Conventional and unconventional mechanisms for capping viral mRNA

Etienne Decroly¹, François Ferron¹, Julien Lescar¹,² and Bruno Canard¹

Abstract | In the eukaryotic cell, capping of mRNA 5′ ends is an essential structural modification that allows efficient mRNA translation, directs pre-mRNA splicing and mRNA export from the nucleus, limits mRNA degradation by cellular 5′–3′ exonucleases and allows recognition of foreign RNAs (including viral transcripts) as ‘non-self’. However, viruses have evolved mechanisms to protect their RNA 5′ ends with either a covalently attached peptide or a cap moiety (7-methyl-Gppp, in which p is a phosphate group) that is indistinguishable from cellular mRNA cap structures. Viral RNA caps can be stolen from cellular mRNAs or synthesized using either a host- or virus-encoded capping apparatus, and these capping assemblies exhibit a wide diversity in organization, structure and mechanism. Here, we review the strategies used by viruses of eukaryotic cells to produce functional mRNA 5′-caps and escape innate immunity.

The cap structure found at the 5′ end of eukaryotic mRNAs consists of a 7-methylguanosine (m7G) moiety linked to the first nucleotide of the transcript via a 5′–5′ triphosphate bridge (FIG. 1a). The cap has several important biological roles, such as protecting mRNA from degradation by 5′ exoribonucleases and directing pre-mRNA splicing and mRNA export from the nucleus. In addition, the cap confers stability to mRNAs and ensures their efficient recognition by eukaryotic translation initiation factor 4E (eIF4E) for translation.

Conversely, RNA molecules with unprotected 5′ ends are degraded in cytoplasmic granular compartments called processing bodies (P-bodies). Uncapped RNAs, such as nascent viral transcripts, may also be detected as ‘non-self’ by the host cell, triggering (in mammalian cells) an antiviral innate immune response through the production of interferons.

During virus–host co-evolution, viral RNA-capping pathways that lead to the same cap structure as that of host mRNAs have been selected for as efficient mechanisms to ensure both escape from detection by the innate immune system and efficient production of viral proteins. Compared with canonical eukaryotic mRNA capping, viral mRNA capping is highly diverse in terms of its genetic components, protein domain organization, enzyme structures and reaction mechanisms. Here, we review what is known about the various mRNA-capping pathways used by viruses that infect eukaryotes, paying particular attention to human pathogens. We also attempt to connect the pathways, machineries, structures and reactions involved in the viral RNA-capping process, and the specific cellular factors that trigger a response from the innate immune system.

Capping, decapping and turnover of host RNA

Nascent cellular mRNAs are generally produced in the nucleus in a 5′-triphosphate form and are modified co-transcriptionally by a set of cap-synthesizing enzymes. These enzymes are recruited by the DNA-dependent RNA polymerase RNA pol II on pausing, once the transcript is approximately 20–25 nucleotides long. Interestingly, two viruses — vaccinia virus (from the Poxviridae family of double-stranded DNA (dsDNA) viruses) and mammalian orthoreovirus (from the dsRNA Reoviridae family) — played a major part in the discovery of the RNA cap.¹⁰,¹¹, because they produce high levels of capped viral mRNAs and encode their own capping machinery, allowing bona fide RNA cap synthesis in vitro. Mammalian orthoreovirus mRNAs and dsRNA genomes were first shown to be blocked at the 5′ end, preventing phosphorylation by polynucleotide kinase.¹⁰,¹¹. Purified mammalian orthoreovirus preparations were then demonstrated to be able to methylate mRNA 5′ ends, and the full mRNA cap structures were deciphered for both dsRNA orthoreoviruses (those targeting humans and insects) and the dsRNA species vaccinia virus, followed by many other viral species.¹⁰,¹¹,¹⁴

The three canonical capping reactions responsible for the synthesis of the cap structure are outlined in
**REVIEWS**

**a Chemical structure of the RNA cap**

![Chemical structure of the RNA cap](image)

- Triphosphate
- RNA 5’ end
- Base 1
- Base 2
- 7-methylguanosine
- CH₃
- HN
- N³
- OH
- HO
- O

**b Conventional RNA-capping pathway**

![Conventional RNA-capping pathway](image)

- Triphosphate
- RNA 5’ end
- Base 1
- Base 2
- 7-methylguanosine
- CH₃
- HN
- N³
- OH
- HO
- O

**y-phosphate**
The third phosphate attached at the 5’ end of the ribose moiety of a nucleotide.

**‘Ping-pong’ mechanism**
A two-step mechanism in which a substrate molecule first forms a (covalent) link with the enzyme and is then transferred to an acceptor molecule to yield a product.

**Poly(A) tail**
A string of AMP that is added to the 5’ end of mRNA.

**NUDIX hydrolase superfamily**
A family of proteins that hydrolyse a wide range of organic pyrophosphates, including NDPs, NTPs, dinucleoside diphosphonositol polyphosphates, nucleotide sugars and RNA caps, with varying degrees of substrate specificity.

**Figure 1** | RNA cap structure and canonical capping mechanisms. a) The mRNA cap consists of a 7-methylguanosine linked to the 5’ nucleoside of the mRNA chain through a 5’–5’ triphosphate bridge. The methyl group at the N7 position of the guanosine is shaded green, and the 2’-O-methyl groups of the first and second nucleotide residues, forming the cap-1 and the cap-2 structures, respectively, are shaded red. b) The cap-0 structure is formed on nascent RNA chains by the sequential action of three enzymes. First, the RNA triphosphatase (RTPase) hydrolyses the y-phosphate of the nascent RNA (pppNp-RNA, in which N denotes the first transcribed nucleotide and p denotes a phosphate group) to yield a diphosphate RNA (ppNp-RNA) and inorganic phosphate (Pᵢ). Then, guanylyltransferase (GTase) reacts with the α-phosphate of GTP (Gppp), releasing pyrophosphate (PPᵢ) and forming a covalent enzyme–guanylate intermediate (Gp–GTase). The GTase then transfers the GMP molecule (Gp) to the 5’-diphosphate RNA to create GpppNp-RNA. In the final step, (guanine-N7)-methyltransferase (N7MTase) transfers the methyl group from 5-adenosyl-L-methionine (AdoMet) to the cap guanine to form the cap-0 structure, 7-methyl-GpppNp (m⁷GpppNp), and releases 5-adenosyl-L-homocysteine (AdoHcy) as a by-product. The capping reaction is completed by methylation of the ribose-2’-O position of the first nucleotide by the AdoMet-dependent (nucleoside-2’-O)-methyltransferase (2’OMTase), generating the cap-1 structure (m⁷GpppNm,Rp). The box contains examples of viruses that acquire their cap structures using the cellular capping machinery or encode their own viral capping machineries that adopt the canonical pathway. Question marks indicate viruses that are likely to follow this conventional pathway. The RNAs capped by viral enzymes are indistinguishable from cellular mRNA and can thus be translated into proteins by the cellular ribosomal machinery. BTV, bluetongue virus; HBV, hepatitis B virus; HCV, hepatitis C virus; SARS CoV, severe acute respiratory syndrome coronavirus.

Following translation, the lifespan of an mRNA molecule is controlled by two main processes in eukaryotic cells: first, the removal of its poly(A) tail and subsequent 3’-to-5’ exonucleolytic degradation, and second, an mRNA-decapping step that allows 5’-to-3’ exonucleolytic degradation (see [REFS 17,18] for a review). Interestingly, decapped RNA may apparently be re-capped by the combined action of an as-yet-unknown 5’-monophosphate kinase[2] interacting with a host cell GTase that is also present in minor amounts in the cytoplasm (see [REF 20] for a review). Cytoplasmic decapping is catalysed by DCPS, a member of the NUDIX hydrolase superfamily, and stimulated by decapping enhancer proteins. Following RNA decapping in P-bodies, transcripts are quickly degraded by 5’-to-3’ exonucleases such as XRN1. Thus, 5’-triphosphate mRNAs are almost completely absent in the cytoplasm. The eukaryotic cell has consequently evolved mechanisms to sense triphosphate RNA as non-self and uses these RNA species to trigger an innate immune response[37]. Viruses are the most common cell invaders to produce cytoplasmic mRNAs and have therefore been under a selective pressure to evolve...
Box 1 | Getting around the lack of capping

Some viruses (such as single-stranded positive-sense RNA (ss(+))RNA viruses from the families Picornaviridae, Caliciviridae and Astroviridae) do not have a cap structure at the 5′ end of their mRNAs or genomic RNAs. Rather, they covalently attach to the 5′ RNA end a protein termed VPg and/or carry an internal ribosome entry site (IRES) structure in the 5′ untranslated region. IRESs have now been found in many different cellular and viral RNAs, including those of ss(+))RNA viruses from the families Picornaviridae and Dicistroviridae, the genus Lentivirus and the Flaviviridae genera Hepacivirus and Pestivirus. In pestiviruses, genomic RNA remains in a 5′-triphosphate form and thus promotes high levels of expression of host interferon-stimulated genes. However, these viruses also trigger several pathways that limit the antiviral response, resulting in a competitive advantage for IRES-dependent translation of viral genes.

countering mechanisms to conceal their RNA 5′ ends from the innate immunity machineries of the host cell.

Some viruses, such as those in the family Picornaviridae (single-stranded positive-sense RNA viruses (ss(+))RNA viruses; for example, polioviruses and encephalomyocarditis virus (EMCV)) recruit the 43S pre-initiation complex in a 5′-cap-independent manner (BOX 1). Other viruses (for example, viruses of the family Caliciviridae, which are also ss(+)RNA viruses) covalently attach their RNA 5′ end to a VPg-like protein, which directly interacts with the cap-binding protein eIF4E and initiates translation of viral mRNAs. In members of the Picornaviridae, VPg may be lost before translation.

However, by far the most common viral mechanism for ensuring efficient translation of viral proteins and avoiding immune surveillance mechanisms is through the acquisition of a cap structure. The remarkable diversity of mechanisms that lead to an RNA cap structure identical to that of the host cell mRNAs is described below.

Conventional capping of viral RNA

Even when the viral replication cycle includes a nuclear phase, any virus entering a host cell must express its genes in the cytoplasm using the host translation machinery. Viral genomes can be made of single-stranded or double-stranded nucleic acids, either RNA or DNA, and the corresponding strategies used for protecting their RNA transcripts are outlined in FIG. 2.

Cap structures can be added to viral RNAs by one of the three following mechanisms. In the first mechanism, most viruses that synthesize their mRNA using cellular RNA pol II use the cellular capping machinery and the corresponding strategies used for protecting their RNA transcripts are outlined in FIG. 2.

**Figure 2 | RNA 5′ ends in the mammalian-virus world.** Mammalian viruses, with the exception of those from the single-stranded positive-sense RNA (ss(+))RNA virus genera Pestivirus and Hepacivirus, use strategies to chemically modify their mRNA 5′ ends through either covalent attachment of a protein (VPg for ss(+))RNA viruses from the families Picornaviridae, Caliciviridae and Astroviridae) or covalent attachment of an RNA cap structure (all other viruses). Arrows indicate the type of RNA 5′ end protection that is used by those viral groups, and the enzymes and mechanisms involved are indicated. Dashed arrows indicate a likely but incompletely demonstrated pathway. Viral and cellular proteins are distinguished by the prefixes v and c, respectively. Yellow shading highlights viral groups for which the life cycle includes a nuclear phase in the host cell. This list of viral taxa is non-exhaustive and used as an example only. ds, double-stranded; E, enzyme; EndoN, endonuclease; GTase, guanylyltransferase; IRES, internal ribosome entry site; m7G, 7-methyl; MTase, methyltransferase; NTPase, nucleotide 5′-triphosphatase; p, phosphate group; PRNTase, polyribonucleotidyl transferase; RTPase, RNA triphosphatase; ss(−)RNA, single-stranded negative-sense RNA.
RNA viruses

The RNA-capping mechanism of negative-sense RNA (−)RNA viruses such as those of the family Rhabdoviridae. The NTPase hydrolyses the γ-phosphate of GTP (Gppp) to yield a GDP (Gpp) and inorganic phosphate (P). Polyribonucleotidyl transferase (PRNTase) reacts with the nascent viral RNA (pppNp-RNA), in which N denotes the first transcribed nucleotide, releasing pyrophosphate (PP), and forming a covalent PRNTase–pNp-RNA intermediate. The PRNTase then transfers the RNA molecule to the GDP to create GpppNp-RNA. (Nucleoside-2′-O)-methyltransferase (2′OMTase) transfers the methyl group from S-adenosyl-L-methionine (AdoMet) to the first nucleotide of the RNA. The capping reaction is then completed by methylation of the cap by the AdoMet-dependent (guanine-N7)-methyltransferase (N7MTase). The box lists examples of viruses that acquire their cap structures using such a capping pathway. Question marks indicate viruses that are likely to follow this conventional pathway.

Significant viruses
- Rabies virus (Rhabdoviridae)
- VSV (Rhabdoviridae)
- Measles virus (Paramyxoviridae)
- Ebola viruses (Filoviridae)

Figure 3 | Unconventional capping pathways. a | The RNA-capping mechanism of negative-sense RNA (−)RNA viruses such as those of the family Rhabdoviridae. The PRNTase hydrolyses the γ-phosphate of GTP (Gppp) to yield a GDP (Gpp) and inorganic phosphate (P). Polyribonucleotidyl transferase (PRNTase) reacts with the nascent viral RNA (pppNp-RNA), in which N denotes the first transcribed nucleotide, releasing pyrophosphate (PP), and forming a covalent PRNTase–pNp-RNA intermediate. The PRNTase then transfers the RNA molecule to the GDP to create GpppNp-RNA. (Nucleoside-2′-O)-methyltransferase (2′OMTase) transfers the methyl group from S-adenosyl-L-methionine (AdoMet) to the first nucleotide of the RNA. The capping reaction is then completed by methylation of the cap by the AdoMet-dependent (guanine-N7)-methyltransferase (N7MTase). The box lists examples of viruses that acquire their cap structures using such a capping pathway. Question marks indicate viruses that are likely to follow this conventional pathway. b | The RNA-capping mechanism of positive-sense RNA (+)RNA viruses such as those of the family Alphaviridae. The RNA triphosphatase (RTase) hydrolyses the γ-phosphate of the viral RNA to yield a diphosphate RNA (ppNp-RNA) and P. A GTP molecule is methylated at its N7 position by the AdoMet-dependent N7MTase. Guanylyltransferase (GTase) then binds the N7-methyl-GTP (m7GpppG), forming a covalent link with a catalytic histidine (m7Gp–GTase) and releasing PP. c | The RNA-capping mechanism of (−)RNA viruses such as those of the family Orthomyxoviridae; this mechanism is referred to as cap snatching. The PB2 subunit of the viral RNA-dependent RNA polymerase (RdRp) binds to the 5′ end of cellular capped mRNAs (which are enriched in the processing (P)-bodies), and the PA subunit then releases short capped RNAs by using its endonuclease (EndoN) activity. These capped RNAs are used as primers by the viral RdRp in viral transcription to generate viral mRNA using the viral (−)RNA as a template. The RdRp then synthesizes the complementary negative-sense strand. The box provides example of viruses that acquire their cap structures using a similar capping pathway. Note that most of the mRNAs that are capped by viral enzymes are indistinguishable from cellular mRNAs and can be translated into proteins by the cellular ribosomal machinery. AdoHcy, S-adenosyl-L-homocysteine; HEV, hepatitis E virus; RVFV, Rift Valley fever virus; VSV, vesicular stomatitis virus.

Single-stranded negative-sense RNA viruses (ss(−)RNA viruses). Viruses that have or produce mRNAs that are complementary to their genomic RNA.

Ambisense RNA viruses
Viruses (such as members of the families Arenaviridae and Bunyaviridae) that have or produce both mRNAs that are co-linear to and mRNAs that are complementary to their genomic RNA, although most mRNAs are complementary in polarity.

A second strategy consists of acquiring cap structures from cellular mRNAs by ‘cap snatching’ (FIGS 2, 3). RNA viruses such as members of the families Orthomyxoviridae, Arenaviridae and Bunyaviridae — which are ss(−)RNA viruses, with the latter two families also referred to as ambisense RNA viruses — steal short, capped cellular mRNAs through endonucleolytic cleavage. The stolen short, capped mRNAs are then used by the viral polymerase to prime synthesis of viral RNA.
For the third method, many viruses encode their own capping machinery and have evolved a diverse set of dedicated enzymes and mechanisms to carry out capping. Accordingly, most viruses with an ssRNA genome synthesize or acquire the RNA cap using their own set of enzymes (FIG. 2). Within this diversity, the capping of viral mRNA can be classified as either ‘conventional’, when it follows the mRNA-capping pathways used by eukaryotes and DNA viruses (that is, the sequential action of RTase, GTase and MTases) (FIGS 1b, 2), or ‘unconventional’, when it does not (see below and FIG. 3).

The best-characterized conventional RNA-capping system is that exemplified by the dsDNA virus vaccinia virus, which expresses a multifunctional mRNA cap-synthesizing enzyme (D1) containing RTase, GTase and N7MTase domains24. The 5′-diphosphate of the nascent mRNA is first hydrolysed by the RTase to yield 5′-diphosphate RNA, which is then sequentially transferred to other internal domains23,24 to be capped and methylated, the latter reaction with allosteric stimulation through direct association with viral D1 protein25,26. Cap assembly is completed by the viral Vp39, a bifunctional protein that catalyses the 2′-O-methylation of the cap-0 structure and also acts as an elongation factor for poly(A) polymerase27. Viruses from the dsDNA virus family Reoviridae share the same pathway as vaccinia virus. The enzymes remain physically associated in an inner capsid (or ‘transcriptionally active core’), which constitutes a molecular machine or ‘assembly line’ that is able to transcribe the genome, synthesize the cap and inject the resulting mRNA into the cytoplasm of the host cell for translation. In addition, further assignment and characterization of the GTase activity in viruses of the genus Flaviviridae and those of the family Coronaviridae may join these viruses, which are ss(+)RNA viruses, to the conventional RNA-capping pathway group.

Conventional viral RNA-cap-synthesizing enzymes

RNA tripolyphosphates. RTases catalyse the cleavage of the interphosphate bond between the β-phosphate and the γ-phosphate of 5′-triphosphorylated mRNA (FIG. 1b). A range of enzyme structures exist (FIG. 4a), indicating that during evolution several independent solutions evolved for this initial step of the capping reaction, which is often found in association with RNA-binding and strand separation activities. Most RTases are also able to hydrolyse NTPs28–30.

RTases can be categorized into four groups, further defined by the metal dependency of the hydrolytic reaction mechanism that they use to cleave the interphosphate bond. Metal-dependent RTases can be organized into three structural groups: the histidine triad (HIT)-like fold (α–β complex), the tripolyphosphate metalloenzyme (TTM) and the viral RNA helicase-like fold (FIG. 4a).

The HIT-like fold is found in NSP2, from the dsRNA rotaviruses, and so far it is the only RTase identified with such a fold. It has an amino-terminal helical domain and a carboxyl terminus with an α–β fold that resembles the ubiquitous cellular HIT group of nucleotidyl hydrolas31. The nucleotide-binding site is located in the cleft between the two domains, which contains a histidine residue involved in binding and in Mg2+-dependent hydrolysis of both NTP and RNA substrates. NSP2 self-assembles into a doughnut-shaped octamer, a quaternary structural organization that creates several RNA-binding sites, which are presumably needed to destabilize RNA duplexes during genome replication and packaging30,32,33.

TTM enzymes hydrolyse NTPs in the presence of Mn2+ or Co2+ and are found in fungi and protozoa, and in most DNA viruses that encode an RTase, including poxviruses, chlorella virus, baculoviruses and mimiviruses34–37. The RTase Cet1 from Saccharomyces cerevisiae is nearly identical to that of mimiviruses and serves as a paradigm for the TTM group. Its structure features a characteristic tunnel lined by eight antiparallel β strands (FIG. 4a) and a double glutamate motif36–38. The mRNA supposedly sits in the tunnel, which harbours several charged and hydrophilic side chains that coordinate Mn2+ and SO42− in the crystal structure. SO42− is thought to indicate the position of the γ-phosphate of mRNA.

The third group comprises proteins with NTase and helicase activity, belonging to the large helicase superfamilies SF1 and SF2 (REFS 39–41). Such NTase–helicase family members are found in flaviviruses (for example, NS3 proteins of dengue virus, yellow fever virus, West Nile virus and Japanese encephalitis virus), coronaviruses (for example, nsP3 of severe acute respiratory syndrome coronavirus (SARS CoV), although for this member of the SF1 helicases the crystal structure is not known), alphaviruses (for example, the protein nsP2) and potexviruses (for example, protein 1A of bamboo mosaic virus), all of which are ss(+)RNA viruses, and also in viruses of the family Reoviridae (for example, the protein λ1), which are dsRNA viruses. The fold adopted by SF2 helicases features two RecA-like subdomains between which a cleft accommodates the nucleotide or 5′-triphosphate RNA substrate35. The RecA subdomains I and II carry the Walker A and B motifs. Residues belonging to the Walker A motif (also named motif I) form the P-loop, which stabilizes the terminal phosphate moiety of the substrate, whereas acidic residues from the Walker B motif (also named motif II or DEEDX box) coordinate the Mg2+ needed for hydrolysis42. Structural and biochemical studies revealed that the RTase and NTase activities of helicase enzymes from flaviviruses (ss(+)RNA) have a common catalytic site39-41. Similar biochemical studies have mapped an associated helicase–RTase activity in other viral families, including alphaviruses (ss(+)RNA)43. However, structural data are needed to better understand how these multifunctional viral proteins coordinate their various activities.

Metal-independent RTPases are found in plants, metazoas and viruses such as baculoviruses (dsDNA)46,48. They belong to the cysteine phosphatases superfamily. The RTPase reaction proceeds in two steps, starting with entry of the 5′-triphosphate RNA into the active-site tunnel and formation of a covalent cysteinyl-S-phosphoester adduct. The catalytic cysteine responsible for the nucleophilic attack on the γ-phosphate of the RNA 5′ end belongs to a P-loop motif (HCXXXXXTR(T/S)). The second step releases inorganic phosphate. Cysteine

© 2011 Macmillan Publishers Limited. All rights reserved
phosphatases adopt a characteristic α–β fold with a central twisted, five-stranded parallel β-sheet flanked by six α-helices. The catalytic cysteine that specifically recognizes the γ-phosphate of the RNA 5′ end resides at the bottom of the substrate-binding pocket, whereas other conserved residues from the P-loop, together with surrounding residues, form a positively charged channel that can accommodate the α-phosphate and the β-phosphate of the RNA 5′ end. The shape of the binding pocket dictates the selectivity for triphosphate RNA and seems too deep to grant diphosphate RNA access to the active site.
Despite the limited overall amino acid sequence identity in the large family of AdoMet-dependent MTases, most of these enzymes share a common structural core made of seven α-helices flanked by three α-helices on each side of the sheet, similar to the core found in the catalyse-O-MTase (a class I AdoMet-dependent MTase (70,71). This core catalytic domain has evolved extensions that consist of structurally non-conserved domains and allow these MTases to accommodate a range of methyl acceptors (9). The MTases exist either as isolated proteins (for example, nsp16 viruses from the Coronaviridae, ss(+)RNA), or as domains of larger proteins (such as NS5 of viruses from the ss(+)RNA virus genus Flaviviridae or the multidomain cap assembly lines of dsRNA viruses from the Reoviridae). In some instances the same protein domain can have dual N7MTase and 2′OMTase activities (for example, the MTase domain in NS5 from flaviviruses (42,43), sharing the same cofactor-binding site. When this is the case, repositioning of the RNA substrate must occur.

Both N7MTases and 2′OMTases share the class I family fold (Fig. 4d). N7-methyl transfer is thought to be promoted by optimal positioning of the reacting groups, mediated by several aromatic residues, and by an electrostatic environment that is favourable for the reaction (25,34). By contrast, 2′OMTases rely on a conserved catalytic tetrad, Lys-Asp-Lys-Glu (69,72,74). The catalytic reaction was deciphered by several structural studies of the 2′OMTase VP39 of vaccinia virus (79,80). It was suggested that residues in the vicinity of the catalytic Lys175 (the second lysine in the Lys-Asp-Lys-Glu motif) decrease the pKₐ value of its ε-amino group. This orients the 2′-hydroxyl group of the ribose for a nucleophilic in-line sn-2 attack on the AdoMet methyl group (81).

Guanylyltransferases. GTases of DNA viruses contain two domains: a nucleotidytransferase (NTase) domain that is conserved in capping enzymes, RNA ligases and DNA ligases, and a C-terminal oligonucleotide-binding domain that is observed in capping enzymes and several DNA ligases. Sequence alignments aided by structural information for several family members identified conserved residues and motifs both in the nucleotide-binding site and in the NTase site (9,52). A lysine-containing motif, KXXDG(U/L), is conserved among the GTases encoded by several DNA viruses (vaccinia virus, Shope fibroma virus and African swine fever virus) and the yeasts S. cerevisiae and Schizosaccharomyces pombe. The lysine in this motif was shown to be the catalytic residue in the GTase of vaccinia virus (46,53) and in the yeast capping enzyme (26). Moreover, this motif is conserved in the active site of polyribonucleotide ligases, which, like capping enzymes, catalyse an enzymatic reaction via the formation of a covalent Lys–NMP intermediate (26).

The structure of the GTases from Paramacium bursaria chlorella virus 1 (dsDNA), humans (48) and the yeast Candida albicans (the protein Cgt1) have been solved individually and, for the chlorella virus GTase and Cig1, in complexes with cap analogues, GTP or as a covalent GTase–GMP reopening to bind a pre-mRNA molecule (model derived from the structure in PDB accession code 1CKN; the intermediate GTase–GMP reopening to bind a pre-mRNA molecule (model derived from the structure in PDB accession code 1YDN9) is the only viral structure available so far. Structures are coloured cyan for α-helices and pink for β-strands.

The dsRNA viruses of the family Reoviridae use one multifunctional capsid protein for the capping reaction. λ2 of the dsRNA virus mammalian orthoreovirus (62), VP4 of bluetongue virus (in the genus Orbivirus) (63) and VP3 from cytoplastic polyhedrosis virus (in the genus Cypovirus) (64) act as assembly lines, such that the RNA substrate is shuttled from one domain to the next (see below). The GTase domains of these proteins feature different folds (see below and Fig. 4c) to DNA virus GTases. Despite these structural differences, viruses from the genera Orthoreovirus and Aquareovirus use a conserved lysine residue (although not part of the signature sequence found in DNA viruses) to form a covalent intermediate with GMP (65,66). In flaviviruses (ss(+)-RNA), the 2′OMTase domain has been proposed to act as a GTase (67,68). However, the proposed catalytic lysine is not conserved in flaviviruses (69).

Methyltransferases. Despite the limited overall amino acid sequence identity in the large family of AdoMet-dependent MTases, most of these enzymes share a common structural core made of seven α-helices flanked by three α-helices on each side of the sheet, similar to the core found in the catalyse-O-MTase (a class I AdoMet-dependent MTase (70,71). This core catalytic domain has evolved extensions that consist of structurally non-conserved domains and allow these MTases to accommodate a range of methyl acceptors (9). The MTases exist either as isolated proteins (for example, nsp16 viruses from the Coronaviridae, ss(+)RNA), or as domains of larger proteins (such as NS5 of viruses from the ss(+)RNA virus genus Flaviviridae or the multidomain cap assembly lines of dsRNA viruses from the Reoviridae (42)). In some instances the same protein domain can have dual N7MTase and 2′OMTase activities (for example, the MTase domain in NS5 from flaviviruses (42,43), sharing the same cofactor-binding site. When this is the case, repositioning of the RNA substrate must occur.

Both N7MTases and 2′OMTases share the class I family fold (Fig. 4d). N7-methyl transfer is thought to be promoted by optimal positioning of the reacting groups, mediated by several aromatic residues, and by an electrostatic environment that is favourable for the reaction (25,34). By contrast, 2′OMTases rely on a conserved catalytic tetrad, Lys-Asp-Lys-Glu (69,72,74). The catalytic reaction was deciphered by several structural studies of the 2′OMTase VP39 of vaccinia virus (79,80). It was suggested that residues in the vicinity of the catalytic Lys175 (the second lysine in the Lys-Asp-Lys-Glu motif) decrease the pKₐ value of its ε-amino group. This orients the 2′-hydroxyl group of the ribose for a nucleophilic in-line sn-2 attack on the AdoMet methyl group (81).
Using cap analogues, the molecular basis for recognition of GpppNp–RNA versus cap–0–RNA has also been characterized for VP39 (REFS 80, 82). The m’G is stacked between two aromatic residues, and electron delocalization and electrostatically enhanced stacking owing to N7 methylation favours the recognition of the cap over GTP (FIG. 4d).

One interesting aspect of MTase activity relates to its regulation: in three cases, interfacial activation is achieved either by a cofactor protein (such as the vaccinia virus D12 subunit, which enhances the N7MTase activity of D1 through an allosteric mechanism21,26, and the SARS CoV (ss(+))RNA metalloprotein nsp10, which acts as a cofactor to activate the 2’OMTase nsp16 but not the N7MTase nsp14 (REF. 83)) or by binding to lipid membranes (as is the case for nsp1 of the ss(+)RNA alphaviruses).

What determines the sequence of the two methylation steps? For flaviviruses78,79 or coronaviruses80,81, which are ss(+)RNA viruses, the order in which methylations are performed is not encoded in the global protein architecture. Rather, variations in kinetics and affinity may dictate the order in which reactions occur82,83. In the case of flaviviruses, RNA secondary structures also seem to be important: whereas the N7MTase activity of the bifunctional NS5 MTase domain requires a long substrate encompassing a specific stem loop RNA structure, christened stem loop A (SLA), the 2’OMTase activity is able to transfer a methyl group to short RNA acceptors84.

Cap assembly lines. Several dsRNA viruses encode structural proteins that are packaged with their genome in the viral particle and are able to perform the four reactions needed to synthesize a cap–1 structure, much like an assembly line (FIG. 4c). The key molecular components of the RNA-capping machinery in members of the dsRNA virus family Reoviridae are RNA-directed RNA polymerase (named VP1 in orbiviruses and rotaviruses, and λ3 in orthoreoviruses) and a multifunctional cap-synthesizing enzyme (named VP4 in orbiviruses, VP3 in rotaviruses and λ2 in orthoreoviruses). Both λ2 and VP4 are composed of four domains that were identified as RTPase, GTase, N7MTase and 2’OMTase, respectively62,63. The spatial arrangement of the different protein domains reflects the time sequence of the enzymatic reactions that are required for mRNA capping following synthesis by the viral polymerase (FIG. 4c). Although it is unclear how and when RTPase activity occurs, a complete pathway has been proposed in which guanylyl transfer occurs near to the base of the pentameric ‘turret’ (formed by λ2 in orthoreoviruses), followed by N7-methylation and 2’-O-methylation of the mRNA62. Therefore, in dsRNA viruses, the sequence of steps in the cap synthesis pathway should remain identical to the sequence in the capping pathway for cellular mRNAs and for DNA viruses such as vaccinia virus, for which the capping machinery is embedded in a multidomain protein complex.

The GTase domain and both MTase domains of bluetongue virus (a dsRNA virus in the genus Orbivirus) were unambiguously mapped on the VP4 structure, but the position of the RTPase domain remains uncertain63. However, it is believed that both the RTPase and GTase activities reside in the same C-terminal domain of VP4, a unique architecture that is reminiscent of, but distinct from, double-domain RTPase–GTase proteins found in metazoans and plants. The enzymatic activity requires Mg2+ (REFS 63, 86, 87). However, VP4 does not adopt a typical metal-dependent RTPase fold. A cysteine residue is found in a deep cavity similar to that harbouring the catalytic motif of the cysteine phosphatase superfamily63. Thus, these assembly line enzymes seem to have incorporated features from various phylogenetic origins.

Unconventional cap synthesis pathways

The first indication that there are deviations from the conventional RNA-capping pathway for viral mRNAs came in the early 1970s, around the time of the discovery of the RNA cap structure. Since then, it has been demonstrated that the ss(–)RNA virus vesicular stomatitis virus (VSV) and ss(+)RNA alphaviruses (from the family Togaviridae) can synthesize a viral RNA cap that is identical to a cellular RNA cap, albeit constructed through a completely different mechanism. Although alphaviruses do not proceed further than synthesizing a cap–0 structure, the fact that divergent biosynthetic pathways converge to the consensus cap structure indicates that the selective pressure to maintain this structure must be high.

The Mononegavirales RNA-capping pathway.

Mononegavirales is a viral order of ss(–)RNA viruses with unsegmented genomes, such as VSV and rabies virus (in the family Rhabdoviridae), measles virus (from the family Paramyxoviridae), bornavirus (from the family Bornaviridae), and Ebola viruses and Marburg viruses (from the family Filoviridae). These viruses encode a multifunctional L protein that carries RNA-dependent RNA polymerase (RdRp) and RNA cap synthesis activities. These enzymes have evolved independently from other known eukaryotic cap-synthesizing enzymes, and the L proteins of VSV88,89, siringaemia of carp virus90, human respiratory syncytial virus91 and Chandipura virus92 transfer GDP rather than GMP to the RNA 5’ end. Part of a domain in the conserved region V of L protein contains the GDP polynucleotidyl transferase (PRNTase) activity, and forms a cognate enzyme–pNp–RNA intermediate (FIG. 5a) with the nascent viral RNA. The cognate bond with RNA involves a conserved histidine residue present in an ‘HR’ motif instead of the lysine residue used by conventional GTases82. The 5’-monophosphorylated viral mRNA start sequence then receives GDP generated from GTP by an as-yet-unknown NTPase. The VSV MTase, present in domain VI of L protein83,84, subsequently methylates the core cap structure at the ribose-2’-O position of the first nucleotide, followed by methylation at the guanine-N7 position, generating GpppAm2–6–9, RNA and m’GpppAm2–6–9, RNA, respectively95–97. The capping reaction seems to be dependent on RNA length, indicating a possible spatial rearrangement in L protein98. Although no crystal structure is available yet, the MTase activities...
Box 2 | The RNA-decapping pathway of viruses

Viruses cap and decap RNA, and many viruses regulate the decapping pathway in order to control the ratio of viral and cellular mRNAs.

Decapping of cellular mRNA by Saccharomyces cerevisiae L-A virus (from the Totiviridae family of double-stranded RNA viruses) proceeds through a decapping enzyme carried by the Gag subunit of the capsid; this Gag subunit is responsible for covalently binding cap structures (7-methyl-GpppG (m7GpppG), in which p is a phosphate group) of cellular mRNA10. The decapping activity of Gag aids in the expression of viral RNA, apparently by producing large amounts of cellular RNA decoys that inhibit the S. cerevisiae enzyme 5'-3' exoribonuclease 1 (Xrn1)19 and compete with degradation of viral RNA. How the viral mRNA is recruited by the eukaryotic translation initiation factor 4E (eIF4E) complex remains to be elucidated. In the case of the family Poxviridae (double-stranded DNA viruses), the decapping enzyme (D10) increases the turnover of host mRNAs and contributes to the shutdown of host protein expression18. Moreover, D10 seems to preferentially degrade m7GpppAmP0, rather than m7GpppApP0, and thereby hydrolyses early-phase viral RNA carrying predominantly m7GpppGmP0, cap structures18. In other words, this viral pathway benefits from having mRNAs (produced by the capping apparatus) that will be recruited by the eIF4E complex, and simultaneously removes the potential competition from cellular mRNAs for ribosome binding. Viral ‘cap-snatching’ (see main text) also results in this imbalance, favouring expression of viral genes. Finally, several viruses have been reported to interfere with the cellular RNA-trafficking and decoy machinery. First, a viral protein, such as DCP1A and the 3'-end deadenylase complex component PAN3. Recently, the ss(+)RNA viruses hepatitis C virus and HIV were also reported to be connected to the cellular decapping machinery and to regulate it (reviewed in REF. 162).

Cell-based replicon assays

Assays that allow one to follow the replication of a ‘minimal viral genome’ encoding the viral replication complex but no structural or envelope proteins, which are usually replaced by reporter genes (such as luciferase or chloramphenicol acetyl transferase genes).

Virus-mediated RNA cap snatching

Among the ss(–)RNA viruses, those of the families Arenaviridae, Bunyaviridae and Orthomyxoviridae have a segmented RNA genome and form the order tentatively named Multinegavirales (FIG. 2). These viruses do not have a cap-synthesizing machinery. However, they have evolved to steal caps from host mRNAs in order to prime their own viral replication, in a process known as cap snatching (REFS 105–107). Cap snatching involves three steps (FIG. 5c). First, the 5’-methylated cap-1 or cap-2 structure of a host mRNA is bound by a specific site in the viral RdRp (or possibly the N protein108). Then, endonucleolytic cleavage of the cellular mRNA occurs several nucleotides downstream from the cap structure. Finally, this short, capped RNA is used as a primer for the synthesis of viral mRNA by the RdRp. The sequence, length and structure of the mRNA 5’ end that comes with the cap varies from one virus to the other. Most sequences are 15–20 nucleotides long106,109–112, but arenaviruses, nairoviruses and thogotoviruses use shorter primers112–114. Following endonucleolytic cleavage, the ‘decapped’ cellular mRNAs (BOX 2) are targeted to the degradation machinery, resulting in the downregulation of cellular mRNAs.

Enzymes from the cap-snatching pathway

Cap snatching was found initially in influenza viruses (ss(–)RNA viruses of the family Orthomyxoviridae), which serve as a model system for the other two viral families known to use snatching (ss(–)RNA viruses of the families Bunyaviridae and Arenaviridae), although differences in the proteins involved and the lengths of snatched sequences are expected, as shown for Thogoto virus117. The influenza virus polymerase is made of three subunits: PA, PB1 and PB2. A cap-binding domain was found in the central region of the PB2 subunit118, and an endonuclease domain at the N terminus of the PA subunit119,120. The structures of both domains shed light on the molecular mechanisms leading to cap snatching (see below). The cap-binding domain has a novel fold, although the mode of m’G binding by aromatic stacking is similar to that used in other cap-binding proteins. By contrast, the endonuclease domain of PA has a fold that is characteristic of the two-metal-dependent PD(D/E)XXK nuclelease superfamily but has the peculiarity of a metal-ligating histidine residue in the active site, conferring Mn2+ specificity121 (FIG. 5).

In contrast to orthomyxoviruses, both arenaviruses and bunyaviruses have a single protein (L) carrying the polymerase and cap-snatching activities. Recent studies showed that a Mn2+-dependent endonuclease that is homologous to that of orthomyxoviruses exists at the N terminus of arenaviral and bunyaviral L protein122,123 (FIG. 5a). Mutational analysis and cell-based replicon assays demonstrated that viral nuclease activities are essential for cap-dependent transcription of viral mRNA124. These domains have a conserved architecture and mechanism, which suggests an evolutionary link between them despite their low sequence identity (FIG. 5b).

A preliminary electron microscopy study of L protein from Machupo virus (an arenavirus) showed a central
Influenza virus PB2

| Orthomyxoviridae | PA
|------------------|----------------|
|                  | Endonuclease   |

| Arenaviridae     | L
|------------------|----------------|
|                  | Endonuclease   |
|                  | Cap-binding domain |

| Bunyaviridae     | L
|------------------|----------------|
|                  | Endonuclease   |

| LCMV L           | Cap-binding domain |

| Influenza virus PA |
|--------------------|
| Endonuclease       |

| La Crosse virus L  |
|--------------------|
| Endonuclease       |

| Influenza virus PB2 |
|--------------------|
| Endonuclease       |

Figure 5 | Unconventional capping machineries. Endonucleases and cap-binding domains (a) Domains involved in the ‘cap-snatching’ mechanism. The organization of cap-snatching domains of influenza viruses (single-stranded negative-sense RNA (ss(–)RNA) viruses of the family Orthomyxoviridae), and corresponding domains of distantly related viruses of the families Arenaviridae and Bunyaviridae (also ss(–)RNA viruses). Influenza virus polymerase is composed of three proteins of multiple domains: PA, PB1 and PB2. PA and PB2 are involved in cap snatching. PA carries the endonuclease domain in its amino terminus, whereas PB2 has an inner domain responsible for cap binding. MAPPING OF THE DOMAIN ORGANIZATION FOR ARENAVIRUSES AND BUNYAVIRUSES IS LESS ADVANCED; IN ITS AMINO TERMINUS, WHEREAS PB2 HAS AN INNER DOMAIN RESPONSIBLE FOR CAP BINDING. MAPPING OF THE DOMAIN ORGANIZATION FOR ARENAVIRUSES AND BUNYAVIRUSES IS LESS ADVANCED; IN ITS AMINO TERMINUS, WHEREAS PB2 HAS AN INNER DOMAIN RESPONSIBLE FOR CAP BINDING. Mapping of the domain organization for arenaviruses and bunyaviruses is less advanced; only the endonuclease domain is mapped to the amino terminus of L protein. The cap-binding domain is not clearly identified, as it is thought to be in either L protein or nucleocapsid (N or NP) protein, depending on the virus. (b) Structures of the different endonuclease domains of viruses from the families Orthomyxoviridae (influenza viruses), Arenaviridae (lymphocytic choriomeningitis virus (LCMV)) and Bunyaviridae (La Crosse virus) (Protein Data Bank (PDB) accession codes 3HW4, 3J5B and 2K15, respectively), and of the cap-binding domain from an influenza virus (PDB accession code 2VGZ). Despite having no sequence similarities, the folds of these endonuclease domains are conserved, suggesting a convergent evolution. Structures are coloured cyan for α-helices and pink for β-strands.

Innate immunity and RNA capping

Mammalian cells have co-evolved with viruses and have developed several mechanisms to detect a viral infection (such as detecting uncapped or partially capped RNA, VPG–RNA, and so on) and induce an antiviral response in neighbouring cells. This innate immunity is based on a small number of receptors called pattern recognition receptors (PRRs), which discriminate self from non-self components. Non-self detection depends on the recognition of a limited set of pathogen-associated molecular patterns (PAMPs), which are molecules or components that are characteristic of infectious agents, such as viral nucleic acids. The presence of non-self nucleic acids is detected through sensors such as Toll-like receptors (TLRs), which recognize DNA or RNA in intracellular compartments that do not usually contain these molecules. Moreover, several PRRs sense the presence of foreign nucleic acids directly in the cytoplasm — namely, the NOD-like receptors (NLRs) and the retinoic acid-inducible gene (RIG)-like receptors (RLRs). As host cell RNA is present in the cytosol, PRRs sense unprocessed RNA structures that are present in infected cells, as such dsRNA, RNA presenting a 5′-triphosphate, RNA with an incompletely methylated cap structure (cap-0, for mammalian standards) or RNA bearing a protein covalently attached to the 5′ end (such as VPG). The detection of PAMPs by PRRs triggers intracellular signalling events that mainly induce the production of type I interferon (IFN), interleukin-1 (IL-1) and pro-inflammatory cytokines, as well as the establishment of a cellular antiviral state in order to limit viral propagation.

Among the PRRs, the TLR family is the best studied innate immunity sensor family. TLRs are transmembrane proteins that are mainly expressed in immune cells, such as macrophages and dendritic cells. They are localized in endosomal compartments or at the cell surface. TLRs contain a leucine-rich repeat motif that recognizes PAMPs, and a Toll–IL-1 receptor (TIR) domain is present in the cytoplasmic part of the protein, ensuring signal transduction through TIR domain interaction with the TIR domains of cytoplasmic adaptor proteins such as MYD88 and TRIF (also known as TICAM1). TLR3, TLR7 and TLR8 are activated by different kinds of RNAs. TLR3 detects dsRNA, whereas TLR7 and TLR8 recognize ssRNA carrying a 5′-triphosphate or a cap-0 structure. In contrast to TLRs, the NLRs and RLRs are localized in the cell cytoplasm and detect the presence of intracellular invaders. Among the RLRs, RIG-I (also known as DDX58) and MDA5 (also known as IFIH1) seem to discriminate non-self RNA from self RNA on the basis of the RNA 5′ end. Metazoan self RNA presents a 5′-cap or cap-2 structure. RIG-I protein is specialized in the detection of 5′-triphosphate RNA, whereas MDA5 senses the presence of RNA with a cap-0 structure or linked to a protein such as VPG. RIG-I and MDA5 consist of two N-terminal caspase-recruitment domains (CARDs), a central DEXH box-containing RNA helicase–ATPase domain and a C-terminal regulatory domain (CTD). It is likely that, in the absence of RNA, RIG-I and MDA5 are incapable of triggering an antiviral response.
of ligands, the CARDs of RIG-I are auto-inhibited by other domains of the protein. For RIG-like proteins, nucleic acid binding to the RNA-binding site of the CTD induces a conformational change resulting in interaction of the CTD with the signalling adaptor molecule IFNβ promoter stimulator 1 (IPS1; also known as CARDIF, MAVS or VISA). IPS1 recruits a signalling complex in order to activate transcription factors such as interferon regulatory factor 3 (IRF3) and nuclear factor-kB (NF-kB), leading to the expression of IFNβ and other proteins that drive the antiviral response (Fig. 6).

The molecular basis for RNA recognition with and without 5′-end modification and/or overhanging nucleotides was analysed for RIG-I and MDA5. Initial studies indicated that RIG-I specifically recognizes 5′-triphosphate-containing ssRNA. It was later found that RIG-I requires base-paired structures in conjunction with a 5′-triphosphate to trigger an antiviral response. The molecular basis of the specific interaction between the RIG-I CTD and 5′-triphosphate dsRNA was deciphered by a crystallographical study of the CTD in complex with RNA. This study revealed
that 5′-triphosphate dsRNA binds to the CTD of RIG-I more tightly than its single-stranded counterpart. The 5′-triphosphate is sequestered in a lysine-rich cleft of the CTD, with a phenylalanine residue stacked to the terminal base pair. Interestingly, 5′-triphosphate dsRNA methylated at the 2′-O position of its first or second nucleotide is expected to create a steric conflict with the CTD of RIG-I. Accordingly, it does not stimulate the RIG-I pathway.

In contrast to RIG-I, MDA5 is thought to recognize either a viral RNA 5′ end carrying structures that are distinct from a 5′-triphosphate or longer, structured (mesh) RNA that would be generated during the viral life cycle. Indeed, MDA5 was reported to sense both dsRNA and ssRNA bearing a 5′ cap-0 [REFS 138,143] or linked to VPg. Structural analysis of the MDA5 CTD indicated that its global fold is similar to that of the RIG-I CTD[144], with amino acid differences in the domains involved in the recognition of the RNA 5′ end.

The involvement of MDA5 and TLR7 in the detection of cap-0-containing RNA was recently characterized in the antiviral response observed for a coronavirus mutant lacking 2′OMTase activity[145]. The replication of this mutant virus was dramatically impaired in infected mice. However, the replication of this virus was restored in MDA5−/−, TLR7−/− or type 1 IFN receptor (IFNIR)-deficient mice. It has also been suggested that other interferon-stimulated genes (ISGs) restrict the replication of 2′OMTase-deficient viruses[146]. Accordingly, IFIT1 (IFN IFN-α/β receptor) induced protein with tetratricopeptide repeats 1; also known as IFI56) and IFIT2 (also known as IFI54) were reported to limit the replication of West Nile virus (an ss(+))RNA virus of the genus Flavivirus), vaccinia virus (a dsDNA virus of the family Poxviridae) and murine hepatitis virus (an ss(+))RNA virus of the family Coronavirus) lacking 2′OMTase activity[147]. IFIT1 was recently shown to bind and sequester 5′-triphosphate RNA into a multiprotein complex containing IFIT2 and IFIT3 (also known as IFI60) in order to exert its antiviral effect[148]. Therefore, it is likely that 2′-O-methylation of the RNA cap promotes escape from the host innate antiviral response through avoidance of IFIT-mediated suppression.

Conclusions

Since the discovery of ‘blocked and methylated’ mRNA ends nearly 40 years ago, viruses have played an essential part in deciphering the process of mRNA capping, as well as its relationship with various cellular processes such as transcription, translation and innate immunity. Viral RNA capping is a field that still has a lot of uncharted territory: whether the RNA 5′ ends are protected or not is still unknown for many neglected viral families, and the GTase resists identification even for some studied human pathogens (for example, the ss(+) RNA viruses of the order Nidovirales).

It is likely that, during their co-evolution with their hosts, viruses evolved different adaptation strategies to protect their RNA transcripts. The diversity of mechanisms expressed in nature to add a cap to an RNA 5′ end is larger than that described here. Future research will also aim to elucidate how the fine-tuning between host-mediated decapping of viral RNA, virus-mediated capping of viral RNA and host innate immunity is performed. Moreover, it has been shown that viral antigens are not capped[149], an observation that has now been extended to many different viruses. Thus, ‘no capping’ signals probably exist. In addition, the abundance of template RNAs must certainly need to be finely regulated for optimal viral replication. The corresponding signals and regulations are largely unknown.

Finally, owing to its spectacular mechanistic diversity, RNA capping is an attractive field for the design of antiviral drugs. Several molecules have been proposed to act directly or indirectly on viral RNA capping. Ribavirin is a broad-spectrum antiviral agent that is active against several viruses that add or snatch RNA caps, and its pleiotropic mechanism includes targeting the RNA-capping machinery[148,149]. So far, efforts to design MTase inhibitors have used the AdoMet and AdoHcy (5-adenosyl-1-homocysteine) backbone to synthesize analogues that are specific to viral enzymes[150]. The increasing knowledge about active-site differences between cellular and viral MTases is expected to provide antiviral selectivity. Last, inhibitors of cap-snatching endonucleases have long been known[151,152,153]. Recently published crystal structures[19,122,123] of their targets should inform antiviral-drug design projects.
activity.

HIT-like fold. Nature replication and packaging exhibits a HIT-like fold. B. V. V. Rotavirus protein involved in genome Taraporewala, Z. F. & Patton, J. T. Histidine triad-like Vasquez-Del Carpio, R., Gonzalez-Nilo, F. D., Riadi, G., site. triphosphate phosphohydrolase activities demonstrate vaccinia virus RNA 5′-phosphatase and nucleotide activities of vaccinia virus are polymerase stimulatory activities of vaccinia virus are mRNA (nucleoside 2′-triphosphatase, and RNA 5′-triphosphatase flanking the catalytic core regulates 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 to RNA. J. Biol. Chem. 270, 2447–2479 (1994).

The discovery and molecular basis of the activation of a viral mTase by a protein cofactor.

Scurrure, B. S., Gershon, P. D. & Moss, B. Cap-specific mRNA (nucleoside-O2′)-methyltransferase and polyA) polymerase stimulatory activities of vaccinia virus are mediated by a single protein. Proc. Natl. Acad. Sci. USA 89, 2897–2901 (1992).

Benarroch, D. et al. The RNA helicase, nucleotide S′-triphosphate, and RNA S′-triphosphate reporter assays of Dengue virus protein NS5 are Mg2+-dependent and require a functional Walker B motif in the helicase catalytic core. Virolology 328, 208–218 (2004).

Myette, J. R. & Niles, E. Characterization of the vaccinia virus RNA S′-triphosphatase and nucleotide triphosphate phosphohydrolase activities demonstrate that both activities are carried out at the same active site. J. Biol. Chem. 271, 11945–11952 (1996).

Vaccinia virus: Niso, F. D., Riazi, G., Taraporewala, Z. F. & Patton, J. T. Histidine triad-like motif of the rotavirus NS2 octamer mediates both RTase and RTase activities. J. Mol. Biol. 362, 559–564 (2006).

Jayaram, H., Taraporewala, Z., Patton, J. T. & Prasad, B. V. V. Rotavirus protein involved in genome replication exhibits a HIT-like fold. Nature 417, 311–315 (2002).

The first crystal structure of an RTase with a HIT-like fold.

Taraporewala, Z., Chen, D. & Patton, J. T. Multimers formed by the rotovirus nonstructural NS2B bind to RNA and have nucleoside triphosphatase activity. J. Virol. 73, 9534–9549 (1999).

Taraporewala, Z. F. & Patton, J. T. Identification and characterization of the helix-stabilizing activity of rotavirus nonstructural protein NS2B. J. Virol. 75, 4519–4527 (2001).

Benarroch, D., Smith, P. & Shuman, S. Characterization of functional murine mRNA capping enzyme and crystal structure of the RNA triphosphatase domain. Structure 16, 501–512 (2008).

Gu, M. & Lima, C. D. Processing the message: structural insights into capping and decapping mRNA molecules of West Nile virus 5′-phosphatase and RNA 5′-phosphatase. Proc. Natl. Acad. Sci. USA 94, 9573–9578 (1997).

Cong, P. & Shuman, S. Covalent catalysis in nucleotidyl transfer: an H3K14-specific catalytic core is conserved at the active sites of RNA and DNA ligases. J. Biol. Chem. 268, 7726–7730 (1993).

Niles, E. G., Lima, C. D. Identification of the vaccinia virus mRNA guanylyltransferase active site lysine. J. Biol. Chem. 266, 24866–24893 (1991).

Shuman, S. & Hurwitz, J. Mechanism of mRNA capping by vaccinia virus guanylyltransferase: characterization of an enzyme–guanylate intermediate. Proc. Natl. Acad. Sci. USA 78, 187–191 (1981).

Schwer, B. & Shuman, S. Mutational analysis of yeast mRNA capping enzyme. Proc. Natl. Acad. Sci. USA 91, 528–532 (1994).

Glindh, T. & Barnes, D. E. Mammalian RNA ligases. Annu. Rev. Biochem. 61, 251–281 (1992).

Chu, C. & Shuman, S. Structure and function of the guanylyltransferase domain of human mRNA capping enzyme. Proc. Natl. Acad. Sci. USA 108, 10104–10108 (2011).

Fabrega, C. & Shuman, S. & Lima, C. D. Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. Mol. Cell 15, 555–561 (2003).

Håkansson, K., Doherty, A. J., Shuman, S. & Wiegley, D. B. X-ray crystallography reveals a large conformational change during guanylyl transfer by mRNA capping enzymes. Nature Struct. Mol. Biol. 14, 545–555 (2007). A study that deciphers the GTase reaction, with corresponding crystal structure snapshots.

Håkansson, K., Doherty, A. J., Shuman, S. & Wiegley, D. B. X-ray crystallography reveals a large conformational change during guanylyl transfer by mRNA capping enzymes. Nature Struct. Mol. Biol. 14, 545–555 (2007). A study that deciphers the GTase reaction, with corresponding crystal structure snapshots.
A paper with polymerase protein.

The structural basis for cap binding of viral RNA is identified as a prime and real-time mechanism for the initiation of viral RNA synthesis. J. Virol. 78, 281–287 (2004).

The authors demonstrate that 2'-O-methylation provides a molecular signature for the distinction of self and non-self RNA-dependent cap-binding activities of RNA sensors. PLoS Pathog. 6, e1000873 (2010).

An article that casts light on how N protein binds preferentially to capped mRNAs, stores and protects these mRNAs in P-bodies, and potentially takes an active role in cap accumulation.

The authors show the sequence requirements for methylation and the chain of events that characterize the mechanism.

The discovery of the unconventional capping pathway of (~)RNA viruses.

The structural basis for cap binding of viral RNA is identified as a prime and real-time mechanism for the initiation of viral RNA synthesis. J. Virol. 78, 281–287 (2004).

The authors demonstrate that 2'-O-methylation provides a molecular signature for the distinction of self and non-self RNA-dependent cap-binding activities of RNA sensors. PLoS Pathog. 6, e1000873 (2010).

An article that casts light on how N protein binds preferentially to capped mRNAs, stores and protects these mRNAs in P-bodies, and potentially takes an active role in cap accumulation.

The authors show the sequence requirements for methylation and the chain of events that characterize the mechanism.

The discovery of the unconventional capping pathway of (~)RNA viruses.

The structural basis for cap binding of viral RNA is identified as a prime and real-time mechanism for the initiation of viral RNA synthesis. J. Virol. 78, 281–287 (2004).

The authors demonstrate that 2'-O-methylation provides a molecular signature for the distinction of self and non-self RNA-dependent cap-binding activities of RNA sensors. PLoS Pathog. 6, e1000873 (2010).

An article that casts light on how N protein binds preferentially to capped mRNAs, stores and protects these mRNAs in P-bodies, and potentially takes an active role in cap accumulation.

The authors show the sequence requirements for methylation and the chain of events that characterize the mechanism.

The discovery of the unconventional capping pathway of (~)RNA viruses.

The structural basis for cap binding of viral RNA is identified as a prime and real-time mechanism for the initiation of viral RNA synthesis. J. Virol. 78, 281–287 (2004).

The authors demonstrate that 2'-O-methylation provides a molecular signature for the distinction of self and non-self RNA-dependent cap-binding activities of RNA sensors. PLoS Pathog. 6, e1000873 (2010).

An article that casts light on how N protein binds preferentially to capped mRNAs, stores and protects these mRNAs in P-bodies, and potentially takes an active role in cap accumulation.
146. Pichlmair, A. et al. IFIT1 is an antiviral protein that recognizes 5’-triphosphate RNA. *Nature Immunol.* 12, 624–630 (2011).

147. Garcin, D. & Kolakofsky, D. A novel mechanism for the initiation of Tacaribe arenavirus genome replication. *J. Virol.* 64, 6196–6203 (1990).

148. Hong, Z. & Cameron, C. E. Pleiotropic mechanisms of ribavirin antiviral activities. *Prog. Drug Res.* 59, 41–69 (2002).

149. Magden, J., Kääriäinen, L. & Ahola, T. Inhibitors of virus replication: recent developments and prospects. *Appl. Microbiol. Biotechnol.* 66, 612–621 (2005).

150. Lim, S. P. et al. Small molecule inhibitors that selectively block dengue virus methyltransferase. *J. Biol. Chem.* 286, 6233–6240 (2011).

151. Kuzuhara, T. et al. Green tea catechins inhibit the endonuclease activity of influenza A virus polymerase. *Antimicrob. Agents Chemother.* 38, 2827–2837 (1994).

152. Balvay, L., Soto Rifo, R., Ricci, E. P. et al. Structural and functional diversity of viral IRESes. *Biochim. Biophys. Acta* 1789, 542–557 (2009).

153. Guidotti, L. G. & Chisari, F. V. Noncytolytic control of viral infections by the innate and adaptive immune response. *Annu. Rev. Immunol.* 19, 65–91 (2001).

154. Parrish, S. et al. Vaccinia virus D10 protein has mRNA decapping activity, providing a mechanism for control of host and viral gene expression. *Proc. Natl Acad. Sci. USA* 104, 2139–2144 (2007).

155. McLennan, A. G. Decapitation: poxvirus makes RNA lose its head. *Trends Biochem. Sci.* 32, 297–299 (2007).

Acknowledgements
The authors are in debt to all previous and current laboratory members, too numerous to name, for their contributions and involvement in the study of viral RNA replication and capping. Special thanks go to B. Selisko for her tireless dedication to and help with the scientific elaboration of this manuscript. This work was supported in part through funding by the Fondation pour la Recherche Médicale (Programme Aide aux équipes), the French Direction Générale de l’Armement (contrat 07co404), Infectiopôle-Sud and the European Union Seventh Framework Programme (FP7/2007–2013) through the project SILVER (Small inhibitor leads against emerging RNA viruses; grant agreement 260644).

Competing interests statement
The authors declare no competing financial interests.

DATABASES
Protein Data Bank: [http://www.cancer.gov/drugdictionary](http://www.cancer.gov/drugdictionary)

FURTHER INFORMATION
Bruno Canard’s homepage: [http://www.afmb.univ-mrs.fr/Bruno-Canard](http://www.afmb.univ-mrs.fr/Bruno-Canard)

PyMOL: [http://www.pymol.org](http://www.pymol.org)

All links are active in the online PDF.