X-ray structure and activities of an essential Mononegavirales L-protein domain

Guido C. Paesen1,*, Axelle Collet2,*, Corinne Sallamand3, Françoise Debart3, Jean-Jacques Vasseur3, Bruno Canard2, Etienne Decroly2,** & Jonathan M. Grimes1,4,**

The L protein of mononegaviruses harbours all catalytic activities for genome replication and transcription. It contains six conserved domains (CR-I to -VI; Fig. 1a). CR-III has been linked to polymerase and polyadenylation activity, CR-V to mRNA capping and CR-VI to cap methylation. However, how these activities are choreographed is poorly understood. Here we present the 2.2-Å X-ray structure and activities of CR-VI+/, a portion of human Metapneumovirus L consisting of CR-VI and the poorly conserved region at its C terminus, the + domain. The CR-VI domain has a methyltransferase fold, which besides the typical S-adenosylmethionine-binding site (SAMP) also contains a novel pocket (NSP) that can accommodate a nucleoside. CR-VI lacks an obvious cap-binding site, and the SAMP-adjoining site holding the nucleotides undergoing methylation (SUBP) is unusually narrow because of the overhanging + domain. CR-VI+ sequentially methylates caps at their 2′O and N7 positions, and also displays nucleotide triphosphatase activity.

DOI: 10.1038/ncomms9749

1 Division of Structural Biology, Wellcome Trust Centre for Human Genetics, Oxford OX3 7BN, UK. 2 AFMB, CNRS, Aix-Marseille University, UMR 7257, Case 925, 163 Avenue de Luminy, 13288 Marseille, France. 3 Department of Nucleic Acids, IBMM, UMR 5247, CNRS, Université Montpellier, ENSCM, Campus Triolet, Place E. Bataillon, 34095 Montpellier, France. 4 Diamond Light Source Limited, Harwell Science and Innovation Campus, Didcot OX11 0DE, UK. * These authors contributed equally to this work. ** These authors jointly supervised this work. Correspondence and requests for materials should be addressed to E.D. (email: etienne.decroly@afmb.univ-mrs.fr) or to J.M.G. (email: jonathan@strubi.ox.ac.uk).
he Mononegavirales order groups five families of monopartite, negative-strand RNA viruses many of which are highly pathogenic and/or contagious; the Filoviridae (of which Ebola virus is a representative), the Bornaviridae (Borna disease virus), the Nyamiviridae (midway virus), the Rhabdoviridae (rabies, vesicular stomatitis virus (VSV)) and the Paramyxoviridae (measles virus, human metapneumovirus (hMPV)). These viruses encode a large RNA polymerase (L) (usually >2,000 amino acids) that is crucial to viral replication (Fig. 1a). It has two distinct roles to replicate the RNA genome and to transcribe viral mRNA. As such it not only polymerizes RNA but also synthesizes fully methylated cap structures. Capping involves the co-transcriptional addition of a guanosine (G) to the first nucleotide (N1) of the nascent RNA chain via a 5′-5′ triphosphate bridge, resulting in a GpppN1- structure. Typically, this is followed by methylation of nitrogen 7 (Gm) of GpppN1-, and of the 2′-oxygen (2′O) of the N1 ribose (m7GpppN1m0). The cap protects mRNAs against 5′-3′-exonucleases and promotes RNA transport and translation, while 2′O-methylation prevents detection by cellular-immunity sensors.

In Rhabdoviridae, CR-V catalyses cap addition by means of an unconventional polyribonucleotyd-transferase (PRNTase) reaction where a conserved histidine in CR-V forms a covalent phosphoamide bond with the transcript, resulting in a CR-V-pRNA intermediate. The capped transcript is released after ligation of a Gpp to the pRNA. This mechanism differs from capping in eukaryotes and most other viruses, in which a guanylyltransferase (GTase) forms a phosphoamide bond with Gp, before transferring it to 5′ppRNA2. Paramyxoviridae also contain a PRNTase signature motif in their CR-V domains, suggesting they use the same capping strategy as Rhabdoviridae. In addition, however, paramyxovirus and filovirus L proteins contain a KxxxKxxG sequence (K-K-G motif) at their C termini, reminiscent of a signature motif for eukaryotic GTases, where one of the lysines forms the transient phosphoamide bond with Gp in the capping reaction3. A C-terminal domain of Rinderpest virus L (containing CR-VI and the downstream K-K-G motif) was shown to form such a bond4, leaving open the possibility that Paramyxoviridae use GTase activity for capping.

Cap methylation is catalysed by S-adenosylmethionine (SAM)-dependent methyltransferases (MTases), which position a SAM molecule next to the target atom on the RNA, enabling the direct transfer of a methyl group and converting SAM into S-adenosylhomocysteine (SAH). In 2′O-MTases, a conserved K-D-K-E tetrad potentiates methyl transfer5. As yet, the boundaries of CR-VI, the putative MTase of L, are not well

---

**Figure 1 | The structure of the MTase domain of hMPV L.** (a) Domain organization of hMPV L, with at its C terminus the 46.5 kDa CR-VI+ domain (residues 1,599-2,005), comprising CR-VI (green), which contains the K-D-K-E motif typical for 2′O-MTases, and the + domain (red), carrying the K-K-G motif (blue). Boundaries of CR-I to -V are approximate. CR-III contains the G-D-N-Q signature motif for polymerase (RdRP) activity, and CR-V contains the HR motif for PRNT activity. (b) Cartoon representation of the CR-VI+ crystal structure, from amino (N) to carboxy (C) terminus (no structure could be assigned to the first ~18 residues). The + domain is shown in red, with K1991 and K1995 of the K-K-G motif in stick format. The CR-VI (MTase) domain is coloured purple (β-strands) and green (helices and loops), except for β2, β3 and β4, (the loop regions C terminal of β-strands 1, 2 and 4 that form SAM-P, orange), and λ1650-1666, (which disengages itself from the main CR-VI-fold to interact with the + domain; yellow). Nomenclature of helices and strands follows that used for other MTases (c). The pale-blue sticks show the K-D-K-E motif. A Zn-ion (silver sphere) is co-ordinated by H1766, H1798, C1802 and C1805. (c) Schematic representation of the secondary structure of a prototypical SAM-dependent MTase (top) and of the hMPV CR-VI domain (bottom). Helices are in green, strands in light purple and coils in blue, except for β3, β4, (orange), and λ1650-1666, (yellow). CR-VI displays some deviations from the prototypical SAM-MTase fold, some of which it shares with other RNA-MTases, including the long N-terminal coil, a longer αD and an extra helix (αX) at the C terminus. αX is absent, whereas, atypical for viral MTases, αZ is fully formed. CR-VI, moreover, has an unusual β-sheet; it lacks β3, but this is compensated for by the addition, at the other end of the sheet, of a new strand (βO), which glues the start of the N-terminal coil to the main structure. Also unusual is the fragmentation of αZ (resulting in the small z'-helix).

---

© 2015 Macmillan Publishers Limited. All rights reserved.
defined, nor is it clear if CR-VI acts on its own or in conjunction with other domains, whether it mediates both N7- and 2′O-methylation and which of these methylation would take place first. Research into the activities of L and the mechanisms underpinning them has historically been hampered by a complete, order-wide absence of high-resolution structures. To redress this, we set out to study L domains of hMPV, a paramyxovirus of the Pneumovirinae subfamily closely related to respiratory syncytial virus (RSV). Like RSV, hMPV is highly contagious and causes respiratory tract disease\(^2\). As a part of this study, we expressed CR-VI +, a 406-residue fragment comprising CR-VI and the adjoining ‘+’ domain, the variable region carrying the K-K-G-motif, investigated its MTase activity and solved its crystal structure. Besides sequentially methylating the 2′O and N7 atoms of small capped RNAs, CR-VI + also 2′O methylates uncapped substrates and displays nucleotide triphosphatase (NTPase) activity. Both the CR-VI domain, which assumes a fairly standard MTase fold, and the K-K-G motif of the (mainly helical) + domain are required for the MTase reactions. Combined, the data provide new insights into the modification of the 5′-ends of transcripts emerging from the polymerase domain of L. This structural information on a mononegavirus L protein, and the new insights in the capping mechanism it provides, should spur the development of novel antiviral drugs against this important group of highly pathogenic viruses.

**Results**

**MTase activities.** *In vitro*, CR-VI + most effectively binds and methylates synthetic RNAs containing the conserved start sequence of hMPV transcripts, preferring a substrate length of nine nucleotides (Fig. 2). The methylation occurs at the GN7 and N7-2′O positions (Figs 2 and 3a), with 2′O-methylation preceding N7-methylation (Fig. 3a), an uncommon order of events also occurring in VSV\(^4\). CR-VI +, moreover, efficiently methylates uncapped RNAs with 5′-phosphate groups (especially pppRNA), primarily at the 2′O atom of N1 since almost no \(^2\)H-methyl

---

**Figure 2 | MTase activity of CR-VI +.** (a) The transfer of tritiated methyl groups from SAM molecules to a capped RNA substrate (GpppGGGACAAGU), containing the consensus start sequence of hMPV transcripts (in red), was monitored over time. The rather slow *in vitro* methylation suggests the reaction is impaired compared with the *in vivo* activity of intact L, regions of which may aid the methyl transfer (for example, by correctly positioning substrate RNAs to the MTase; see main text). The bars and error bars correspond to the mean values from three independent measurements and their s.d.’s, respectively. (b). Substrate specificity was determined as above, but using various synthetic RNA substrates, and allowing the reactions to proceed for 16 h. Substrates were compared with GpppGGGACAAGU (red and blue panels) and pppGGGACAAGU (green panel), for which the degree of methylation was set at 100% (#, marked bars). The red-shaded panel compares the degree of methylation of nine-nucleotide-long hMPV start sequences with different 5′-ends and methylation states (the lighter bars represent uncapped RNAs). The results indicate efficient methylation of RNAs that already carried a (cold) methyl atom of N1 since almost no \(^2\)H-methyl
were methylated beforehand at the K-K-G lysines K1991, K1992 and K1995 and phase for a better separation of green or red background, to indicate that the altered residue belongs to stage, and m

activities of CR-VI

also grouped according to whether they change the K-D-K-E tetrad, SAMP, SUBPo or NSP. The results highlight the importance of essential SUBP residues (such

rather than the core MTase fold, were also found to be essential

was solved to 2.2 Å resolution in space group P212121 with two

MTases9. Other residues, many of which belong to the

2

caps at the

0

D1779, K1817 and E1848) are essential for 2

methylation target in the

G

that observed with

G

and vaccinia virus 2

methylation of pppGGGACAAGU represents

a

Thin-layer chromatograms. Following CR-VI + -mediated methyl transfer from SAM onto GpppGGGACAAGU (in which G was [32P]-labelled), nucleotides 2–9 were removed by nuclease P1 digestion, and the caps were separated by thin-layer chromatography (TLC). The controls (GpppG, GpppGm, GpppG and GpppGm, from left to right) were obtained with the same substrate, using MTases that specifically methylate caps at the N7 or 2′O positions (human N7- and vaccinia virus 2′-O-MTase). The TLC experiment used 0.65 M LiCl as mobile phase, allowing a clear separation of GpppG and GpppGm (top). The caps on the TLC plate were subsequently further resolved, this time using 0.45 M (NH4)2SO4 as mobile phase for a better separation of GpppG and GpppGm (bottom). GpppGm appears first (after a 1 h incubation), GpppGm becomes prominent at a later stage, and GpppG was not observed, indicating that 2′-O-methylation of N1 precedes N7-methylation of G. (b) The effect of point mutations on the MTase activities of CR-VI +, measured after a 16 h incubation period (by means of a filter-binding assay, as in Fig. 2), using GpppGGGACAAGU substrates that were methylated beforehand at N7 of G or 2′O of N1, to specifically monitor 2′O or N7-MTase activities, respectively. Mutants are listed against a yellow, green or red background, to indicate that the altered residue belongs to λ1650–1666, the rest of the CR-VI domain, or the + domain, respectively. They are also grouped according to whether they change the K-D-K-E tetrad, SAMP, SUBPo or NSP. The results highlight the importance of essential SUBP residues (such as the K-K-G lysines K1991, K1992 and K1995 and λ1650–1666 residues H1659 and R1662) for both 2′O- and N7-methylation. All tetrad residues are crucial for 2′O-methylation, while D1779 in particular is important for N7-MTase activity. The bars and error bars correspond to the mean values from three independent measurements and their s.d.’s, respectively.

Figure 3 | N7- and 2′O-methylation. (a) Thin-layer chromatograms. Following CR-VI + -mediated methyl transfer from SAM onto GpppGGGACAAGU (in which G was [32P]-labelled), nucleotides 2–9 were removed by nuclease P1 digestion, and the caps were separated by thin-layer chromatography (TLC). The controls (GpppG, GpppGm, GpppG and GpppGm, from left to right) were obtained with the same substrate, using MTases that specifically methylate caps at the N7 or 2′O positions (human N7- and vaccinia virus 2′-O-MTase). The TLC experiment used 0.65 M LiCl as mobile phase, allowing a clear separation of GpppG and GpppGm (top). The caps on the TLC plate were subsequently further resolved, this time using 0.45 M (NH4)2SO4 as mobile phase for a better separation of GpppG and GpppGm (bottom). GpppGm appears first (after a 1 h incubation), GpppGm becomes prominent at a later stage, and GpppG was not observed, indicating that 2′-O-methylation of N1 precedes N7-methylation of G. (b) The effect of point mutations on the MTase activities of CR-VI +, measured after a 16 h incubation period (by means of a filter-binding assay, as in Fig. 2), using GpppGGGACAAGU substrates that were methylated beforehand at N7 of G or 2′O of N1, to specifically monitor 2′O or N7-MTase activities, respectively. Mutants are listed against a yellow, green or red background, to indicate that the altered residue belongs to λ1650–1666, the rest of the CR-VI domain, or the + domain, respectively. They are also grouped according to whether they change the K-D-K-E tetrad, SAMP, SUBPo or NSP. The results highlight the importance of essential SUBP residues (such as the K-K-G lysines K1991, K1992 and K1995 and λ1650–1666 residues H1659 and R1662) for both 2′O- and N7-methylation. All tetrad residues are crucial for 2′O-methylation, while D1779 in particular is important for N7-MTase activity. The bars and error bars correspond to the mean values from three independent measurements and their s.d.’s, respectively.

Structure of the + domain. The crystal structure of CR-VI + was solved to 2.2 Å resolution in space group P212121 with two molecules, disulphide-linked at residue C1877, in the asymmetric unit, assuming a ‘head-to-toe’ conformation (Supplementary Fig. 1). There are no significant differences between the molecules, which have a slightly twisted, bi-lobed shape, composed of two globular domains, the larger corresponding to CR-VI (residues ~1,616–1,883) and the smaller to the + domain (1,884–2,005; Fig. 1b). Although composed of only ~120 residues in Pneumovirinae, the size of the + domain varies greatly within the Mononegavirales order, reaching ~240 residues in Rabdoviridae. In hMPV, it consists of six α-helices (α1+6–1). Helix α+6 contains the K-K-G motif (Supplementary Fig. 2) and together with α+1 leans over the active cleft of CR-VI. Helix α+4, and to a lesser extent helices α+1, α+5 and α+6, packs down on λ1650–1666, the second half of a long loop (residues 1,635–1,666) that swerves around the CR-VI domain (Fig. 1b and Supplementary Fig. 3; ‘l’ is used throughout the paper to denote loops). The λ1650–1666 region, which contains a small helix (α-l), acts as a fulcrum allowing the + domain to pivot relative to the CR-VI domain. Helix α+3 varies in length, from 4½ turns (in most crystals) to 6 (in Protein Data Bank (PDB) 4UCY), and the loop between α+2 and α+3 is always disordered (Supplementary Fig. 3). In crystals of the monomeric C1877A mutant (space group P3121), this disordered, unstable
region is further enlarged, as helix α+2 completely unfolds and α+3 unwinds to 3½ turns.

Structure of the CR-VI domain. The CR-VI domain shares some peripheral characteristics with the 2O-MTases of SARS coronavirus (PDB: 2XYQ)10, vaccinia virus (1VP3) (ref. 11) and bluetongue virus (VP4-subunit; 2JHA)12, such as the long N-terminal loop and the position of helix αX. In its active core, however, it better resembles RrmJ-type flavivirus MTases (for example, 3EVF9; Supplementary Fig. 4). Most notably, hMPV and flaviviruses share an unusually long (~10 residues), flexible β3λ (that is, the loop immediately following the β2-strand), which forms the SAM-binding pocket (SAMP) along with loops β1λ and β5λ, shielding it from the solvent (Fig. 1b, c). In Wesselsbron (flavi-)virus, β3λ is found in closed or open conformations, either packing up against SAM (PDB: 3ELW) or exposing it to the solvent (3EMB)13, changes that may assist SAM uptake and/or SAH expulsion. In CR-VI+, β2λ similarly assumes alternate conformations (Fig. 4c,d). In the ‘closed’ form, the ligand’s ribose group is hydrogen bonded to D1725, and—via a water—to D1722. Loops β1λ and β5λ also show a degree of flexibility. β1λ residue E1697, conserved in paramyxov- and filoviruses, forms a hydrogen bond with the NH2 group of SAM and is essential for both MTase activities (Fig. 3b). However, in the absence of SAM, its side chain either moves into the sub-pocket that normally accommodates the NH2 group, or, more markedly, turns towards the solvent in the direction of β5λ, which in this case assumes the ‘open’ position (Fig. 4d). β5λ forms a side wall of SAMP, and together with αD and β5 also defines a deep, hydrophobic cavity not present in other MTases, termed NSP (or nucleoside-binding pocket). Although a role for NSP has yet to be determined, the pocket binds the adenosine moiety of SAM or ATP, soaked at high (25 mM) concentrations into CR-VI+ crystals, causing β4λ to impinge onto SAMP (Fig. 4c,d; when SAM is used for soaking, one SAM molecule occupies SAMP, and another binds to NSP). GTP was not observed in NSP, but this may reflect a lower soaking concentration (2.5 mM), due to GTP’s poor solubility. With the exception of E1781, the amino acids lining NSP are poorly conserved beyond the Pneumovirinae, and mutating key NSP residues barely affects the MTase activities (Fig. 3b). CR-VI, finally, contains a C-terminal Zn-finger, which is not conserved beyond the Pneumovirinae subfamily and links the small α-ξ′ to the rest of the structure (Fig. 1b).

**Figure 4 | CR-VI+ binding pockets.** The cartoon representations show the + domain in red and CR-VI in green, with λ1650–1666 in yellow. SAM (in SAMP), GTP (in SUBP) and adenosine (ADN; in NSP) are shown as sticks, with the C atoms coloured gold, slate and magenta, respectively. Hydrogens (in white) accentuate the methyl group of SAM. 2Fo-Fc electron density maps around the ligands are represented in grey mesh (contoured at 1σ). (a) The relative positions of the pockets in the protein. (b) Close-up of SUBP, which is defined by residues of the + domain (particularly the K-K-G motif), λ1650–1666, and the CR-VI domain. Residues involved in ligand binding are shown as sticks. GTP is fitted in different orientations into the density in the PDB 4UCZ structure (main figure, and top figure to the right, where the guanosine ring is turned 180°) and in the PDB 4UCJ structure (bottom right, where the ligand lays in the opposite direction), highlighting that the ligand can bind in different orientations within the spacious pocket. (c) SAMP and NSP containing a SAM and ADN ligand, respectively (PDB 4UCJ, in which SUBP is also occupied). Residues lining the pockets are shown as sticks. The loops delineating SAMP (β5λ, β3λ, β1λ and β4λ) and the β-strands they originate from are shown in magenta. The dashed yellow lines show putative hydrogen bonds. (d) Superposition of three other CR-VI+ structures onto that in c, highlighting the flexibility of β1λ (especially E1697), β3λ and β5λ. The structure in blue (PDB 4UCK) contains SAM, whereas those in yellow (4UCJ) and aquamarine (4UCU) have empty SAMPs (this suggests that there is no strict correlation between SAMP occupancy and the position of β3λ). NSP is empty in the three superposed structures, which apparently affects the position of their β4λ loops and especially of the R1785 side group, which closes the pocket when occupied. All overlaid structures have empty SUBPs.
During the preparation of this manuscript, a 3.8-Å structure of VSV L, obtained by electron cryo-microscopy, was published (11). The CR-VI (or MTase) part of VSV resembles that of hMPV, but apparently lacks a deep NSp pocket (PDB: 5a22, Fig. 3). The + domain (C-terminal domain) is more elaborate in VSV than in hMPV, containing extra regions N and C terminal of helix a 2. The helix itself, however, appears well conserved, both in length and in position. Although a K-K-G motif is not present in VSV, it does contain an arginine (R1038) strictly conserved among the Rhabdoviridae, which is structurally equivalent to hMPV’s K1995, the second lysine of K-K-G motif.

The absence of a classical cap-binding site. A common feature of MTases involved in cap methylation is a defined cap-binding pocket that binds G with high affinity, enabling subsequent, low-affinity interactions with the triphosphate bridge and the first few nucleotides, thus precluding methylation of uncapped RNAs (5,10). In CR-VI+, however, an open, solvent-exposed area is found where this pocket is normally located (Fig. 6a). Moreover, GpppG- or m7GpppG binding was not observed in co-crystallization or soaking experiments, suggesting that CR-VI+ has a weak affinity for G at best and that the cap is not required for substrate recognition. This is consistent with CR-VI+ binding capped and uncapped RNAs with similar strength and being able to 2′O-methylate uncapped RNAs (Fig. 2b). Strong binding would also prevent translocation of G into SUBp for N7-methylation. Although a high-affinity cap-binding site is clearly absent from CR-VI+, the low-affinity nucleoside binding to NSp and the convenient location of this pocket relative to SUBp suggest it could provide space for G without forming strong interactions (Supplementary Fig. 5).

The absence of a high-affinity cap-binding pocket appears partly compensated for by the narrowing of the groove that in related MTases accommodates the first few nucleotides of the transcripts, by the overhanging + domain (Figs 4 and 6a).

Figure 5 | Comparison of the CR-VI+ domains of hMPV and VSV. (a) Cartoon representations. The CR-VI domains are similar, and share the unusual, strand-0 containing β-sheet (purple), the rather large β2L, indicated by arrow (1) and the long N-terminal loop (2), which runs somewhat differently in VSV. The λ1650-1666 peptide on which the + domain rests (yellow) also has a homologue in VSV. The Zn-finger, however, is not conserved, and α-helices B and Z are not fragmented. Helix αE, an element of the standard MTase topology (Fig. 1c), is present in VSV (3), as a result of which NSp may have disappeared. αX is at a different location (4), and is preceded by an extra helix (αX′(5)). E865 is expected to belong to the K-D-E tetrad from sequence alignments, is buried in the structure and does not reach the surface of the catalytic pocket (6), and the position normally taken by the K-D-E glutamate is occupied by T1833. The + domain of VSV is tilted, compared with that of hMPV, and more elaborate. Helices α5+1, α5+2, α5+3, α5+5 and α5+6 are conserved, but the α5+1-α5+2 loop is replaced by an extra helix (α5′ Y(7)). α5+4 is absent, whereas α5+5 is enlarged and immediately follows α5+3 (8). The 2-residue loop connecting α5+5 and α5+6 in hMPV is replaced by a 34-residue coil carrying a small three-stranded β-sheet (9). Helix α5+6 seems best conserved between the two + domains, although a K-K-G motif is not present in VSV. However, R2038, which is strictly conserved in Rhabdoviridae, takes the place of K1995 (10) and alignment below). In VSV, the + domain is extended beyond α5+6 with a 65-residue, partly helical, but mainly unstructured polypeptide (in grey (11)). Colour scheme and labelling are as in Fig. 1b. (b) Alignment of αX+6-helices from Mononegavirales L proteins. K-K-G motif residues are highlighted in red; the arginine replacing the second lysine of the motif in Filoviridae and Rhabdoviridae is highlighted in magenta. Red letters indicate other (less strictly) conserved residues, except for the G that replaces the first lysine of the K-K-G motif in most Rhabdoviridae (blue).
particular, the site adjoining $\text{SAMP}$, which holds the nucleotides undergoing methylation, has become a more elaborate, but narrower and possibly therefore, higher-affinity substrate-binding pocket (termed $\text{SUBP}$) as it accommodates the nucleotide undergoing methylation. Consistently, electron density is found in $\text{SUBP}$ following soaking or co-crystallization with GTP, whereas in other MTases added GTP predominantly shows up in the cap-binding pocket. In particular, helix $\alpha + 6$ and the $+\text{domain}$-affiliated $\lambda_{1650-1666}$ help shape $\text{SUBP}$ through the side chains of K$_{1991}$ and K$_{1995}$ (of the K-K-G motif), and of H$_{1665}$ and R$_{1662}$, respectively (Fig. 4b). The marked decrease in MTase activity of mutants altered at these residues (Fig. 3b) illustrates the importance of $\text{SUBP}$ in correctly presenting the substrate nucleotides to SAM. Nevertheless, the pocket is too spacious for a single nucleotide, and the electron density in $\text{SUBP}$ from a number of soaked crystals suggests that bound GTP often assumes more than one orientation. In structures where GTP could be fitted with confidence, the guanosine moiety predominately interacts with $\lambda_{1650-1666}$ residues H$_{1659}$ and R$_{1662}$, and with K$_{1991}$ and (K-D-K-E residue) K$_{1673}$, which clamp the guanine (Fig. 4b). Unusually for cap-MTases, K$_{1673}$ is not part of $\alpha Z$, but instead resides on the small $Z^\prime (310)$-helix (Fig. 1b,c). Whether any of the observed positions of GTP reflects in vivo binding of N$_1$ (as part of a transcript) is unclear; in MTase–RNA complexes (PDB: 1AV6 (ref. 17), 4N49 (ref. 16)), N$_1$ is situated much closer to the K-D-K-E tetrad (Fig. 6a).

**Figure 6 | RNA-binding site comparisons.** (a). Comparison of the RNA-binding sites in vaccinia virus cap-MTase (PDB 1AV6) and CR-VI$+$+. The vaccinia virus MTase (white surface, left) has a narrow cap-binding pocket (in between the red arrows) and a large, open RNA-binding site (adjoining the SAH-containing $\text{SAMP}$). In CR-VI$+$+ (coloured surface, middle) the cap-binding pocket is not present, whereas the RNA-binding site is narrowed (into $\text{SUBP}$) by the $+\text{domain}$ overhang (in red). A structural superposition (obtained by aligning the K-D-K-E tetrads, right) shows that the GTP ligand in CR-VI$+$+ is situated at a considerably greater distance from the tetrad than the first transcribed nucleotide (N$_1$) in the vaccinia virus MTase–RNA complex (shown in light pink). The Z$^\prime$O atoms of the nucleotides are shown as transparent, red spheres. (b) $\text{SUBP}$ conservation within the Mononegavirales order. The surface presentation on the left shows the basic (blue) and acidic (red) charge distribution on CR-VI$+$+. The ligands are in yellow. The other cartoons show the surface of the hMPV CR-VI$+$+ domain in the same orientation, but in white. Residues that are conserved in the hRSV, Measles, Ebola or Rabies virus homologues of CR-VI$+$+ are coloured dark red (identical residues) or pink (similar residues), and cluster around $\text{SUBP}$ and $\text{SAMP}$.

**Figure 7 | NTPase activity.** Autoradiographs of urea–PAGE gels show CR-VI$+$+ mediated conversion of radiolabelled GTP to GDP and ATP to ADP, over time. The smaller autoradiograph shows the requirement of Mg$^{2+}$ for the reaction (allowed to proceed for 1h), and also shows efficient GDP generation by washed, dissolved CR-VI$+$+ crystals (CR-VI$+$+*), dispelling the possibility that the activity is due to contaminants. The diagram on the right (obtained by phosphorimage analysis following electrophoresis) further illustrates the requirement of Mg$^{2+}$ and shows the effect of the metal-ion chelator EDTA on the GTPase reaction. The bars and error bars correspond to the mean values from three independent measurements and their s.d.’s, respectively.
Mg\(^{2+}\), resulted in radioactive protein bands on denaturing SDS gels, the level of radioactivity was low, and was not diminished by acid treatment before SDS-polyacrylamide gel electrophoresis (SDS-PAGE), implying it is not due to phosphomimic bond formation. A second, strong argument against a GTase-based capping mechanism in hMPV (and in favour of a PRNTase-based one) is the fact that in the closely related RSV the cap is formed by Gpp ligation to pRNA\(^20\).

We observed that CR-VI + also displays NTPase activity, converting GTP into GDP and ATP into ADP (Fig. 7). GTase activity, which is required for PRNTase-based capping, was previously reported in Mononegavirales L (ref. 21), but as yet could not be linked to a specific domain within the protein. The reaction observed with CR-VI +, however, is quite slow, possibly because other parts of L, or other co-factors, are needed for full activity. In line with this, we were not able to identify key active-site residues. Using the mutants listed in Fig. 3b, the greatest reductions in GTase activity were obtained with E\(_{SAMP}\)(corresponding to the flexible residue at the bottom of SAMP; Fig. 4d) and D\(_{SUBP}\)(part of SUBP; Fig. 4b) to 54 (± 20) and 57 (± 16) % of the wild-type activity, respectively (t = 3). The deletion of the dipeptide G\(_{1645}\)K\(_{1646}\) from the long N-terminal site resulted in a somewhat more pronounced ~70% reduction (Supplementary Fig. 6a). The NTPase activity was confirmed using crystallized CR-VI + (Fig. 7).

**Discussion**

As a GDP is transferred onto a PRNTase-bound pRNA intermediate during cap synthesis in VSV\(^21\), the presence of NTPase activity in CR-VI + would suggest that this domain is involved in cap addition. It is unclear from the structure of VSV L how capping, cap methylation and RNA synthesis are coordinated\(^14\), but an involvement of CR-VI + in cap addition is consistent with the dynamic nature of the multi-domain polymerases of RNA viruses in general, exemplified by flu where the C-terminal two-thirds of PB2 has been shown to be extremely mobile\(^22,23\). The presentation of uncapped (but CR-V-linked) pRNA to CR-VI + would also explain why hMPV L does not require a high-affinity cap-binding site. The integrity of both the PRNTase and K-K-G motifs, and thus represents an attractive target for the structure-based design of (potentially broad-spectrum) antiviral compounds.

**Methods**

**Cloning and expression.** The sequence encoding CR-VI + was PCR amplified with primers that added a C-terminal 5GHHHHTTHT tag to the translation product, from a synthetic hMPV L gene (hMPV isolate 00-1, GenBank: AF371337.2), codon optimized for expression in HEK293T mammalian cells (following transfection using Lipofectamine 2000; Invitrogen), in BL21 Star (DE3) (Novagen) and in Sf\(_21\) (Insect) cells) following expression with (flashBACULTRA) baculovirus (Oxford Expression Systems) using Cellfection II (Invitrogen)\(^25\). Mutants were generated by PCR, using primers carrying the mutation (Supplementary Fig. 7). The CR-VI + DNA was used as a template for mutagenesis, except for the K\(_{1998}\)Q/K\(_{1999}\)Q and K\(_{1998}\)Q/K\(_{1999}\)Q/K\(_{2000}\)Q mutants, which were obtained using the K\(_{1998}\)Q DNA. For each mutation, the forward primer was combined with a vector-specific reverse primer (5’-AGTGGATTTTG TGGACCGAGG-3’), and in a second PCR, the reverse primer was used together with a vector-specific forward primer (5’-CCITTAATTACACCAACAC-3’). The two PCR products thus generated were digested with either BstQII or BstI (New England Biolabs), and the ligation products were PCR amplified using the vector-specific primers for insertion in the popIN-E vector.

**Selenomethionine incorporation.** Twenty-four hours after injection of an infected Sf\(_21\) suspension culture (27.8 °C, with an agitation speed of 1300 r.p.m.,0.25g), cells were collected (70g, 10 min, 22 °C) and resuspended in cytosine- and methionine-free Sf9 medium (Gibco) supplemented with dialysed fetal bovine serum (Gibco; 7% v/v) and 150mM l-cysteine (Sigma). Following an additional 4 h at 27.8 °C in the shaking incubator, 250mg l-1 selenomethionine (Sigma) was added. Protein expression was allowed to continue for another 48 h. Proteins were purified as above.

**Crystallization, structure solving, refinement and validation.** Crystallization was carried out by vapour diffusion at 20.5 °C using 96-well sitting drop plates (Greiner). Protein pellets were dissolved in water to 5–10 mg ml\(^{-1}\), and initial trials were obtained by equilibrating 100 ml of protein with 100 ml of reservoir solution C11 of the PGA-HT screen (Molecular Dimensions; pH 6.5) supplemented with guanidine hydrochloride (to 0.1 M), against 0.1 ml of reservoir. Glycogen was added (to 20% v/v) for cryoprotection. Diffraction data were collected at 100 K on Diamond beamlines I02, I03, I04 and I24 (Harwell, UK), and processed using the Xia2 programme suite\(^27\). A single-wavelength anomalous dispersion experiment allowed determination and refinement of the positions of selenium atoms, as well as calculation of the phases with autoSHARP\(^28\). An initial model was obtained using MR-ROSETTA\(^29\), enabling manual model building using COOT\(^30\). Refinement was performed using autoBUSTER\(^31\), and the program employed COOT and NOLPROB\(^2\). Molecular replacement (using PHASER\(^33\)) was used to solve additional CR-VI + structures. Refinement statistics are given in Table 1, and a portion of the electron density map is shown in Supplementary Fig. 8.

**Synthesis of RNA substrates.** RNA sequences were chemically synthesized on a solid support using an ABI 394 synthesizer. After RNA elongation with 2’-O-pivaloyloxymethyl phosphoramidite monomers \(^34\)–\(^37\) (Chemgenes, USA), the 5’ hydroxy group was phosphorylated and the resulting 5’-H-phosphate-derivatives oxidized and activated into a phosphoromonoazidololate derivative to react with either phosphoric acid (for pRNA synthesis), pyrophosphate (pppRNA)\(^38\) or guanosine diphosphate (GpppRNA)\(^37,38\). N7-methylation of the purified GpppRNA was performed enzymatically using N7-hmTase\(^39\). To prepare phosphorylated RNA for cross-linking experiments, the RNA was treated with a mixture of N\(_2\)-O-bis-trimethylethylamide (0.4 ml), CH\(_3\)CN (0.8 ml) and triethylamine (0.1 ml) at 35 °C for 15 min, and then oxidized with a tert-butyl
hydroperoxide solution (5–6 M in decane, 0.4 mL; 35 °C, 15 min). After deprotection and release from the solid support, RNA sequences were purified by IEX-HPLC (>95% pure) and their identity were confirmed by MALDI-TOF (Matrix-Assisted Laser Desorption/Ionization Time-of-Fight) spectrometry.

**MTase activity assays.** These were performed by combining 4 μM CR-VI+ with 0.7 μM of the purified and validated synthetic RNAs, 10 μM SAM and 0.33 μM 1H-SAM (Perkin Elmer) in 40 mM Tris-HCl (pH 8.5) and 1 mM diethiothreitol (DTT). Reactions (at 30 °C) were stopped by adding an equal volume of 1 M ice-cold SAH and the samples were transferred to DEAE filters (Perkin Elmer) using a Filtermat Harvester (Packard Instruments). The RNA-containing mats were washed twice with 10 mM ammonium formate (pH 8.0), twice with water and once with ethanol. Then they were then soaked with liquid scintillation fluid, allowing the

**References**

1. Rahmeh, A. A. et al. Molecular architecture of the vesicular stomatitis virus RNA polymerase. *Proc. Natl Acad. Sci. USA* **107**, 20075–20080 (2010).
2. Decroly, E., Ferron, F., Lescar, J. & Canard, B. Conventional and unconventional mechanisms for capping viral mRNAs. *Nat. Rev. Microbiol.* **10**, 51–65 (2012).
3. Züst, R. et al. Ribose 2′-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat. Immunol.* **12**, 137–143 (2011).
4. Daffis, S. et al. 2′-O-methylation of the viral mRNA cap evades host restriction by IFTT family members. Nature 468, 452–456 (2010).
5. Nishio, M. et al. Human parainfluenza virus type 2L protein regions required for interaction with other viral proteins and mRNA capping. J. Virol. 85, 725–732 (2011).
6. Gopinath, M. & Shaila, M. S. RNA triphosphatase and guanylyl transferase activities are associated with the RNA polymerase protein L of rinderpest virus. J. Gen. Virol. 90, 1748–1756 (2009).
7. Feuillet, F., Lina, B., Rosa-Catalava, M. & Boivin, G. Ten years of human metapneumovirus research. J. Clin. Virol. 53, 97–105 (2012).
8. Ogino, T. & Banerjee, A. K. Unconventional pathway of mRNA cap formation by vesiculoviruses. Virus Res. 162, 100–109 (2011).
9. Ray, D. et al. West Nile virus 5′-cap structure is formed by sequential guanine N-7 and ribose 2′-O-methylation by nonstructural protein 5. J. Virol. 80, 8362–8370 (2006).
10. Decroly, E. et al. Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2′-O-methyltransferase nsp10/nsp16 complex. PLoS Pathog. 7, e1002059 (2011).
11. Hodel, A. E., Gershon, P. D., Shi, X., Wang, S. M. & Quiocho, F. A. Specific protein recognition of an mRNA cap through its alkylated base. Nat. Struct. Biol. 4, 350–354 (1997).
12. Sutton, G., Grimes, I. M., Stuart, D. I. & Roy, P. B. Lentivirus vector VP4 is a RNA-capping assembly line. Nat. Struct. Mol. Biol. 14, 449–451 (2007).
13. Bollati, M. et al. Recognition of RNA cap in the Wesselsbron virus NS5 methyltransferase domain: implications for RNA-capping mechanisms in flaviviruses. J. Mol. Biol. 385, 140–152 (2009).
14. Liang, B. et al. Structure of the L protein of vesicular stomatitis virus from electron cryomicroscopy. Cell 162, 314–327 (2015).
15. Hodel, A. E., Gershon, P. D., Shi, X. & Quiocho, F. A. The 1.85Å structure of influenza polymerase. Proc. Natl Acad. Sci. USA 108, 10104–10108 (2011).
16. Hodel, A. E., Gershon, P. D., Shi, X. & Quiocho, F. A. Structural basis for sequence-nonspecific recognition of 5′-capped mRNA by a cap-modifying enzyme. Mol. Cell 1, 443–447 (1998).
17. Chu, C. et al. Structure of the guanylyltransferase domain of human mRNA capping enzyme. Proc. Natl Acad. Sci. USA 105, 10104–10108 (2011).
18. Hakansson, K., Doherty, A. J., Shuman, S. & Wigley, D. B. X-ray crystallography reveals a large conformational change during guanylyl transfer by mRNA capping enzymes. Cell 89, 545–553 (1997).
19. N-7 and ribose 2′-O-methylation of the viral mRNA cap evades host restriction by IFTT family members. Nature 468, 452–456 (2010).
20. Reich, S. et al. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. Nature 516, 361–366 (2014).
21. Cheng, N. et al. Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. Nature. doi:10.1038/nature15525.
22. Gupta, K. C. & Roy, P. Synthesis of capped and uncapped methylated oligonucleotides by the virion transcriptase of spring viremia of carp virus, a rhabdovirus. Proc. Natl Acad. Sci. USA 78, 4758–4762 (1981).
23. Barik, S. The structure of the 5′ terminal cap of the respiratory syncytial virus mRNA. J. Gen. Virol. 74(3): 485–490 (1993).
24. Ogino, T. & Banerjee, A. K. Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. Mol. Cell. 25, 85–97 (2007).
25. Reich, S. et al. Structural insights into the functional interactions between the 5′ cap and the RNA polymerase by cryo-electron microscopy. Nature 516, 361–366 (2014).
26. Cheng, N. et al. Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. Nature. doi:10.1038/nature15525.
27. Barik, S. The structure of the 5′ terminal cap of the respiratory syncytial virus mRNA. J. Gen. Virol. 74(3): 485–490 (1993).
28. Vonrhein, C., Blanc, E., Roversi, P. & Briegleb, G. Automated structure solution with autoSHARP. Methods Mol. Biol. 364, 213–230 (2007).
29. DeMaio, F. Advances in Rosetta structure prediction for difficult molecular-replacement problems. Acta Crystallogr. D Biol. Crystallogr. 69, 2202–2208 (2013).
30. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501 (2010).
31. BUSTER version 2.9. Global Phasing Ltd. (Cambridge, United Kingdom, 2010).
32. Davis, I. W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 35, W375–W383 (2007).
33. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).
34. Lavergne, T., Bertrand, J.-R., Vasseur, J.-J. & Debart, F. A labile-basal cap group for 2′-OH protection of ribonucleosides: a major challenge for RNA synthesis. Chem. Eur. J. 14, 9135–9138 (2008).
35. Debart, F., Lavergne, T., Janin, M., Dupouy, C. & Vasseur, J.-J. in Current Protocols in Nucleic Acid Chemistry Vol. 43 (eds Beaucauge, S. et al.) 3.19.11–3.19.27 (John Wiley & Sons, 2010).
36. Zlatov, I. et al. Chemical solid-phase synthesis of 5′-triphosphates of DNA, RNA, and their analogs. Org. Lett. 12, 2199–2193 (2010).
37. Thullier, Y. et al. Synthesis of 5′-cap-0 and cap-1 RNAs using solid-phase chemistry coupled with enzymatic methylation by human (guanine-N7)-methyltransferase. RNA 18, 856–868 (2012).
38. Barral, K. et al. Development of specific Dengue virus 2′-O- and N7-methyltransferase assays for antiviral drug screening. Antiviral Res. 99, 292–300 (2013).