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

Guido C. Paesen, Axelle Collet, Corinne Sallamand, Francoise Debart,
Jean-Jacques Vasseur, Bruno Canard, Etienne Decroly, Jonathan M. Grimes

HAL Id: hal-01439027
https://hal.science/hal-01439027

Submitted on 7 Jan 2020

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

Guido C. Paesen[1,*], Axelle Collet[2,*], Corinne Sallamand[3], Françoise Debart[3], Jean-Jacques Vasseur[3], Bruno Canard[2], Etienne Decroly[2,**] & Jonathan M. Grimes[1,4,**]

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 (^{SAM}P) also contains a novel pocket (^{NS}P) that can accommodate a nucleoside. CR-VI lacks an obvious cap-binding site, and the ^{SAM}P-adjoining site holding the nucleotides undergoing methylation (^{SUB}P) 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.

[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).

The *Mononegavirales* order groups five families of mono-partite, 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[1]. Capping involves the co-transcriptional addition of a guanosine (**G**) to the first nucleotide (N_1) of the nascent RNA chain via a 5'-5' triphosphate bridge, resulting in a $GpppN_1$- structure. Typically, this is followed by methylation of nitrogen 7 (*N7*) of **G**, giving rise to $^{m}GpppN_1$-, and of the 2'-oxygen (*2'O*) of the N_1 ribose ($^{m}GpppN_{1m}$-). The cap protects mRNAs against 5'-exonucleases and promotes RNA transport and translation, while *2'O*-methylation prevents detection by cellular-immunity sensors[2–4]. In *Rhabdoviridae*, CR-V catalyses cap addition by means of an unconventional polyribonucleotydyl-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 **G**pp to the pRNA. This mechanism differs from capping in eukaryotes and most other viruses, in which a guanylyltransferase (GTase) forms a phosphoamide bond with **G**p, before transferring it to 5'ppRNA[2]. *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 **G**p in the capping reaction[5]. A C-terminal domain of Rinderpest virus *L* (containing CR-VI and the downstream K-K-G motif) was shown to form such a bond[6], 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 transfer[2]. As yet, the boundaries of CR-VI, the putative MTase of *L*, are not well
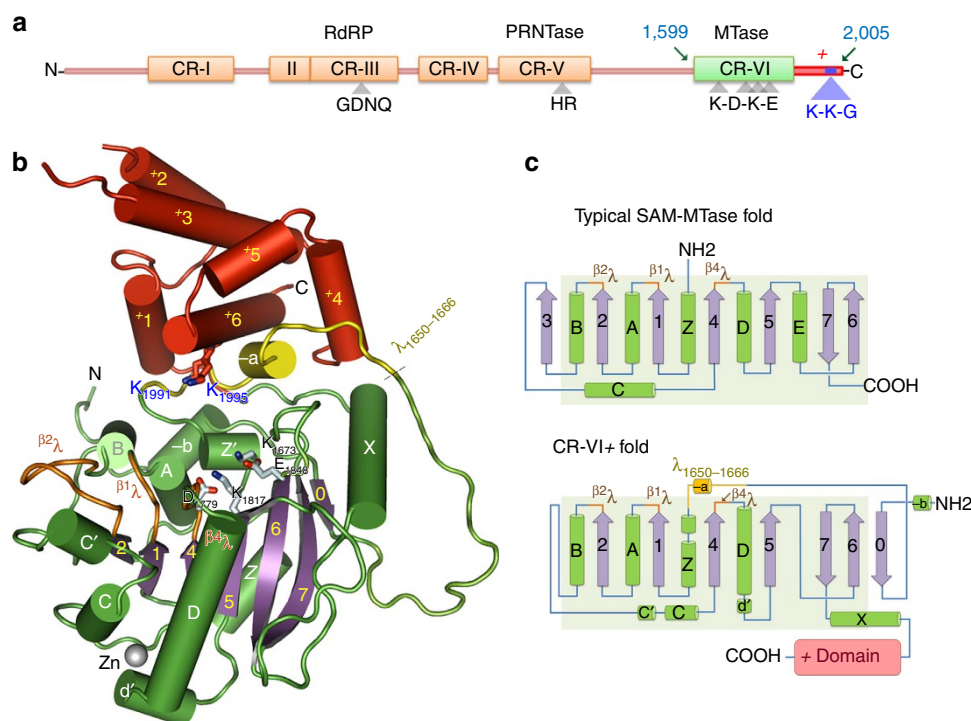

Figure 1 | The structure of the MTase domain of hMPV L. (a) Domain organization of hMPV *L*, with at its C terminus the 46.5 kDa CR-VI+ domain (residues 1,599–2,005), comprising CR-VI (green), which contains the K-D-K-E motif typical for 2'O-MTases, and the +domain (red), carrying the K-K-G motif (blue). 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 PRNTase 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 K_{1991} and K_{1995} 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 $^{β1}λ$, $^{β2}λ$ and $^{β4}λ$ (the loop regions C terminal of β-strands 1, 2 and 4 that form ^{SAM}P; orange), and $λ_{1650-1666}$ (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 H_{1766}, H_{1798}, C_{1802} and C_{1805}. **(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 $^{β1}λ$, $^{β2}λ$ and $^{β4}λ$ (orange) and $λ_{1650-1666}$ (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. αE is absent, whereas, atypical for viral MTases, αB 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 (β0), 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).

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 methylations 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[7]. 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 G$N7$ and N_1-$2'O$ positions (Figs 2 and 3a), with $2'O$-methylation preceding $N7$-methylation (Fig. 3a), an uncommon order of events also occurring in VSV[8]. CR-VI+, moreover, efficiently methylates uncapped RNAs with 5'-phosphate groups (especially pppRNA), primarily at the $2'O$ atom of N1 since almost no ^{3}H-methyl
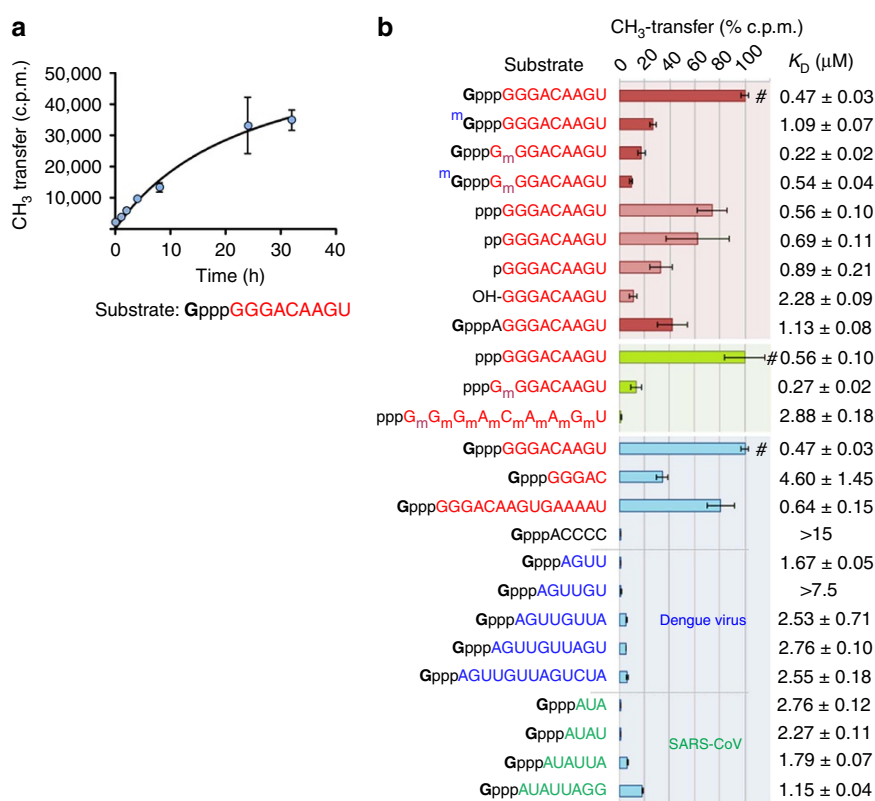
Figure 2 | MTase activity of CR-VI+. (a). The transfer of tritiated methyl groups from SAM molecules to a capped RNA substrate (**G**pppGGGACAAGU), containing the consensus start sequence of hMPV transcripts (in red), was monitored over time. The rather slow *in vitro* methylation suggests the reaction 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 **G**pppGGGACAAGU (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 group, either at their $N7$-guanine or $2'O$-ribose position (confirming the occurrence of $2'O$- and $N7$-methylation, respectively), and of uncapped RNAs (especially pppRNA). The pppRNA substrate is predominantly methylated at N_1, as substrates that were methylated beforehand at this nucleotide do not undergo substantial additional methylation (green panel). CR-VI+ prefers the hMPV start sequence over the short **G**pppACCCC sequence, and over RNA-start sequences of Dengue virus and SARS coronavirus, irrespective of their lengths (blue panel). A nine-nucleotide hMPV substrate, however, is much preferred over one with only five nucleotides, indicating that additional interactions take place between the protein and nucleotides 6–9. Consistently a 10-times lower K_D (dissociation constant) characterizes the interaction of CR-VI+ with the 9-mer, compared with that with the 5-mer. The K_Ds, measured in triplicate using a dot-blot assay and listed at the right of the diagram (± s.d.'s), also show that capped and uncapped hMPV sequences are bound with comparable affinities, and that $2'O$-methylated substrates are preferred over unmethylated ones.
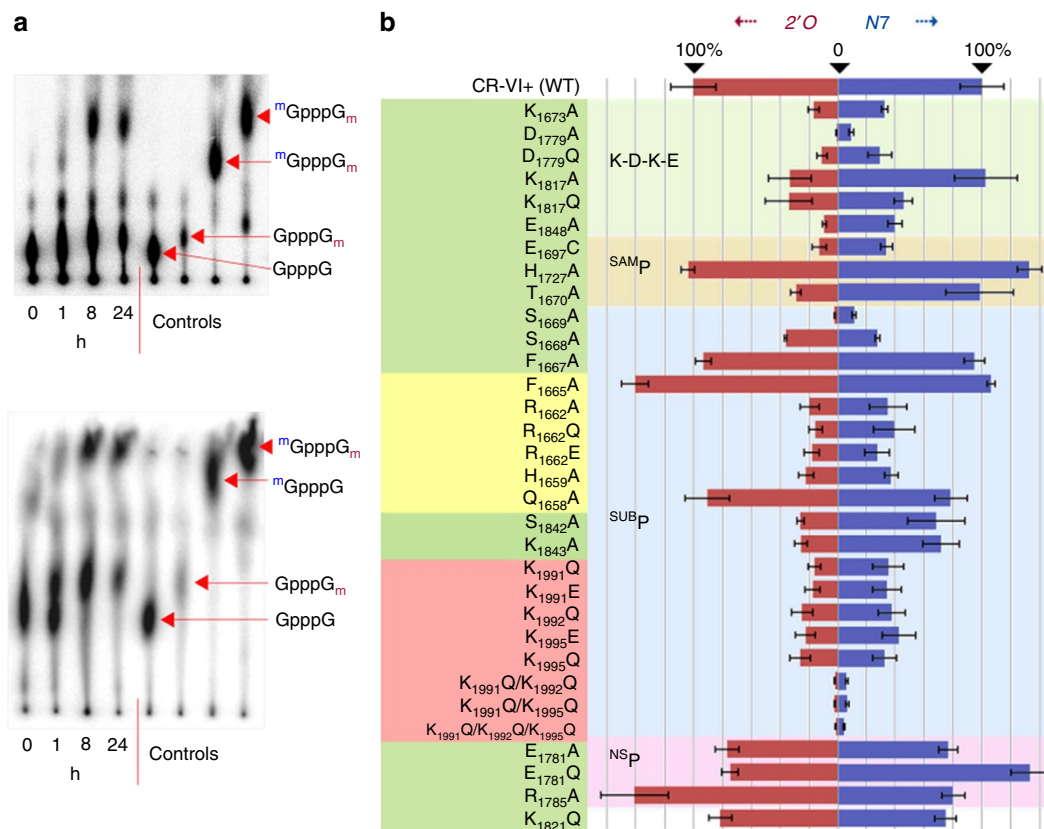
Figure 3 | N7- and 2′O-methylation. (a) Thin-layer chromatograms. Following CR-VI + -mediated methyl transfer from SAM onto **G**pppGGGACAAGU (in which **G** was [^{32}P]-labelled), nucleotides 2–9 were removed by nuclease P1 digestion, and the caps were separated by thin-layer chromatography (TLC). The controls (**G**pppG, **G**pppG$_m$, ^{m}**G**pppG and ^{m}**G**pppG$_m$, 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 ^{m}**G**pppG and ^{m}**G**pppG$_m$ (top). The caps on the TLC plate were subsequently further resolved, this time using 0.45 M (NH$_2$)$_2$SO$_4$ as mobile phase for a better separation of **G**pppG and **G**pppG$_m$ (bottom). **G**pppG$_m$ appears first (after a 1-h incubation), ^{m}**G**pppG$_m$ becomes prominent at a later stage, and ^{m}**G**pppG was not observed, indicating that 2′O-methylation of N$_1$ 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 **G**pppGGGACAAGU 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 λ$_{1650-1666}$, 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, ^{SAM}P, ^{SUB}P or ^{NS}P. The results highlight the importance of essential ^{SUB}P residues (such as the K-K-G lysines K$_{1991}$, K$_{1992}$ and K$_{1995}$ and λ$_{1650-1666}$ residues H$_{1659}$ and R$_{1662}$) for both 2′O- and N7-methylation. All tetrad residues are crucial for 2′O-methylation, while D$_{1779}$ 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.

transfer takes place onto the pppG$_m$GGACAAGU substrate where this atom is blocked (Fig. 2b, green bars). The 2′O-methylation of pppGGGACAAGU represents ∼50% of that observed with **G**pppGGGACAAGU (which has an extra methylation target in the **G**N7 atom; Fig. 2b, red bars). Mutagenesis studies show that all K-D-K-E residues (K$_{1673}$, D$_{1779}$, K$_{1817}$ and E$_{1848}$) are essential for 2′O-methylation, while the aspartic acid (D$_{1779}$) in particular is important for N7-methylation (Fig. 3b), as was also observed in *Flaviviridae* MTases[9]. Other residues, many of which belong to the + domain rather than the core MTase fold, were also found to be essential for the MTase activities (the '^{SUB}P' residues in Fig. 3b, described below).

Structure of the + domain. The crystal structure of CR-VI + was solved to 2.2 Å resolution in space group P2$_1$2$_1$2$_1$ with two molecules, disulphide-linked at residue C$_{1877}$, 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 *Rhabdoviridae*. In hMPV, it consists of six α-helices (α$^+$1–6). Helix α$^+$6 contains the K-K-G motif (Supplementary Fig. 2) and together with α$^+$1 leans over the active cleft of CR-VI. Helix α$^+$4, and to a lesser extent helices α$^+$1, α$^+$5 and α$^+$6, packs down on λ$_{1650-1666}$, the second half of a long loop (residues 1,635–1,666) that swerves around the CR-VI domain (Fig. 1b and Supplementary Fig. 3; 'λ' is used throughout the paper to denote loops). The λ$_{1650-1666}$ region, which contains a small helix (α-a), acts as a fulcrum allowing the + domain to pivot relative to the CR-VI domain. Helix α$^+$3 varies in length, from 4½ turns (in most crystals) to 6 (in Protein Data Bank (PDB) 4UCY), and the loop between α$^+$2 and α$^+$3 is always disordered (Supplementary Fig. 3). In crystals of the monomeric C$_{1877}$A mutant (space group P3$_1$21), 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 2'O-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, 3EVF[9]; Supplementary Fig. 4). Most notably, hMPV and flaviviruses share an unusually long (~10 residues), flexible $^{\beta 2}\lambda$ (that is, the loop immediately following the β2-strand), which forms the SAM-binding pocket (^{SAM}P) along with loops $^{\beta 1}\lambda$ and $^{\beta 4}\lambda$, shielding it from the solvent (Fig. 1b, c). In Wesselsbron (flavi-)virus, $^{\beta 2}\lambda$ 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+, $^{\beta 2}\lambda$ similarly assumes alternate conformations (Fig. 4c,d). In the 'closed' form, the ligand's ribose group is hydrogen bonded to D_{1725}, and—via a water—to D_{1722}. Loops $^{\beta 1}\lambda$ and $^{\beta 4}\lambda$ also show a degree of flexibility. $^{\beta 1}\lambda$ residue E_{1697}, conserved in paramyxo- and

filoviruses, forms a hydrogen bond with the NH_2 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 NH_2 group, or, more markedly, turns towards the solvent in the direction of $^{\beta 2}\lambda$, which in this case assumes the 'open' position (Fig. 4d). $^{\beta 4}\lambda$ forms a side wall of ^{SAM}P, and together with αD and β5 also defines a deep, hydrophobic cavity not present in other MTases, termed ^{NS}P (or nucleoside-binding pocket). Although a role for ^{NS}P 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 $^{\beta 4}\lambda$ to impinge onto ^{SAM}P (Fig. 4c,d; when SAM is used for soaking, one SAM molecule occupies ^{SAM}P, and another binds to ^{NS}P). GTP was not observed in ^{NS}P, but this may reflect a lower soaking concentration (2.5 mM), due to GTP's poor solubility. With the exception of E_{1781}, the amino acids lining ^{NS}P are poorly conserved beyond the *Pneumovirinae*, and mutating key ^{NS}P residues barely affects the MTase activities (Fig. 3b). CR-VI, finally, contains a Zn-finger, which is not conserved beyond the *Pneumovirinae* subfamily and links the small α–d′ to the rest of the structure (Fig. 1b).
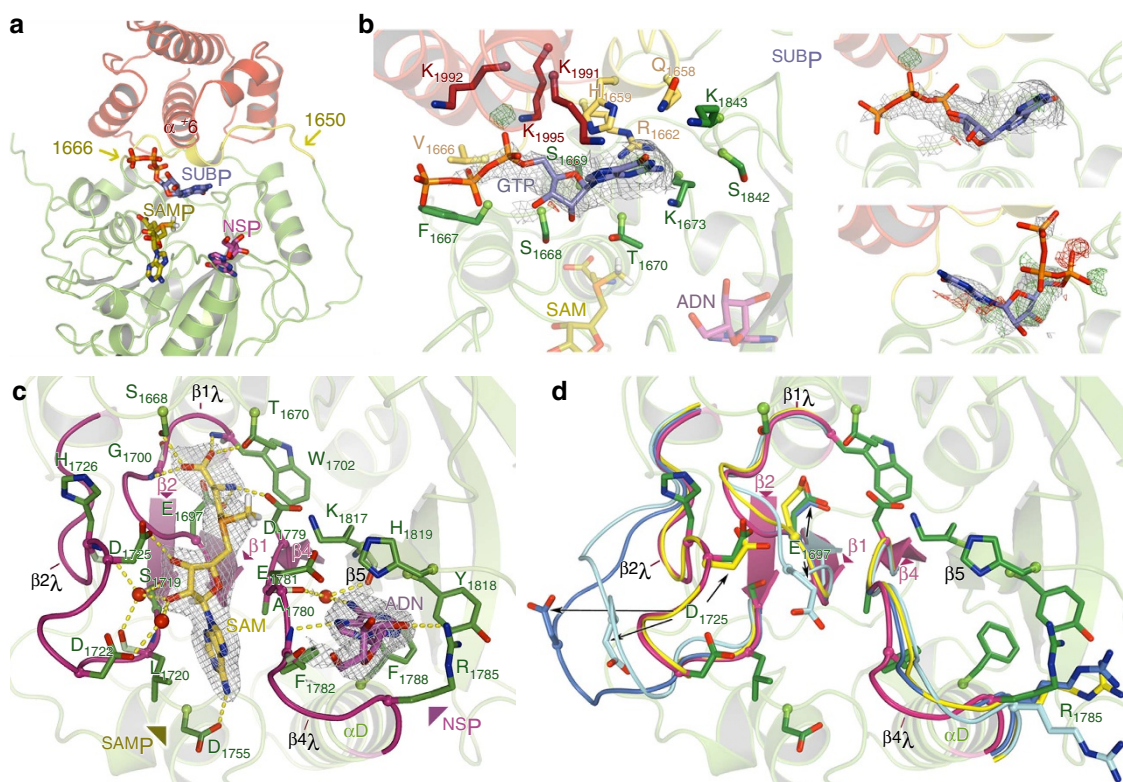

Figure 4 | CR-VI+-binding pockets. The cartoon representations show the +domain in red and CR-VI in green, with $\lambda_{1650-1666}$ in yellow. SAM (in ^{SAM}P), GTP (in ^{SUB}P) and adenosine (ADN; in ^{NS}P) are shown as sticks, with the C atoms coloured gold, slate and magenta, respectively. Hydrogens (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 ^{SUB}P, which is defined by residues of the +domain (particularly the K-K-G motif), $\lambda_{1650-1666}$ 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 4UCI 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**) ^{SAM}P and ^{NS}P containing a SAM and ADN ligand, respectively (PDB 4UCI, in which ^{SUB}P is also occupied). Residues lining the pockets are shown as sticks. The loops delineating ^{SAM}P ($^{\beta 1}\lambda$, $^{\beta 2}\lambda$ and $^{\beta 4}\lambda$) 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 $^{\beta 1}\lambda$ (especially E_{1697}), $^{\beta 2}\lambda$ and $^{\beta 4}\lambda$. The structure in blue (PDB 4UCK) contains SAM, whereas those in yellow (4UCL) and aquamarine (4UCJ) have empty ^{SAM}Ps (this suggests that there is no strict correlation between ^{SAM}P occupancy and the position of $^{\beta 2}\lambda$). ^{NS}P is empty in the three superposed structures, which apparently affects the position of their $^{\beta 4}\lambda$ loops and especially of the R_{1785} side group, which closes the pocket when occupied. All overlaid structures have empty ^{SUB}Ps.

During the preparation of this manuscript, a 3.8-Å structure of VSV L, obtained by electron cryo-microscopy, was published[14]. The CR-VI (or MTase) part of VSV resembles that of hMPV, but apparently lacks a deep ^{NS}P pocket (PDB: 5a22, Fig. 5). The + domain (C-terminal domain) is more elaborate in VSV than in hMPV, containing extra regions N and C terminal of helix α^+6. 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 (R_{2038}) strictly conserved among the *Rhabdoviridae*, which is structurally equivalent to hMPV's K_{1995}, 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[15,16].

In CR-VI+, however, an open, solvent-exposed area is found where this pocket is normally located (Fig. 6a). Moreover, **GpppG**- or ^{m}**GpppG** 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 ^{SUB}P for N7-methylation. Although a high-affinity cap-binding site is clearly absent from CR-VI+, the low-affinity nucleoside binding to ^{NS}P and the convenient location of this pocket relative to ^{SUB}P 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). In
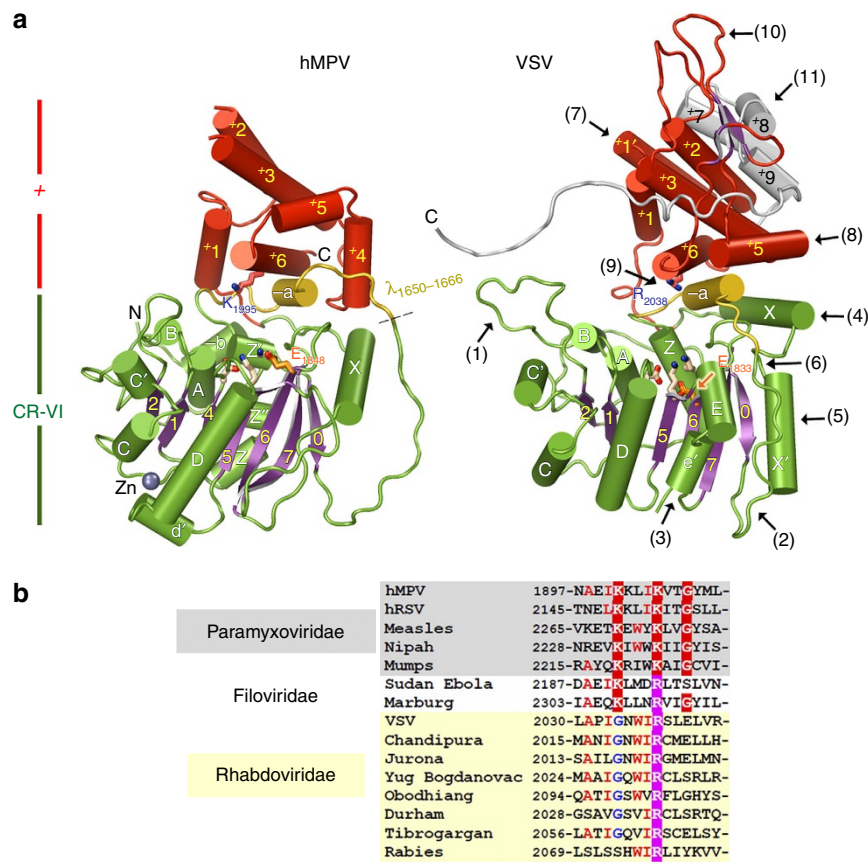

Figure 5 | Comparison of the CR-VI+ domains of hMPV and VSV. (a). Cartoon representations. The CR-VI domains are similar, and share the unusual, strand-0 containing β-sheet (purple), the rather large $^{\beta2}\lambda$—indicated by arrow (**1**)—and the long N-terminal loop (**2**), which runs somewhat differently in VSV. The $\lambda_{1650-1666}$ 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 ^{NS}P may have disappeared. αX is at a different location (**4**), and is preceded by an extra helix (αx′(**5**)). E_{1833}, expected to belong to the K-D-K-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-K-E glutamate is occupied by T_{1831}. The + domain of VSV is tilted, compared with that of hMPV, and more elaborate. Helices α^+1, α^+2, α^+3, α^+5 and α^+6 are conserved, but the α^+1-α^+2 loop is replaced by an extra helix (α^+1' (**7**)). α^+4 is absent, whereas α^+5 is enlarged and immediately follows α^+3 (**8**). The 2-residue loop connecting α^+5 and α^+6 in hMPV is replaced by a 34-residue coil carrying a small three-stranded β-sheet (**9**). Helix α^+6 seems best conserved between the two + domains, although a K-K-G motif is not present in VSV. However, R_{2038}, which is strictly conserved in *Rhabdoviridae*, takes the place of K_{1995} ((**10**) and alignment below). In VSV, the + domain is extended beyond α^+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 α^+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 ^{SAM}P, which holds the nucleotides undergoing methylation, has become a more elaborate, but narrower and possibly therefore, higher-affinity substrate-binding pocket (termed ^{SUB}P as it accommodates the nucleotide undergoing methylation). Consistently, electron density is found in ^{SUB}P following soaking or co-crystallization with GTP, whereas in other MTases added GTP predominantly shows up in the cap-binding pocket. In particular, helix α^+6 and the +domain-affiliated $\lambda_{1650-1666}$ help shape ^{SUB}P through the side chains of K_{1991} and K_{1995} (of the K-K-G motif), and of H_{1659} and R_{1662}, respectively (Fig. 4b). The marked decrease in MTase activity of mutants altered at these residues (Fig. 3b) illustrates the importance of ^{SUB}P in correctly presenting the substrate nucleotides to SAM. Nevertheless, the pocket is too spacious for a single nucleotide, and the electron density in ^{SUB}P 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 predominantly 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 αZ, but instead resides on the small z' (3_{10})-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).

A potential role of CR-VI+ in cap addition. K-K-G residue K_{1995} corresponds to the **G**p binding lysine in the signature motif of eukaryotic GTases, but here is in a loop instead of an α-helix (PDB: 3S24 (ref. 18), 1CKN (ref. 19)) and appears ideally placed to target the phosphate tail of GTP. Although incubation of CR-VI+ with [α-^{32}P]- and [β-^{32}P]-GTP, in the absence of
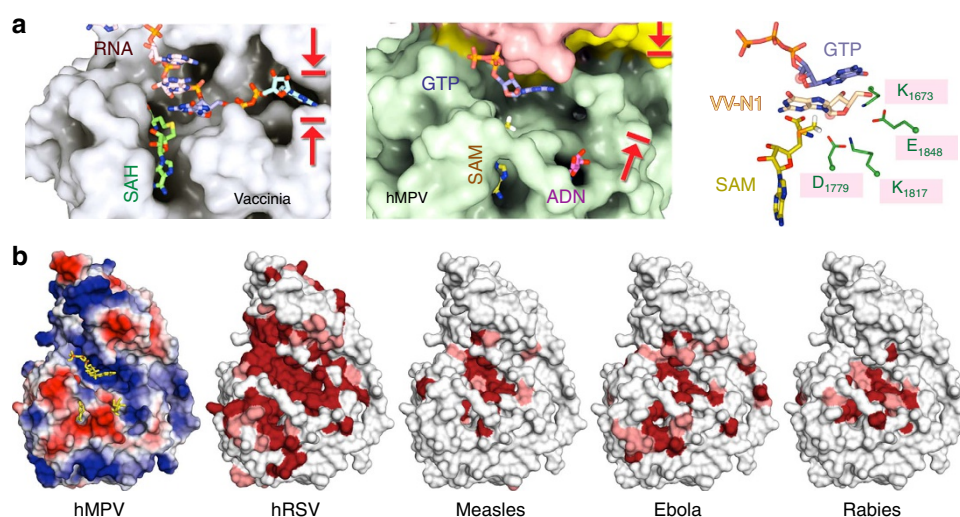
Figure 6 | RNA-binding site comparisons. (a). Comparison of the RNA-binding sites in vaccinia virus cap-MTase (PDB 1AV6) and CR-VI+. The vaccinia virus MTase (white surface, left) has a narrow cap-binding pocket (in between the red arrows) and a large, open RNA-binding site (adjoining the SAH-containing ^{SAM}P). In CR-VI+ (coloured surface, middle) the cap-binding pocket is not present, whereas the RNA-binding site is narrowed (into ^{SUB}P) by the +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 2'O atoms of the nucleotides are shown as transparent, red spheres. **(b)**. ^{SUB}P 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 ^{SUB}P and ^{SAM}P.
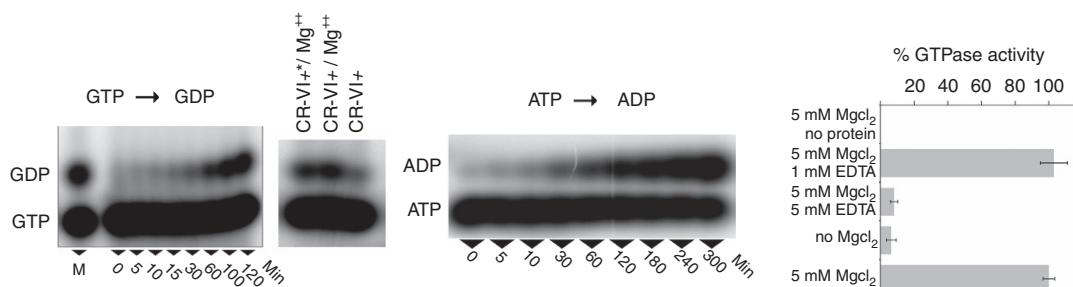
Figure 7 | NTPase activity. Autoradiographs of urea–PAGE gels show CR-VI+-mediated conversion of radiolabelled GTP to GDP and ATP to ADP, over time. The smaller autoradiograph shows the requirement of Mg^{++} for the reaction (allowed to proceed for 1 h), and also shows efficient GDP generation by washed, 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^{++} 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^{++}, 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 phosphoamide 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 **G**pp ligation to pRNA[20].

We observed that CR-VI+ also displays NTPase activity, converting GTP into GDP and ATP into ADP (Fig. 7). GTPase 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 GTPase activity were obtained with $E_{1697}C$ (corresponding to the flexible residue at the bottom of ^{SAM}P; Fig. 4d) and $S_{1669}A$ (part of ^{SUB}P, Fig. 4b) to 54 (± 20) and 57 (± 16) % of the wild-type activity, respectively ($n = 3$). The deletion of the dipeptide $G_{1645}K_{1646}$ from the long N-terminal loop resulted in a somewhat more pronounced $\sim 70\%$ reduction (Supplementary Fig. 6). 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, moreover, has been shown to be essential for viable mRNA synthesis in human parainfluenzavirus 2 (hPIV-2)[5], in line with capping involving a concerted action of CR-V and CR-VI+. Assuming the transcript is presented to CR-VI+ by the PRNTase, $2'O$-methylation may actually occur before the cap is added (as a cap is not required for the reaction; Fig. 2b), which would also be consistent with $2'O$-methylation preceding $N7$-methylation. The possibility that $2'O$-methylation precedes cap addition was also suggested by the ability of virions of the rhabdovirus spring viremia of carp virus to synthesize uncapped, N1-methylated oligonucleotides[24]. We did not observe significant methylation of free GTP by CR-VI+ (levels were $<5\%$ of that of **G**pppGGGAC AAGU methylation), suggesting **G**$N7$-methylation takes place after **G** is added to the RNA. In addition $2'O$-methylated substrates (**G**pppG_mGGACAAGU and pppG_mGGACAAGU) show an ~ 2-fold higher affinity for CR-VI+, compared with their unmethylated counterparts (Fig. 2b), which presumably reflects an RNA repositioning mechanism allowing the **G**$N7$ atom access to the SAM methyl donor in the second methylation step.

In conclusion, CR-VI+ is a dynamic part of the L protein that potentially completes PRNTase-initiated cap addition, and methylates the cap at its $2'O$ and $N7$ positions. The ^{SUB}P pocket, which is conserved among *Paramyxoviridae* and *Filoviridae* (Fig. 6b), is clearly instrumental in key activities of the domain, 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 using primers that added a C-terminal SGHHHHHH-tag to the translation product, from a synthetic hMPV L gene (hMPV isolate 00-1, GenBank: AF371337.2), codon optimized for expression in *Spodoptera frugiperda* (*Sf*) cells (Geneart). The amplicon was cloned into pOPIN-E for expression in HEK293T mammalian cells (following transfection using Lipofectamine 2000; Invitrogen), in BL21 Star bacteria (Invitrogen) and in *Sf*21 (insect cells) following co-transfection with (flashBACULTRA) baculovirus (Oxford Expression Systems) using Cellfectin 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 $K_{1991}Q/K_{1995}Q$ and $K_{1991}Q/K_{1992}Q/K_{1995}Q$ mutants, which were obtained using the $K_{1995}Q$ DNA. For each mutation, the forward primer was combined with a vector-specific reverse primer (5'-AGTGGTATTTG TGAGCCAGG-3'), and, in a second PCR, the reverse primer was used together with a vector-specific forward primer (5'-CCTTTAATTCAACCCAACAC-3'). The two PCR products thus generated were digested with either BspQI or BtsI (New England Biolabs), restriction enzymes that cut within the CR-VI+ -specific primer regions. The amplicons were then rejoined using T4 DNA ligase (New England Biolabs), and the ligation products were PCR amplified using the vector-specific primers for insertion in the pOPIN-E vector.

Purification. Seventy-two hours after infection of the suspension cultures with recombinant baculovirus, *Sf*21 cells were spun down (1,000g, 7 min, 22 °C) and lysed in one-fourth volume of buffer L (1.5% triton X-100, 5% glycerol, 50 mM arginine, 300 mM KCl, 10 mM imidazole and 20 mM Tris (pH 8.0)). After clarification (10,000g, 25 min, 4 °C), the lysate was incubated with benzonase (Novagen; to 0.2 U ml^{-1}) and NiNTA resin (Qiagen; 0.3 ml l^{-1} culture) for 5 h at 4 °C, with gentle shaking. The beads were transferred to a 10-ml column and washed with 3×10 ml buffer W (20 mM Tris (pH 8.0), 1.5 M NaCl, 10 mM imidazole and 7.5% glycerol) and 1.5 ml of 0.1 M arginine (pH 8.0). Protein was eluted in 0.8 M arginine (pH 8.0). An equal volume of ice-cold 3 M AmSO$_4$ was added to the eluent, and the precipitated protein collected (10,000g, 10 min, 4 °C). Protein pellets were stored at -20 °C. A similar lysis and purification protocol was used to obtain recombinant protein from HEK293T cells. BL21 Star cells were lysed using Bugbuster Protein Extraction Reagent (Novagen). Expression and purification were monitored by SDS–PAGE and western blotting. Blots were developed with a 1/10,000 dilution of an anti-histidine tag antibody (Penta · His Antibody, Qiagen, catalogue number 34660). Supplementary Fig. 1 shows purified protein following SDS–PAGE under reducing and non-reducing conditions. An uncropped western blot showing the unreduced, insect-cell-expressed protein next to a BenchMark (Invitrogen) molecular weight marker is also shown.

Selenomethionine incorporation. Twenty hours following infection of an *Sf*21 suspension culture (27.8 °C, with an agitation speed of 130 r.p.m./0.25g), cells were collected (70g, 10 min, 22 °C) and resuspended in cysteine- and methionine-free SF900II medium (Gibco) supplemented with dialysed fetal bovine serum (Gibco; 7% v/v) and 150 mg l^{-1} L-cysteine (Sigma). Following an additional 4 h at 27.8 °C in the shaking incubator, 250 mg 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)[26]. Protein pellets were dissolved in water to 5–6 mg ml^{-1}, and initial crystals were obtained by equilibrating 100 nl of protein with 100 nl 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. Glycerol 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 validation employed COOT and Molprobity[32]. 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[34,35] (Chemgenes, USA), the 5'-hydroxyl group was phosphorylated and the resulting H-phosphonate derivative[36] oxidized and activated into a phosphoroimidazolidate derivative to react with either phosphoric acid (for ppRNA synthesis), pyrophosphate (pppRNA)[36] or guanosine diphosphate (**G**pppRNA)[37,38]. $N7$-methylation of the purified **G**pppRNA was performed enzymatically using $N7$-hMTase[37,38]. To prepare monophosphate RNA (pRNA), the 5'-H-phosphonate RNA was treated with a mixture of N,O-bis-trimethylacetamide (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

Table 1 | Crystallographic data collection and refinement statistics.

PDB code	4UCI	4UCL	4UCJ	4UCK	4UCZ	4UCY	4UD0
Construct	Native*	Native†	Native‡	$K_{397}E$§	Native‖	Native¶	$C_{279}A$#
Ligands fitted (^{SAM}P/ ^{SUB}P/ ^{NS}P)	SAM/GTP/ADN	–/SO_4/–	–/SO_4/–	SAM/–/–	–/GTP (A) GMP (B)/–	–/SO_4/–	SAH/SO_4/–
Data collection							
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P3_121$
Cell dimensions							
$a/b/c$ (Å)	80.2/83.6/182.6	78.8/83.4/182.9	82.1/84.5/181.2	80.4/84.4/181.2	78.8/83.7/182.6	78.3/82.2/183.1	75.6/75.6/187.9
$\alpha/\beta/\gamma$ (°)	90/90/90	90/90/90	90/90/90	90/90/90	90/90/90	90/90/90	90, 90, 90
Resolution (Å)	83.6–2.2 (9.9–2.2)	182.9–2.8 (2.9–2.8)	42.3–3.3 (3.3–3.3)	84.4–2.7 (2.7–2.7)	78.8–3.0 (3.1–3.0)	82.2–2.8 (2.9–2.8)	62.7–3.2 (3.3–3.2)
R_{meas}	0.086 (2.33)	0.098 (2.53)	0.23 (2.13)	0.13 (3.25)	0.16 (1.32)	0.11 (2.00)	0.17 (3.15)
R_{pim}	0.024 (0.626)	0.098 (2.527)	0.064 (0.594)	0.037 (0.895)	0.035 (0.436)	0.038 (0.687)	0.029 (0.689)
$CC_{1/2}$	0.999 (0.622)	0.998 (0.596)	0.996 (0.528)	0.999 (0.616)	0.978 (0.759)	0.999 (0.440)	1.000 (0.577)
$I/\sigma(I)$	17.4 (1.4)	16.5 (1.1)	9.7 (1.4)	15.8 (1.1)	45.2 (1.6)	13 (1.1)	21.2 (1.4)
Completeness (%)	99.9 (99.9)	99.8 (99.8)	99.9 (100)	99.9 (99.8)	100.0–99.9	99.9 (100)	99.8 (99.9)
Redundancy	13.2 (13.7)	13.1 (13.5)	12.6 (12.8)	13.1 (13.0)	21.5 (9.1)	8.6 (8.3)	35.1 (20.4)
Refinement							
Resolution (Å)	44.2–2.2	91.5–2.8	42.3–3.3	76.5–2.7	76.1–3.0	61.2–2.8	62.7–3.2
No. of reflections	62,253	30,379	20,304	35,718	24,839	28,861	10,817
R_{work}/R_{free} (%)	19.0/20.1	18.9/21.7	18.6/22.0	18.6/20.5	19.7/23.8	19.3/22.0	19.9/22.9
No. of atoms (protein/ligand/water)	5,943/608/382	6,000/101/89	6,014/90/58	6,065/212/157	6,142/147/80	6,260/98/86	2,840/52/20
R.m.s. deviations							
Bond lengths (Å)/bond angles (°)	0.008/0.97	0.008/1.02	0.008/0.98	0.008/1.02	0.008/1.01	0.009/1.04	0.008/1.04
Ramachandran plot							
Preferred/allowed/ outlier (%)	96.8/2.9/0.3	96.0/3.6/0.4	95.3/4.4/0.3	95.7/4.2/0.1	95.2/4.3/0.5	94.9/4.8/0.3	96.5/3.5/0.00
Molprobity scores							
Protein geometry	0.96	1.66	1.94	1.52	1.79	1.82	2.04
Clash score all atoms	3.65	3.65	3.65	3.65	2.97	3.09	3.65

R.m.s, root mean square.
Outer shell values are given between parantheses.
*Morpheus HT G11 crystallization condition (molecular dimensions), guanidine (300 mM); crystal soaked with 25 mM SAM.
†Morpheus HT G11, guanidine (300 mM).
‡SALTRx HT F01 (Hampton Research), 10 mM SAM
§PEGRx HT F01 (Hampton Research), 10 mM SAM
‖Morpheus HT G11, guanidine (120 mM); crystal soaked with 2.5 mM GTP.
¶Morpheus HT G11, guanidine (120 mM).
#ProPlex HT F7 (Molecular Dimensions), 10 mM SAH.

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 ^{3}H-SAM (Perkin Elmer) in 40 mM Tris-HCl (pH 8.5) and 1 mM dithiothreitol (DTT). Reactions (at 30 °C) were stopped by a 10-fold dilution in 100 µM ice-cold SAH and the samples were transferred to DEAE filtermats (Perkin Elmer) using a Filtermat Harvester (Packard Instruments). The RNA-retaining mats were washed twice with 10 mM ammonium formate (pH 8.0), twice with water and once with ethanol. They were then soaked with liquid scintillation fluid, allowing the measurement of ^{3}H-methyl transfer to the RNA substrates using a Wallac MicroBeta TriLux Liquid Scintillation Counter[13].

Methylation of GTP was determined in the same buffer (minus the ^{3}H-SAM and RNA), using the EPIgeneous methyltransferase assay kit (Cisbio), which measures the generation of SAH, as it competes with d2-coupled SAH for binding to a Lumi4-Tb-labelled anti-SAH antibody, affecting the TR-FRET signal between these compounds. In practice, MTase reactions were stopped after 16 h by addition of the detection reagents, and 1 h later the TR-FRET signal was monitored using a PHERAstar Flashlamp plate reader.

Thin-layer chromatography analysis of cap structures. G*pppGGGACAAGU (in which the asterisk indicates the [^{32}P]-labelled phosphate) was synthesized by incubating pppGGGACAAGU (10 µM) with vaccinia virus capping enzyme (New England Biolabs) in the presence of 1.65 µCi [α-^{32}P]-GTP (Perkin Elmer). The capped RNA was purified by precipitation in 3 M sodium acetate supplemented with 1 µg µl^{-1} of glycogen (Thermo Scientific), and submitted to methylation by CR-VI+ (as above), after which it was precipitated again (stopping the reactions), and digested with 1 U of Nuclease P1 (US Biologicals) in 30 mM sodium acetate (pH 5.3), 5 mM ZnCl$_2$ and 50 mM NaCl (4 h, 37 °C). The products were spotted onto polyethylenimine cellulose thin-layer chromatography plates (Macherey

Nagel), and resolved in two steps, first using 0.65 M LiCl, then 0.45 M (NH$_2$)$_2$SO$_4$ as mobile phase. The radiolabelled caps released by nuclease P1 were visualized using a Fluorescent Image Analyzer FLA3000 (Fuji) phosphor-imager.

Characterization of RNA-CR-VI+ interactions. Using T4 RNA ligase 1 (20 units; New England Biolabs), [5′-^{32}P]-pCp (1.1 µCi) was ligated to the 3′-ends of the RNA substrates (10 µM) in 50 mM Tris-HCl (pH 7.8), 10 mM MgCl$_2$, 10 mM DTT and 1 mM ATP (16 °C, overnight). Ligase was removed by RNA precipitation in 3 M sodium acetate supplemented with 1 µg µl^{-1} of glycogen (Thermo Scientific). The radiolabelled RNA was incubated (15 min, 37 °C) with increasing concentrations of CR-VI+, in 20 mM Tris-HCl (pH 8.5), 1 mM DTT, 10% glycerol and 30 mM NaCl. The reaction mixtures were spotted onto nitrocellulose (GE Healthcare) using a manifold-1 dot-blot apparatus (Whatman) and washed with 20 mM Tris-HCl (pH 8.5), 1 mM DTT and 50 mM NaCl. Membrane-bound RNA was quantified by phosphor-imaging. Dissociation constants (K_Ds) were determined using Hill slope curve fitting (Prism).

NTPase assay. Samples combining CR-VI+ (4 µM) with 0.5 µCi [α-^{32}P]-GTP or 0.5 µCi [α-^{32}P]-ATP in 40 mM Tris-HCl (pH 8.5), 1 mM DTT and 5 mM MgCl$_2$ were incubated at 37 °C. Reactions were stopped by adding an equal volume of formamide/EDTA gel-loading buffer, and hydrolysis products were separated over a 20% polyacrylamide/8 M urea gel before phosphor-imaging.

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 mRNA. *Nat. Rev. Microbiol.* **10,** 51–65 (2012).
3. Zust, 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).

9

4. Daffis, S. et al. 2′-O-methylation of the viral mRNA cap evades host restriction by IFIT 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-Calatrava, M. & Boivin, G. Ten years of human metapneumovirus research. J. Clin. Virol. 53, 97–105 (2012).
8. Ogino, T. & Banerjee, A. K. An 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 methylations 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, J. M., Stuart, D. I. & Roy, P. Bluetongue virus VP4 is an 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 Flavivirus. 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 1.85A structure of vaccinia protein VP39: a bifunctional enzyme that participates in the modification of both mRNA ends. Cell 85, 247–256 (1996).
16. Smietanski, M. et al. Structural analysis of human 2′-O-ribose methyltransferases involved in mRNA cap structure formation. Nat. Commun. 5, 3004 (2014).
17. Hodel, A. E., Gershon, P. D. & Quiocho, F. A. Structural basis for sequence-nonspecific recognition of 5′-capped mRNA by a cap-modifying enzyme. Mol. Cell 1, 443–447 (1998).
18. Chu, C. et al. Structure of the guanylyltransferase domain of human mRNA capping enzyme. Proc. Natl Acad. Sci. USA 108, 10104–10108 (2011).
19. Hakansson, K., Doherty, A. J., Shuman, S. & Wigley, D. B. X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes. Cell 89, 545–553 (1997).
20. Barik, S. The structure of the 5′ terminal cap of the respiratory syncytial virus mRNA. J. Gen. Virol. 74(Pt 3): 485–490 (1993).
21. 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).
22. Reich, S. et al. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. Nature 516, 361–366 (2014).
23. Hengrung, N. et al. Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. Nature. doi:10.1038/nature15525.
24. 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).
25. Berrow, N. S. et al. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. Nucleic Acids Res. 35, e45 (2007).
26. Walter, T. S. et al. A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization. Acta Crystallogr. D Biol. Crystallogr. 61, 651–657 (2005).
27. Winter, G., Lobley, C. M. & Prince, S. M. Decision making in xia2. Acta Crystallogr. D Biol. Crystallogr. 69, 1260–1273 (2013).
28. Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. Automated structure solution with autoSHARP. Methods Mol. Biol. 364, 215–230 (2007).
29. DiMaio, 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 base-labile 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 Beaucage, S. et al.) 3.19.11–13.19.27 (John Wiley & Sons, 2010).
36. Zlatev, I. et al. Chemical solid-phase synthesis of 5′-triphosphates of DNA, RNA, and their analogs. Org. Lett. 12, 2190–2193 (2010).
37. Thillier, 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).

Author contributions

G.C.P., A.C., E.D. and J.M.G. conceived and designed the study; G.C.P., A.C. and E.D. performed experiments, using material prepared by C.S., F.D. and J.-J.V.; G.C.P., A.C., E.D, J.M.G. and B.C. analysed data and wrote the paper.

Additional information

Accession codes: Coordinates and structure factors are deposited in the Protein Data Bank under accession codes 4UCI, 4UCL, 4UCJ, 4UCK, 4UCZ, 4UCY and 4UD0.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Paesen, G. C. et al. X-ray structure and activities of an essential Mononegavirales L-protein domain. Nat. Commun. 6:8749 doi: 10.1038/ncomms9749 (2015).