501–512. [PubMed: 18400173]
36. Takagi T, Moore CR, Diehn F, Buratowski S. An RNA 5'-triphosphatase related to the protein tyrosine phosphatases. *Cell*. 1997; 89:867–873. [PubMed: 9200605]

37. Changela A, Ho CK, Martins A, Shuman S, Mondragon A. Structure and mechanism of the RNA triphosphatase component of mammalian mRNA capping enzyme. *EMBO J*. 2001; 20:2575–2586. [PubMed: 11350947]
38. Shuman S, Lima CD. The polynucleotide ligase and RNA capping enzyme superfamily of covalent nucleotidyltransferases. *Curr Opin Struct Biol*. 2004; 14:757–764. [PubMed: 15582400]
39. Fabrega C, Shen V, Shuman S, Lima CD. Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol Cell*. 2003; 11:1549–1561. [PubMed: 12820968]
40. Hakansson K, Doherty AJ, Shuman S, Wigley DB. X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes. *Cell*. 1997; 89:545–553. [PubMed: 9160746]
41. Shuman S, Schwer B. RNA capping enzyme and DNA ligase: a superfamily of covalent nucleotidyl transferases. *Mol Microbiol*. 1995; 17:405–410. [PubMed: 8559059]
42. Shuman S. NAD⁺ specificity of bacterial DNA ligase revealed. *Structure*. 2004; 12:1335–1336. [PubMed: 15296724]
43. Fabrega C, Hausmann S, Shen V, Shuman S, Lima CD. Structure and mechanism of mRNA cap (guanine-N7) methyltransferase. *Mol Cell*. 2004; 13:77–89. [PubMed: 14731396]
44. Zheng S, Hausmann S, Liu Q, Ghosh A, Schwer B, et al. Mutational analysis of *Encephalitozoon cuniculi* mRNA cap (guanine-N7) methyltransferase, structure of the enzyme bound to sinefungin, and evidence that cap methyltransferase is the target of sinefungin's antifungal activity. *J Biol Chem*. 2006; 281:35904–35913. [PubMed: 16971388]
45. Cong P, Shuman S. Methyltransferase and subunit association domains of vaccinia virus mRNA capping enzyme. *J Biol Chem*. 1992; 267:16424–16429. [PubMed: 1322901]
46. Higman MA, Bourgeois N, Niles EG. The vaccinia virus mRNA (guanine-N7-)-methyltransferase requires both subunits of the mRNA capping enzyme for activity. *J Biol Chem*. 1992; 267:16430–16437. [PubMed: 1322902]
47. Higman MA, Christen LA, Niles EG. The mRNA (guanine-7-)-methyltransferase domain of the vaccinia virus mRNA capping enzyme. Expression in *Escherichia coli* and structural and kinetic comparison to the intact capping enzyme. *J Biol Chem*. 1994; 269:14974–14981. [PubMed: 8195132]
48. Higman MA, Niles EG. Location of the S-adenosyl-L-methionine binding region of the vaccinia virus mRNA (guanine-7-)-methyltransferase. *J Biol Chem*. 1994; 269:14982–14987. [PubMed: 8195133]
49. Mao X, Shuman S. Intrinsic RNA (guanine-7) methyltransferase activity of the vaccinia virus capping enzyme D1 subunit is stimulated by the D12 subunit. Identification of amino acid residues in the D1 protein required for subunit association and methyl group transfer. *J Biol Chem*. 1994; 269:24472–24479. [PubMed: 7929111]
50. De la Pena M, Kyrieleis OJ, Cusack S. Structural insights into the mechanism and evolution of the vaccinia virus mRNA cap N7 methyl-transferase. *EMBO J*. 2007; 26:4913–4925. [PubMed: 17989694]
51. Zheng S, Shuman S. Mutational analysis of vaccinia virus mRNA cap (guanine-N7) methyltransferase reveals essential contributions of the N-terminal peptide that closes over the active site. *RNA*. 2008; 14:2297–2304. [PubMed: 18799596]
52. Mao X, Schwer B, Shuman S. Mutational analysis of the *Saccharomyces cerevisiae* ABD1 gene: cap methyltransferase activity is essential for cell growth. *Mol Cell Biol*. 1996; 16:475–480. [PubMed: 8552073]
53. Shibagaki Y, Itoh N, Yamada H, Nagata S, Mizumoto K. mRNA capping enzyme. Isolation and characterization of the gene encoding mRNA guanylyltransferase subunit from *Saccharomyces cerevisiae*. *J Biol Chem*. 1992; 267:9521–9528. [PubMed: 1315757]
54. Ho CK, Sriskanda V, McCracken S, Bentley D, Schwer B, et al. The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem*. 1998; 273:9577–9585. [PubMed: 9545288]

55. Pillutla RC, Yue Z, Maldonado E, Shatkin AJ. Recombinant human mRNA cap methyltransferase binds capping enzyme/RNA polymerase II complexes. *J Biol Chem.* 1998; 273:21443–21446. [PubMed: 9705270]
56. Lehman K, Schwer B, Ho CK, Rouzankina I, Shuman S. A conserved domain of yeast RNA triphosphatase flanking the catalytic core regulates self-association and interaction with the guanylyltransferase component of the mRNA capping apparatus. *J Biol Chem.* 1999; 274:22668–22678. [PubMed: 10428848]
57. Ho CK, Schwer B, Shuman S. Genetic, physical, and functional interactions between the triphosphatase and guanylyltransferase components of the yeast mRNA capping apparatus. *Mol Cell Biol.* 1998; 18:5189–5198. [PubMed: 9710603]
58. Cho EJ, Rodriguez CR, Takagi T, Buratowski S. Allosteric interactions between capping enzyme subunits and the RNA polymerase II carboxy-terminal domain. *Genes Dev.* 1998; 12:3482–3487. [PubMed: 9832501]
59. Ho CK, Lehman K, Shuman S. An essential surface motif (WAQKW) of yeast RNA triphosphatase mediates formation of the mRNA capping enzyme complex with RNA guanylyltransferase. *Nucleic Acids Res.* 1999; 27:4671–4678. [PubMed: 10572165]
60. Takase Y, Takagi T, Komarnitsky PB, Buratowski S. The essential interaction between yeast mRNA capping enzyme subunits is not required for triphosphatase function in vivo. *Mol Cell Biol.* 2000; 20:9307–9316. [PubMed: 11094081]
61. Hausmann S, Ho CK, Schwer B, Shuman S. An essential function of *Saccharomyces cerevisiae* RNA triphosphatase Cet1 is to stabilize RNA guanylyltransferase Ceg1 against thermal inactivation. *J Biol Chem.* 2001; 276:36116–36124. [PubMed: 11463793]
62. Cho EJ, Takagi T, Moore CR, Buratowski S. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev.* 1997; 11:3319–3326. [PubMed: 9407025]
63. McCracken S, Fong N, Rosonina E, Yankulov K, Brothers G, et al. 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.* 1997; 11:3306–3318. [PubMed: 9407024]
64. Gu M, Rajashankar KR, Lima CD. Structure of the *Saccharomyces cerevisiae* Cet1-Ceg1 mRNA capping apparatus. *Structure.* 2010; 18:216–227. [PubMed: 20159466]
65. Hausmann S, Pei Y, Shuman S. Homodimeric quaternary structure is required for the in vivo function and thermal stability of *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* RNA triphosphatases. *J Biol Chem.* 2003; 278:30487–30496. [PubMed: 12788946]
66. Pei Y, Hausmann S, Ho CK, Schwer B, Shuman S. The length, phosphorylation state, and primary structure of the RNA polymerase II carboxyl-terminal domain dictate interactions with mRNA capping enzymes. *J Biol Chem.* 2001; 276:28075–28082. [PubMed: 11387325]
67. Martins A, Shuman S. Mechanism of phosphoanhydride cleavage by baculovirus phosphatase. *J Biol Chem.* 2000; 275:35070–35076. [PubMed: 10954717]
68. Changela A, Martins A, Shuman S, Mondragon A. Crystal structure of baculovirus RNA triphosphatase complexed with phosphate. *J Biol Chem.* 2005; 280:17848–17856. [PubMed: 15713658]
69. Guarino LA, Jin J, Dong W. Guanylyltransferase activity of the LEF-4 subunit of baculovirus RNA polymerase. *J Virol.* 1998; 72:10003–10010. [PubMed: 9811738]
70. Guarino LA, Xu B, Jin J, Dong W. A virus-encoded RNA polymerase purified from baculovirus-infected cells. *J Virol.* 1998; 72:7985–7991. [PubMed: 9733837]
71. Jin J, Dong W, Guarino LA. The LEF-4 subunit of baculovirus RNA polymerase has RNA 5'-triphosphatase and ATPase activities. *J Virol.* 1998; 72:10011–10019. [PubMed: 9811739]
72. He JG, Deng M, Weng SP, Li Z, Zhou SY, et al. Complete genome analysis of the mandarin fish infectious spleen and kidney necrosis iridovirus. *Virology.* 2001; 291:126–139. [PubMed: 11878882]
73. Martins A, Shuman S. Mutational analysis of baculovirus capping enzyme Lef4 delineates an autonomous triphosphatase domain and structural determinants of divalent cation specificity. *J Biol Chem.* 2001; 276:45522–45529. [PubMed: 11553638]

74. Ramadevi N, Burroughs NJ, Mertens PP, Jones IM, Roy P. Capping and methylation of mRNA by purified recombinant VP4 protein of bluetongue virus. *Proc Natl Acad Sci U S A*. 1998; 95:13537–13542. [PubMed: 9811835]
75. Ramadevi N, Roy P. Bluetongue virus core protein VP4 has nucleoside triphosphate phosphohydrolase activity. *J Gen Virol*. 1998; 79:2475–2480. [PubMed: 9780054]
76. Reinisch KM, Nibert ML, Harrison SC. Structure of the reovirus core at 3.6 Å resolution. *Nature*. 2000; 404:960–967. [PubMed: 10801118]
77. Sutton G, Grimes JM, Stuart DI, Roy P. Bluetongue virus VP4 is an RNA-capping assembly line. *Nat Struct Mol Biol*. 2007; 14:449–451. [PubMed: 17417654]
78. Coppola JA, Field AS, Luse DS. Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped. *Proc Natl Acad Sci U S A*. 1983; 80:1251–1255. [PubMed: 6572384]
79. Perales R, Bentley D. “Cotranscriptionality”: the transcription elongation complex as a nexus for nuclear transactions. *Mol Cell*. 2009; 36:178–191. [PubMed: 19854129]
80. Broyles SS, Moss B. Sedimentation of an RNA polymerase complex from vaccinia virus that specifically initiates and terminates transcription. *Mol Cell Biol*. 1987; 7:7–14. [PubMed: 3031483]
81. Hagler J, Shuman S. A freeze-frame view of eukaryotic transcription during elongation and capping of nascent mRNA. *Science*. 1992; 255:983–986. [PubMed: 1546295]
82. Phatnani HP, Greenleaf AL. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev*. 2006; 20:2922–2936. [PubMed: 17079683]
83. Buratowski S. Progression through the RNA polymerase II CTD cycle. *Mol Cell*. 2009; 36:541–546. [PubMed: 19941815]
84. Egloff S, Murphy S. Cracking the RNA polymerase II CTD code. *Trends Genet*. 2008; 24:280–288. [PubMed: 18457900]
85. Lima CD. Inducing interactions with the CTD. *Nat Struct Mol Biol*. 2005; 12:102–103. [PubMed: 15702066]
86. Corden JL. Tails of RNA polymerase II. *Trends Biochem Sci*. 1990; 15:383–387. [PubMed: 2251729]
87. Gerber HP, Hagmann M, Seipel K, Georgiev O, West MA, et al. RNA polymerase II C-terminal domain required for enhancer-driven transcription. *Nature*. 1995; 374:660–662. [PubMed: 7715709]
88. Chapman RD, Conrad M, Eick D. Role of the mammalian RNA polymerase II C-terminal domain (CTD) nonconsensus repeats in CTD stability and cell proliferation. *Mol Cell Biol*. 2005; 25:7665–7674. [PubMed: 16107713]
89. West ML, Corden JL. Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. *Genetics*. 1995; 140:1223–1233. [PubMed: 7498765]
90. Ho CK, Shuman S. Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol Cell*. 1999; 3:405–411. [PubMed: 10198643]
91. Komarnitsky P, Cho EJ, Buratowski S. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev*. 2000; 14:2452–2460. [PubMed: 11018013]
92. Schroeder SC, Zorio DA, Schwer B, Shuman S, Bentley D. A function of yeast mRNA cap methyltransferase, Abd1, in transcription by RNA polymerase II. *Mol Cell*. 2004; 13:377–387. [PubMed: 14967145]
93. Yue Z, Maldonado E, Pillutla R, Cho H, Reinberg D, et al. Mammalian capping enzyme complements mutant *Saccharomyces cerevisiae* lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II. *Proc Natl Acad Sci U S A*. 1997; 94:12898–12903. [PubMed: 9371772]
94. Takagi T, Cho EJ, Janoo RT, Polodny V, Takase Y, et al. Divergent subunit interactions among fungal mRNA 5'-capping machineries. *Eukaryot Cell*. 2002; 1:448–457. [PubMed: 12455993]

95. Iyer LM, Aravind L. The catalytic domains of thiamine triphosphatase and CyaB-like adenylyl cyclase define a novel superfamily of domains that bind organic phosphates. *BMC Genomics*. 2002; 3:33. [PubMed: 12456267]
96. Hothorn M, Neumann H, Lenherr ED, Wehner M, Rybin V, et al. Catalytic core of a membrane-associated eukaryotic polyphosphate polymerase. *Science*. 2009; 324:513–516. [PubMed: 19390046]
97. Schubert HL, Blumenthal RM, Cheng X. Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem Sci*. 2003; 28:329–335. [PubMed: 12826405]
98. Ogino T, Banerjee AK. Formation of guanosine(5')tetraphospho(5')adenosine cap structure by an unconventional mRNA capping enzyme of vesicular stomatitis virus. *J Virol*. 2008; 82:7729–7734. [PubMed: 18495767]
99. Ogino T, Yadav SP, Banerjee AK. Histidine-mediated RNA transfer to GDP for unique mRNA capping by vesicular stomatitis virus RNA polymerase. *Proc Natl Acad Sci U S A*. 2010; 107:3463–3468. [PubMed: 20142503]
100. Plotch SJ, Bouloy M, Ulmanen I, Krug RM. A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell*. 1981; 23:847–858. [PubMed: 6261960]
101. Rao P, Yuan W, Krug RM. Crucial role of CA cleavage sites in the cap-snatching mechanism for initiating viral mRNA synthesis. *EMBO J*. 2003; 22:1188–1198. [PubMed: 12606583]
102. Cianci C, Tiley L, Krystal M. Differential activation of the influenza virus polymerase via template RNA binding. *J Virol*. 1995; 69:3995–3999. [PubMed: 7769657]
103. Sugiyama K, Obayashi E, Kawaguchi A, Suzuki Y, Tame JR, et al. Structural insight into the essential PB1-PB2 subunit contact of the influenza virus RNA polymerase. *EMBO J*. 2009; 28:1803–1811. [PubMed: 19461581]
104. Guilligay D, Tarendeau F, Resa-Infante P, Coloma R, Crepin T, et al. The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat Struct Mol Biol*. 2008; 15:500–506. [PubMed: 18454157]
105. Dias A, Bouvier D, Crepin T, McCarthy AA, Hart DJ, et al. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature*. 2009; 458:914–918. [PubMed: 19194459]
106. Li ML, Rao P, Krug RM. The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits. *EMBO J*. 2001; 20:2078–2086. [PubMed: 11296240]
107. Yuan P, Bartlam M, Lou Z, Chen S, Zhou J, et al. Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. *Nature*. 2009; 458:909–913. [PubMed: 19194458]
108. Ahola T, Kaariainen L. Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. *Proc Natl Acad Sci U S A*. 1995; 92:507–511. [PubMed: 7831320]
109. Gershon PD, Ahn BY, Garfield M, Moss B. Poly(A) polymerase and a dissociable polyadenylation stimulatory factor encoded by vaccinia virus. *Cell*. 1991; 66:1269–1278. [PubMed: 1670500]
110. Schnierle BS, Gershon PD, Moss B. Cap-specific mRNA (nucleoside-O2')-methyltransferase and poly(A) polymerase stimulatory activities of vaccinia virus are mediated by a single protein. *Proc Natl Acad Sci U S A*. 1992; 89:2897–2901. [PubMed: 1313572]
111. Barbosa E, Moss B. mRNA(nucleoside-2')-methyl transferase from vaccinia virus. Characteristics and substrate specificity. *J Biol Chem*. 1978; 253:7698–7702. [PubMed: 701282]
112. Lockless SW, Cheng HT, Hodel AE, Quijcho FA, Gershon PD. Recognition of capped RNA substrates by VP39, the vaccinia virus-encoded mRNA cap-specific 2'-O-methyltransferase. *Biochemistry*. 1998; 37:8564–8574. [PubMed: 9622508]
113. Hodel AE, Gershon PD, Quijcho FA. Structural basis for sequence-nonspecific recognition of 5'-capped mRNA by a cap-modifying enzyme. *Mol Cell*. 1998; 1:443–447. [PubMed: 9660928]
114. Hodel AE, Gershon PD, Shi X, Quijcho FA. The 1.85 Å structure of vaccinia protein VP39: a bifunctional enzyme that participates in the modification of both mRNA ends. *Cell*. 1996; 85:247–256. [PubMed: 8612277]

115. Hodel AE, Gershon PD, Shi X, Wang SM, Quirocho FA. Specific protein recognition of an mRNA cap through its alkylated base. *Nat Struct Biol.* 1997; 4:350–354. [PubMed: 9145102]
116. Hu G, Gershon PD, Hodel AE, Quirocho FA. mRNA cap recognition: dominant role of enhanced stacking interactions between methylated bases and protein aromatic side chains. *Proc Natl Acad Sci U S A.* 1999; 96:7149–7154. [PubMed: 10377383]
117. Cleaves GR, Dubin DT. Methylation status of intracellular dengue type 2 40 S RNA. *Virology.* 1979; 96:159–165. [PubMed: 111410]
118. Egloff MP, Benarroch D, Selisko B, Romette JL, Canard B. An RNA cap (nucleoside-2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. *EMBO J.* 2002; 21:2757–2768. [PubMed: 12032088]
119. Peyrane F, Selisko B, Decroly E, Vasseur JJ, Benarroch D, et al. High-yield production of short GpppA- and 7MeGpppA-capped RNAs and HPLC-monitoring of methyltransfer reactions at the guanine-N7 and adenosine-2'O positions. *Nucleic Acids Res.* 2007; 35:e26. [PubMed: 17259217]
120. Ray D, Shah A, Tilgner M, Guo Y, Zhao Y, et al. West Nile virus 5'-cap structure is formed by sequential guanine N-7 and ribose 2'-O methylations by nonstructural protein 5. *J Virol.* 2006; 80:8362–8370. [PubMed: 16912287]
121. Zhou Y, Ray D, Zhao Y, Dong H, Ren S, et al. Structure and function of flavivirus NS5 methyltransferase. *J Virol.* 2007; 81:3891–3903. [PubMed: 17267492]
122. Rahmeh AA, Li J, Kranzusch PJ, Whelan SP. Ribose 2'-O methylation of the vesicular stomatitis virus mRNA cap precedes and facilitates subsequent guanine-N-7 methylation by the large polymerase protein. *J Virol.* 2009; 83:11043–11050. [PubMed: 19710136]
123. Ullu E, Tschudi C. Trans splicing in trypanosomes requires methylation of the 5' end of the spliced leader RNA. *Proc Natl Acad Sci U S A.* 1991; 88:10074–10078. [PubMed: 1719544]
124. Ullu E, Tschudi C. Accurate modification of the trypanosome spliced leader cap structure in a homologous cell-free system. *J Biol Chem.* 1995; 270:20365–20369. [PubMed: 7657609]
125. Zeiner GM, Sturm NR, Campbell DA. Exportin 1 mediates nuclear export of the kinetoplastid spliced leader RNA. *Eukaryot Cell.* 2003; 2:222–230. [PubMed: 12684371]
126. Mittra B, Zamudio JR, Bujnicki JM, Stepinski J, Darzynkiewicz E, et al. The TbMTr1 spliced leader RNA cap 1 2'-O-ribose methyltransferase from *Trypanosoma brucei* acts with substrate specificity. *J Biol Chem.* 2008; 283:3161–3172. [PubMed: 18048356]
127. Mouaikel J, Bujnicki JM, Tazi J, Bordonne R. Sequence-structure-function relationships of Tgs1, the yeast snRNA/snoRNA cap hypermethylase. *Nucleic Acids Res.* 2003; 31:4899–4909. [PubMed: 12907733]
128. Franke J, Gehlen J, Ehrenhofer-Murray AE. Hypermethylation of yeast telomerase RNA by the snRNA and snoRNA methyltransferase Tgs1. *J Cell Sci.* 2008; 121:3553–3560. [PubMed: 18840651]
129. Monecke T, Dickmanns A, Ficner R. Structural basis for m7G-cap hypermethylation of small nuclear, small nucleolar and telomerase RNA by the dimethyltransferase TGS1. *Nucleic Acids Res.* 2009; 37:3865–3877. [PubMed: 19386620]
130. Monecke T, Dickmanns A, Strasser A, Ficner R. Structure analysis of the conserved methyltransferase domain of human trimethylguanosine synthase TGS1. *Acta Crystallogr D Biol Crystallogr.* 2009; 65:332–338. [PubMed: 19307714]
131. Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, et al. Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell.* 2007; 27:262–274. [PubMed: 17643375]
132. Shuman S. Transcriptional networking captures the 7SK RNA 5'-gamma-methyltransferase. *Mol Cell.* 2007; 27:517–519. [PubMed: 17707222]
133. Furuichi Y, Shatkin AJ. Viral and cellular mRNA capping: past and prospects. *Adv Virus Res.* 2000; 55:135–184. [PubMed: 11050942]
134. Shuman S. The mRNA capping apparatus as drug target and guide to eukaryotic phylogeny. *Cold Spring Harb Symp Quant Biol.* 2001; 66:301–312. [PubMed: 12762032]
135. De Clercq E. John Montgomery's legacy: carbocyclic adenosine analogues as SAH hydrolase inhibitors with broad-spectrum antiviral activity. *Nucleosides Nucleotides Nucleic Acids.* 2005; 24:1395–1415. [PubMed: 16438025]

136. Liu S, Yuan CS, Borchardt RT. Aristeromycin-5'-carboxaldehyde: a potent inhibitor of S-adenosyl-L-homocysteine hydrolase. *J Med Chem.* 1996; 39:2347–2353. [PubMed: 8691429]
137. Chrebet GL, Wisniewski D, Perkins AL, Deng Q, Kurtz MB, et al. Cell-based assays to detect inhibitors of fungal mRNA capping enzymes and characterization of sinefungin as a cap methyltransferase inhibitor. *J Biomol Screen.* 2005; 10:355–364. [PubMed: 15964937]

FIGURE 1.

Structures and synthesis of RNA caps. Cap 0 or m⁷G caps (GMP colored blue) are formed in sequential steps by three enzymatic activities that act on the 5' triphosphate end (colored red) of nascent transcripts. Transfer of a methyl group (colored green) from *S*-adenosylmethionine (AdoMet) completes the synthesis of cap 0. Transfer of two methyl groups (colored magenta) from *S*-adenosylmethionine (AdoMet) is required to form the TMG cap. Caps 1 and 2 structures require methylation (colored magenta) of cap 0 at the ribose 2'-O at the first and second nucleosides, respectively. Cap 4 structure is generated by six rounds of methylation (colored magenta). The first three rounds of methylation (colored magenta) occur at two positions on the base and ribose of the first nucleoside of the primary transcript. The next three rounds of methylation (colored magenta) occur on ribose 2'-O positions of the next three nucleosides. The γ -methyl phosphate cap is formed by transfer of a methyl group (colored magenta) from AdoMet to the γ -phosphate of the primary transcript.

FIGURE 2.

Mechanisms of cap 0 synthesis. Triphosphatase activities: Metal-dependent (top) or metal-independent RNA triphosphatases (bottom) catalyze removal of the γ -phosphate (colored red) from pppRNA to generate ppRNA and release of inorganic phosphate (colored black). The tunnel architecture of the metal-dependent RNA triphosphatase Cet1 indicating side chains that coordinate pppRNA and two metals, one (solid sphere) derived from structural studies and the other (circle containing M) derived from biochemical studies. Metal-independent RNA triphosphatases catalyze removal of the γ -phosphate through a two-step reaction through a covalent protein-cysteinylyl-S-phosphate intermediate (step 1, shown in the figure) colored as in top panel. Guanylyltransferase activities: RNA guanylyltransferase catalyzes capping in a two-step reaction. In step 1, the enzyme binds GTP (colored blue) and magnesium (not shown) to catalyze transfer of GMP to the active site lysine to form a covalent enzyme(lysyl-N)-GMP intermediate. In step 2, the enzyme binds the ppRNA (colored red) to catalyze transfer of the GMP to form GpppRNA. Methyltransferase activities: RNA guanine-N7 methyltransferase binds S-adenosylmethionine (AdoMet) (colored green) and GpppRNA (colored as above) to catalyze transfer of the methyl group (colored green) to the guanine N7 position.

FIGURE 3.

Structures of yeast, viral, and metazoan RNA triphosphatases. (a) Structure of homodimeric metal-dependent RNA triphosphatase Cet1 (PDB 1D8H) from *S. cerevisiae* with a view looking into the triphosphate tunnels. The structure is depicted in ribbon representation with loops and α -helices colored gray and β -strands colored salmon. The divalent cation manganese and sulfate ion are labeled and depicted in ball-and-stick representation. Basic and acidic residues that point into the tunnel are shown in sticks. Amino- and carboxy-termini are denoted 'N' and 'C', respectively. Disordered segments are shown as dashed lines. (b) Structure of the monomeric RNA triphosphatase domain (amino acid 1–237) of *mimivirus* capping enzyme, MimiCE (PDB 3BGY) colored and structurally depicted as in (a). (c) The RNA triphosphatase domain of mammalian mRNA capping enzyme Mce1 (PDB 1I9S). Ribbon representation colored pink with the active site cysteine (Cys126) denoted in ball-and-stick representation and labeled. Structural graphics prepared using PyMOL (<http://pymol.sourceforge.net/>).

FIGURE 4.

Structures of viral and yeast RNA guanylyltransferases. (a) Structure of *Chlorella virus* capping enzyme PbCV-1 in 'open' configuration (PDB 1CKN) depicting the nucleotidyltransferase (NTase) domain (blue) and OB domain (yellow) bound to GTP (labeled). The active site lysine (Lys82) and GTP are shown in stick representation. The position of the OB-fold domain (green) in the 'closed' conformation was superimposed on the 'open' conformation based on alignment of the respective NTase domains. (b) Structure of the *Candida albicans* RNA guanylyltransferase, Cgt1 (PDB 1P16; chain A). The covalent lysyl-GMP adduct is represented in stick with NTase (blue) and OB-fold (yellow) domains shown in ribbon representation. View of Cgt1 generated by aligning its NTase domain to that in (a). OB-fold domains of (c) *C. albicans* Cgt1 and (d) *S. cerevisiae* Ceg1 (PDB 3KYH) to illustrate conserved OB elements (yellow) and insertion elements (salmon) that were shown to contribute to interactions with the triphosphatase. Disordered regions are represented by dashed lines.

FIGURE 5.

Structures of viral, protozoan, and metazoan guanine-N7 methyltransferases. (a) Orthogonal views of *vaccinia* virus guanine-N7 methyltransferase (PDB 2VDW) depicting the D1 catalytic subunit (marine) and D12 regulatory subunit (yellow) in ribbon diagram. S-adenosyl-homocysteine (AdoHcy) and a sulfate ion in the D1 active site are represented in ball-and-stick. The N-terminal peptide (red) that covers AdoHcy depicted as a ribbon and transparent molecular surface. Dashed lines denote disordered regions of the structure. (b) Structure of RNA guanine-N7 methyltransferase from *Encephalitozoon cuniculi*, Ecm1 (PDB 1RI1) shown in a similar view to that presented for the D1 subunit in the left panel of (a). GTP and AdoHcy are shown in the Ecm1 active site in ball-and-stick. (c) Structure of human RNA guanine-N7 methyltransferase Hcm1 (PDB 3EPP) aligned as in (b) with AdoHcy shown in ball-and-stick. Amino- and carboxy-termini denoted as 'N' and 'C', respectively.

FIGURE 6.

Schematic diagram depicting various architectures of the capping apparatus. Domain organizations in different taxa showing the tunnel shaped metal-dependent RNA triphosphatase (TPase; pink with white interior), the nucleotidyltransferase (NTase; blue), the OB-fold (OB; yellow), and guanine-N7 methyltransferase (N7-MTase; green). For ISKNV only the metal-independent TPase and NTase domains are shown, as the virus is not known to encode an N7-MTase. For the *bluetongue* virus capping enzyme, the KL domain (gold) and 2'-O MTase (gray) are also included. Domains are shown as boxes, catalytic residues are labeled and indicated by red vertical lines, and flexible linkers between domains denoted by dashed lines.

FIGURE 7.

Structures of the capping apparatus from *S. cerevisiae* and *bluetongue* virus and a complex between the phosphorylated RNAP II CTD and *C. albicans* guanylyltransferase. (a) Structure of a complex between *S. cerevisiae* Cet1 and Ceg1. The Cet1 homodimer is shown with Cet1 protomers colored cyan and magenta. The Ceg1 NTase domain is colored blue and OB domain is colored yellow. Dashed lines indicate disordered amino acids. The Cet1 WAQKW motifs are represented in ball-and-stick and specified amino acid side chains are labeled. Lys70 is shown in Ceg1 to indicate the position of the guanylyltransferase active site. (b) Structure of *bluetongue* virus VP4 (PDB 2HJA) in ribbon representation indicating the following domains: kinase-like (KL; gold), N7-methyltransferase (N7 MTase; green), 2'-O methyltransferase (2'-O MTase; gray), and triphosphatase/nucleotidyltransferase (TPase/NTase; blue). GpppG shown in the 2'-O MTase active site with S-adenosyl-homocysteine (AdoHcy, PDB 2HJP) shown in both N7 and 2'-O MTase active sites in stick representation. Dashed lines indicated disordered regions. (c) The structure of *C. albicans* Cgt1 (PDB 1P16; chain B) with NTase (blue), OB domain (yellow) bound to a Ser5 phosphorylated RNAP II CTD peptide (green). Close-up view of the CTD binding site with positions for Tyrosine 1, Proline 3, and phosphorylated Ser5 indicated. The CTD and GTP in the active site shown in stick representation. Dashed lines denote disordered regions.

FIGURE 8.

Schematic models for co-transcriptional mRNA capping in *S. cerevisiae*, *S. pombe* and mammals. In *S. cerevisiae*, the RNAP II CTD (light-green) is shown hyperphosphorylated at Ser2 and Ser5 positions (gold spheres with P). The RNA guanylyltransferase interacting with the phosphorylated CTD via the NTase domain (blue) and Cet1 (magenta) via the OB domain (yellow). Cet1 and Ceg1 are tethered by a flexible linker (dashed magenta line). The RNA guanine-N7 methyltransferase Abd1 (green) shown independently interacting with the phosphorylated RNAP II CTD. In *S. pombe*, the RNA triphosphatase Pct1 (magenta), RNA guanylyltransferase Pce1 (NTase blue, OB yellow), and RNA guanine-N7 methyltransferase Pcm1 (green) are shown interacting independently with the phosphorylated RNAP II CTD. The bifunctional mammalian capping enzyme shown with the *N*-terminal RNA triphosphatase domain (pink) and RNA guanylyltransferase domain linked by a flexible tether (dashed line) with the RNA guanylyltransferase NTase domain (blue) and OB domain (yellow). Interactions between the bifunctional mammalian capping enzyme and phosphorylated RNAP II CTD are mediated by the NTase domain. The mammalian RNA

guanine-N7 methyltransferase (green) shown interacting independently with the phosphorylated RNAP II CTD. The nascent mRNA is denoted by red lines in each of the panels with the triphosphate terminus indicated by gold spheres.

FIGURE 9.

Unconventional pathways for cap formation. (a) Schematic of the proposed mechanism of mRNA cap formation in *rhabdovirus* with protein colored purple, catalytic histidine residue denoted in red, nascent mRNA colored red, terminal phosphates of nascent mRNA and phosphatases of the guanine nucleotide shown as yellow balls with phosphate positions labeled. (b) Schematic of *influenza virus* 'cap snatching' indicating the host capped mRNA (blue), the viral RNA (red), the PB2 subunit (green), the PA subunit (orange), and the PB1 subunit (colored in cyan). (c) Schematic of the proposed pathway for capping in *alphavirus*. The nsP1 protein (pink) transfers a methyl group (green) from AdoMet (black) to GTP (blue and orange spheres) to generate N7-methyl GTP. The N7-methyl GTP is then used as a substrate to cap the viral RNA (red).