Structural insight into molecular inhibitory mechanism of InsP₆ on African Swine Fever Virus mRNA-decapping enzyme g5Rp

Yan Yang^a, Changhui Zhang^a, Li Li^a, Xuehui Li^a, Xin Yang^b, Yao Zhao^c, Cheng Chen^d,
Wei Wang^b, Zhihui Zhong^a, Cheng Yang^a, Zhen Huang^e, Dan Su^{a,f,*}

^a State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University and Collaborative Innovation Center of Biotherapy, Chengdu, Sichuan 610041, P.R. China; ^b Institute of Life Sciences, Chongqing Medical University, Chongqing 400016, P.R. China; ^c Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology, ShanghaiTech University, Shanghai, China. ^d School of Life Sciences, Tianjin University, Tianjin 300072, P.R. China; ^e College of Life Sciences, Sichuan University, Chengdu, 610041, China. ^f Tianjin International Joint Academy of Biotechnology and Medicine, Tianjin 300457, P.R. China

* Address correspondence to Dan Su

Email: sudan@scu.edu.cn.

Running title: Structural and functional insights into g5Rp

Keywords: g5Rp, ASFV, mRNA decapping enzyme, Nudix hydrolases, InsP₆

Abstract

Removal of 5' cap on cellular mRNAs by the African Swine Fever Virus (ASFV) decapping enzyme g5R protein (g5Rp) is beneficial to viral gene expression during the early stages of infection. As the only nucleoside diphosphate linked moiety X (Nudix) decapping enzyme encoded in the ASFV genome, g5Rp works in both the degradation of cellular mRNA and hydrolyzation of the diphosphoinositol polyphosphates. Here, we report the structures of dimeric g5Rp and its complex with inositol hexakisphosphate (InsP₆). The two g5Rp protomers interact head-to-head to form a dimer, and the dimeric interface is formed by extensive polar and nonpolar interactions. Each protomer composed a unique N-terminal helical domain and C-terminal classic Nudix domain. As a mRNA decapping enzyme, we identified key residues, including K⁸, K⁹⁴, K⁹⁵, K⁹⁸, K¹⁷⁵, R²²¹, and K²⁴³ located on the substrate RNA binding interfaces of g5Rp, are important to RNA binding and decapping enzyme activity. Furthermore, we identified that the g5Rp-mediated mRNA decapping was inhibited by the InsP₆. The g5Rp–InsP₆ complex structure showed that the InsP₆ molecules occupy the same regions that primarily mediate g5Rp-RNA interaction, elucidating the roles of InsP₆ in the regulation of the viral decapping activity of g5Rp in mRNA degradation. Collectively, these results provide the structural basis of interaction between RNA and g5Rp and highlight the inhibitory mechanism of InsP₆ on mRNA decapping by g5Rp.

Introduction

African swine fever virus (ASFV), which is an enveloped double-stranded DNA virus, has a genome that varies between 170 and 193 kilobase pairs with 151–167 open reading frames depending on the virus strain ^{1,2}. As the only known DNA arbovirus, ASFV is the sole member of the *Asfarviridae*, a family of African swine fever-like viruses that are relatively independent of the host cell transcriptional machinery for viral replication ^{3,4}. The ASFV infection of domestic swine can result in various disease forms, ranging from highly lethal to subclinical depending on the contributing viral and host factors ⁵. Since 2018, ASFV has spread into China and led to a high mortality rate in domestic pigs ^{6,7}. Currently, there are still no effective vaccines or specific drugs available against this particular virus ^{8,9}.

During an ASFV infection, protein synthesis in the host cell is inhibited as a result of a massive degradation of host cellular mRNAs in the cytoplasm of infected cells ^{10,11}. As part of its strategy to inhibit host cellular translation and promote viral protein synthesis instead, the virus targets the mRNAs of the host cell using specific enzymes¹². Hydrolysis of the 5' cap structure (m⁷GpppN) on eukaryotic mRNAs, a process known as decapping, is considered to be a crucial and highly regulated step in the degradation of mRNA¹³. Some viruses including ASFV and vaccinia virus (VACV) can harbor decapping enzymes for control of viral and cellular gene expression¹⁴. Two poxvirus Nudix hydrolases D9, D10 have been confirmed with intrinsic mRNA-decapping activity,

although the two decapping enzymes appear to have some differences in substrate recognition^{15,16}.

Nudix hydrolases (nucleoside diphosphate-linked moiety X) are widely present in bacteria, archaea, and eukarya, where they belong to a superfamily of hydrolytic enzymes that catalyze the cleavage of nucleoside diphosphates and the decapping of the 5' cap of mRNAs, the latter of which plays a pivotal role in mRNA metabolism^{17,18}. Mammalian cells have about 30 different genes with Nudix motifs, including Dcp2, Nudt16, and NUDT3/DIPP1, which cleaves mRNA caps in mRNA degradation by the 5'-3' decay pathway *in vivo*¹⁹⁻²¹. The mRNA-decapping enzyme g5Rp, which is the only Nudix hydrolase in ASFV, shares sequence similarity to the mRNA-decapping enzymes Dcp2 in *Schizosaccharomyces pombe* and D9 or D10 in the VACV²²⁻²⁴. However, g5Rp and its Nudix homologs D9 and D10 exhibit higher hydrolytic activity toward diphosphoinositol polyphosphates and dinucleotide polyphosphates than toward cap analogs^{25,26}. Similar to Dcp2, these Nudix hydrolases cleave the mRNA cap attached to an RNA moiety predicated that RNA binding is crucial for performing its mRNA decapping activity¹⁶. Recently, structural study has confirmed that the Nudix protein CFI_m25 has a sequence-specific RNA-binding capability²⁷. The requirement of RNA binding for majority of the Nudix decapping enzymes suggest that the members of Nudix family also belong to RNA-binding proteins.

The viral mRNA-decapping enzyme g5Rp is expressed in the endoplasmic reticulum

from the early stage of ASFV infection and accumulates throughout the infection process, playing an essential role in the machinery assembly of mRNA regulation and translation initiation²³. Likes other members of Nudix family, g5Rp has a broader range of nucleotide substrate specificity, including that for a variety of guanine and adenine nucleotides and dinucleotide polyphosphates²⁵. Generally, g5Rp has two distinct enzymatic activities *in vitro* (viz., diadenosine hexaphosphate hydrolase activity and mRNA-decapping activity), implying that it plays roles in viral membrane morphogenesis and mRNA regulation during viral infections²⁸. In light of these biochemical observations, the elucidation of the structure of g5Rp is of fundamental importance toward our understanding of the molecular mechanisms through which it degrades cellular RNAs and regulates viral gene expression.

Here, we report the crystal structure of g5Rp and its complex structure with InsP₆. Combing with biochemical experiments, the dimeric form of g5Rp and three RNA binding surfaces on each protomer are critical to substrate RNA binding of g5Rp. The g5Rp-InsP₆ complex structure shows that two of RNA binding surfaces are occupied by InsP₆ indicating that InsP₆ may play a role in its ability to inhibit g5Rp-RNA binding activity. Meanwhile, we evaluate the inhibitor effect of InsP₆ on the mRNA decapping enzyme activity of g5Rp. Therefore, we proposed that such inhibition could be caused by the competition of InsP₆ with substrate mRNA for binding to the g5Rp. Furthermore, we show detail how InsP₆ inhibits g5Rp activity by occupied the RNA-binding interfaces on

g5Rp, thereby competitively blocking the binding of substrate mRNA to the enzyme.

These results suggest InsP₆ or its structural analogs may be involved in the manipulation of the mRNA decapping process during viral infections and provide an essential structural basis for the development of ASFV chemotherapies in the future.

Results

Characterization of recombinant ASFV g5Rp protein.

Recombinant wild type ASFV g5Rp (residues 1–250) was expressed in *Escherichia coli* (*E. coli*) with a N-terminal His₆-tag. The purified g5Rp protein was eluted from a Superdex 200 column (GE Healthcare) with a major elution volume of 15.6 mL, indicating an approximate molecular weight of 32.1 kDa (Fig. S1A). The fractions were further analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS/PAGE), showing a g5Rp band of 29.9 kDa (Fig. S1B). Crosslinking assay confirmed that the g5Rp protein exists as a stable homodimer in solution (Fig. S1C).

We first characterized the nucleic acid binding ability of g5Rp with different length of single-stranded RNA (12-mer, 26-mer ssRNA). Electrophoretic Mobility Shift Assays (EMSAs) results demonstrated that g5Rp binds ssRNA (0.25 μM) at the lowest concentration of 0.5 μM (Fig. 1A and B). Furthermore, we measured the binding affinity of wild type (WT) g5Rp for ssRNA by using surface plasmon resonance (SPR). (Fig. 1D and E). The enzyme exhibited a stronger binding affinity to ssRNAs with equilibrium dissociation constant 12-mer K_D = 164.0 nM, 26-mer K_D = 44.8 nM. The kinetic analysis

of the binding experiments is shown in Supplementary Table 1. These results indicate that g5Rp possesses a higher affinity with long ssRNA. Next, we re-evaluated the decapping activity of recombinant g5Rp by incubating the protein with a ^{32}P -cap-labeled RNA substrate in a reaction. The products of the reaction were resolved by polyethyleneimine (PEI)-cellulose thin layer chromatography (TLC) and detected by autoradiography²³. As shown in Fig 1C, the recombinant g5Rp in the decapping reaction released 7-methylguanosine cap (m^7GDP) product efficiently. On the contrary, the ^{32}P -cap-labeled RNA substrate as control remained at the origin of the plate. These results suggest that the recombinant g5Rp possesses efficient mRNA decapping enzyme activity.

Overview of the ASFV g5Rp structure

To investigate structural insights into the catalytic mechanism of g5Rp, we determined its dimeric structure by single-wavelength anomalous diffraction (SAD) phases using selenomethionine-labeled protein (supplementary result). As shown in Fig. 2A, the g5Rp dimer composed two protomers that each adopts a baseball glove shape with distinct N-terminal helical domain and C-terminal Nudix domain (Fig. 2B). The helical domain (residues 36–124) forms a globin-fold-like feature composed of six α -helices ($\alpha 1$ – $\alpha 6$) that connect to the Nudix domain by two hinge linkers (linker I: residues 32–35; linker II: residues 119–139). The Nudix domain (residues 1–35 and 125–250) consists of a central curved β -sheet ($\beta 1$, $\beta 2$, $\beta 3$, $\beta 4$) surrounded by five α -helices ($\alpha 7$ – $\alpha 11$) and several loops, thereby forming a classic α – β – α sandwich structure. Linker II splits the top of the β -

sheet to connect $\alpha 6$ and $\alpha 7$ together (Fig. 2C). The Nudix motif located in center of the Nudix domain is a highly conserved comprising the loop-helix-loop architecture that contains the Nudix signature sequence extending from residues ¹³²GKPKEDESDLTCAIREFEEETGI¹⁵⁴ in g5Rp (Fig. 2D). The sequence of the g5Rp Nudix motif matches the classic pattern of the Nudix motif in the Nudix hydrolase superfamily; that is, GX₅EX₇REUXEEXGU, where X is any residue and U is Ile, Leu, or Val ^{29,30}. Using the Dali server³¹, we compared the structure of the g5Rp with that of other proteins in the Protein Data Bank (PDB), whereupon 46 structures were found to be likely homologous to the enzyme, with Z-scores in the range of 8–20 (Supplementary, Table 2). However, all the listed protein structures only shared high architectural similarity with the Nudix domain of g5Rp. Therefore, a search on the Dali server was carried out for the N-terminal domain alone, whereupon no homologous structure with a Z-score above 4 was found, suggesting that the N-terminus of g5Rp adopts a novel fold. Comparing with the structures of Dcp2 in a number of different conformations, the g5Rp shows a unique N-terminal globin-fold-like domain (Fig. S1D and E).

A previously study showed that the N-terminus of g5Rp is the major mediator of RNA interaction ²⁸. However, the positively charged surface of g5Rp structure overlaps both the N- and C-terminal region of protein that may exhibit RNA-binding activity (Fig. S2A). We proposed that both positively charged regions could contribute to g5Rp-RNA interaction. To test the hypothesis, we measured the binding of the truncation variants

g5Rp Δ C (helical domain, residues 36-124) and g5Rp Δ N (Nudix domain, connecting residues 1-35 and 125-250 directly) to single-stranded ssRNA (12-mer, 26-mer), respectively. Our EMSA results showed that both helical domain (g5Rp Δ C) and Nudix domain (g5Rp Δ N) of g5Rp are involved in ssRNA interaction (Fig. S2B). The helical domain exhibited K_D values of 39.0 and 50.7 nM for the surface-immobilized 12- and 26-mer ssRNAs measured by SPR, respectively (Fig. S2C and D). By contrast, the K_D values of g5Rp wild-type for the ssRNAs (12-mer K_D = 164.0nM, 26-mer K_D = 44.8nM) are slightly lower than that of the helical domain with ssRNAs, which indicating that both full-length and truncated g5Rp associated with RNA with high affinity.

The dimeric structure of g5Rp

When recombinant g5Rp was subject to gel filtration chromatogram to estimate molecular weight, it migrated as a single population of molecules at a molecular mass consistent with a monomer. However, g5Rp dimerization was consistent with cross-linking experiments (Fig. S1A and C). To obtain more information about the interfaces and likely biological assemblies of g5Rp, we analyzed its structure using the PDB-related interactive tool Proteins, Interfaces, Structures and Assemblies (PDBePISA)³². The results suggested that the g5Rp forms a stable symmetric dimer in crystal packing. The stable dimer was composed of two protomers (A and B) positioned in an orientation similar to two baseball gloves stuck together back-to-back (Fig. 2A). The dimer interfaces were stabilized mainly by hydrophobic interactions. Furthermore, a network of hydrogen bonds conferred

additional stability to the interface. One interface was composed of four α -helices ($\alpha 3$ and $\alpha 4$ from each A and B protomer) from the N-terminus of each protomer. Residues Ile⁶⁴, Asn⁶⁵, Arg⁶⁷, Leu⁶⁸, Leu⁶⁹, Lys⁷¹, Thr⁷², Arg⁷⁷, Tyr⁸⁰, His⁸¹ and Ile⁸⁴ located in helices $\alpha 3$ and $\alpha 4$ played pivotal roles in stabilizing the dimeric form of the protein. The other dimer interface was located at the linker II portion of the C-terminus, and one solvent contact surface of the C-terminal domain. Residues Ile¹¹⁶, Asn¹¹⁷, Ala¹¹⁹, Lys¹²⁰, Gly¹²¹, Ser¹²², Gly¹²³, and Thr¹²⁴ located on linker II, and residues Asn¹⁹⁵, Met¹⁹⁶, Leu¹⁹⁸, Ser¹⁹⁹, Leu²⁰⁰, Gln²⁰¹, Ile²⁰⁶, Ser²¹⁰, Lys²¹¹, Gln²¹⁵, Glu²¹⁸ and Ala²¹⁹ at the C-terminus, forming hydrogen bonds in the dimer interface, with further contributions from a hydrophobic patch composed of Ile²⁰⁶, Ile²⁰⁹, Phe²²², and Ile²²³ (Fig. S3A). To determine the multimeric state of g5Rp in solution and to examine which of its termini is critical for its dimerization, we measured the multimerization of two g5Rp truncation variants (g5Rp Δ N and g5Rp Δ C) using cross-linking experiments. The results showed that the wild type, N-terminus, and C-terminus of g5Rp all formed a dimeric conformation in solution (Fig. S3B and C). The g5Rp mutant I84A/I116A/L200A/I206A/F222A that prepared to dissociate the dimeric form of g5Rp was successful to alter a monomeric state, even the dimeric total buried area of 3050 Å². Wild-type g5Rp and mutants were subjected to gel-filtration chromatography resulting that the mutant I84A/I116A/L200A/I206A/F222A has a larger retention volume, corresponding to a smaller molecular weight (Fig. S4A). The protein cross-linking experiment showed that the dimeric conformation was significantly reduced in solution of

mutant (Fig. S4B). The ssRNA binding ability of monomeric mutant has been measured by SPR and EMSA. Various concentration of monomeric mutant were passed over immobilized ssRNA. The resultant sensorgrams are shown in (Fig. S4C and D), and kinetic analysis is shown in Supplementary Table 1. EMSA data is shown in Fig. S4E. Both measurements produced consistent results indicating that the g5Rp mutant I84A/I116A/L200A/I206A/F222A partially impaired the RNA binding ability. Therefore, we proposed that the dimeric g5Rp is preferred to efficient RNA-binding.

Structure of the g5Rp-InsP₆ complex

The g5Rp was originally characterized through its ability to dephosphorylate 5-PP-InsP₅ (InsP₇) to produce InsP₆²⁵. We were surprised to find a tight interaction between InsP₆ and g5Rp by microscale thermophoresis (MST) (Fig. 3A). To gain insight into the molecular basis of the interaction, we determined the crystal structure of the g5Rp-InsP₆ complex and found that each asymmetric unit contained one g5Rp-InsP₆ complex in space group *P*4₁22. PDBePISA analysis revealed that an identical dimeric conformation exists in g5Rp-InsP₆ complex structure (Fig. 3B). Two InsP₆ molecules were situated on the edge of the β 1 strand of each g5Rp protomer through interactions with residues Gln⁶, Lys⁸, and Lys¹³³ (Fig. 3C). Due to the two-fold symmetry in the crystal, each of the g5Rp protomers shared two InsP₆ molecules (InsP₆ and InsP₆^{asym}) with its neighboring g5Rp protomer in the crystal lattice. Besides the InsP₆ binding on the β 1 strand located on the edge of Nudix domain, an extra InsP₆ molecule from the neighboring molecule also

interacted with g5Rp through residues Lys⁹⁴ and Lys⁹⁸ on $\alpha 5$ helix in helical domain (Fig. 3C). In this way, each InsP₆ molecule is surrounded by four Lys residues in complex structure. The solvent-accessible surface of the InsP₆-binding region of g5Rp was calculated according to the electrostatic potential. It was apparent that both InsP₆ molecules were situated on the highly positively charged area located in the protein cleft between helical domain and Nudix domain of g5Rp (Fig. S5A). The local conformational changes of g5Rp in the complex structure induced by its interaction with InsP₆ are illustrated in Fig. S5B. In the complex structure, the $\beta 1$, $\beta 3$, and $\beta 5$ strands located in the C-terminal region had moved closer to the helical domain, and $\alpha 2$ was pushed away from the InsP₆-binding sites. These changes rendered the g5Rp conformation more stable in the complex.

To assess their relative importance on g5Rp-InsP₆ interaction, amino acid residues involved in InsP₆ binding pockets were substituted by single point mutation (Q6A, K8A, K94A, K98A, K133A). Each mutant was tested their binding affinity of InsP₆ by MST, respectively. Fig. 3D, and Fig. S6 showed that mutants resulted in a notable decrease in g5Rp-InsP₆ interaction, what's more, the quintuple mutant Q6A/K8A/K94A/K98A/K133A totally lost the binding ability with InsP₆. Taken together, the mutagenesis work indicates that positively charged residues Lys8, Lys94, Lys98, Lys133 forming a cluster to mediate the g5Rp-InsP₆ interaction.

Analysis of residues involved in g5Rp-RNA interfaces

To characterize RNA binding surface on g5Rp, we analyzed the electrostatic potential at the surface of the g5Rp, which indicated that three highly positively charged areas (Area I-III) may play roles in g5Rp-RNA interaction (Fig. S2A). Area I is located on the helical domain, containing residues Lys94, Lys95, Lys98, Arg100, Lys101 located on helix $\alpha 5$. Area II compose of residues Lys8, Lys131, Lys133, Lys135, Arg146, Lys175, Lys179, and His180 mostly located on the β -strand 1,3, which close to the Nudix motif; Area III is located the very end of C-terminus of g5Rp, comprising residues Arg²²¹, Lys²²⁵, Arg²²⁶, Lys²⁴³, and Lys²⁴⁷ on helices $\alpha 10$ and $\alpha 11$ (Fig. 4A). To identify the mRNA binding surfaces on the g5Rp further, the residues mentioned above located in three positively charged areas of g5Rp were mutated respectively. The EMSA pattern showed that some mutants reduce the RNA binding affinity of g5Rp. Specifically, residues Lys⁸, Lys⁹⁴, Lys⁹⁵, Lys⁹⁸, Lys¹³¹, Lys¹³³, Lys¹⁷⁵, Arg²²¹, and Lys²⁴³ were critical for single stranded RNA binding, albeit with different efficiencies (Fig. 4B and C), implying that the g5Rp–RNA interaction interfaces are mainly located at area I, II, and III. These results also agree with our hypothesis that residues Lys⁸, Lys⁹⁴, Lys⁹⁸, Lys¹³³ of g5Rp involved in both RNA and InsP₆ interaction.

We further explore that these key residues were responsible for cap cleavage in a manner dependent on the RNA moiety interaction. Mutant proteins including Q6E/K8E, K94E, K95E, K98E, K175E, R221D, and K243E were expressed and purified. Consistent with our previous data, incubation of the ³²P-cap-labeled RNA substrate with wild-type

g5Rp resulted in cap cleavage, as observed by m⁷GDP release. When equivalent amounts of the mutant proteins of the g5Rp were included in the decapping reaction, the amount of m⁷GDP released is reduced variously in each lane. Mutant K95E decrease the decapping activity almost 50% (Fig. 4D and E), indicating that these residues of g5Rp play pivotal role in mRNA decapping by interacting with substrates mRNA.

Residues Gly¹³², Lys¹³³, and Glu¹⁴⁷ in Nudix motif impact the decapping activity

The Nudix motif of hydrolases contains crucial residues involved in catalytic activity. However, the residues in the catalytic pocket of g5Rp are still elusive from view of structure. To elucidate the function of the key residues in g5Rp, three substrate-binding structures from the Nudix superfamily were selected to identify homologous domains with high similarity at the potential catalytic pockets (Fig. S7A, B and C), as shown in Supplementary Table 2 (viz., Ap4A hydrolase (*Aquifer aeolicus*, PDB: 3I7V)³³, nudix hydrolase DR1025 (*Deinococcus radiodurans*, PDB: 1SZ3), and MTH1 (*Mus musculus*, PDB: 5MZE)^{34,35}, (all belonging to the Nudix superfamily). Superposition of the C-terminus of g5Rp with that of MTH1, Ap4A hydrolase, and Nudix hydrolase DR1025 resulted in Cα backbone root-mean-square deviation values of 0.50, 3.08, and 5.6 Å, respectively, despite the low sequence identities among these proteins (Fig. S7D). Therefore, the potential substrate-binding site of g5Rp was proposed on the basis of the superpositions of these substrate-binding protein structures in the Nudix superfamily. Residues Gly¹³², Lys¹³³, and Glu¹⁴⁷ located on the Nudix motif of g5Rp may be

responsible for cap cleavage.

To investigate the potential roles of these key residues located on the Nudix motif in the decapping activity, we substituted g5Rp residues G132, K133 and E147 from Nudix motif (Figure. 5A and S7D) with Ala, Glu, and Gln, respectively. As expected, the substitution of the residue K133 by glutamate resulted in a 30% decrease in the decapping activity. And the substitution of the residues G132, E147 by alanine and glutamine respectively, inactivate the decapping function of g5Rp completely (Fig.5B and C). No m⁷GDP was observed when the two mutants of the g5Rp were included in the decapping reaction, validating that the decapping activity was dependent on these two key residues located in Nudix hydrolase motif. Interestingly, EMSA results showed that mutant K133E reduces g5Rp's binding affinity to RNA which suggests that the loop region of Nudix motif takes part in substrates mRNA binding. (Fig.5D and E)

InsP₆ inhibits the decapping activity by disrupting g5Rp-mRNA interaction

The finding that residues located on mRNA binding regions of g5Rp are also playing pivotal roles in g5Rp-InsP₆ interaction suggests that the InsP₆ may inhibit the g5Rp decapping activity through preventing g5Rp from binding to its mRNA substrate (Fig. 6A). This prediction was confirmed by decapping and EMSA assays using recombinant g5Rp protein, InsP₆ and RNA substrates *in vitro*. The increasing amounts of InsP₆ were added to the decapping reactions to analyze its effect on RNA decapping by g5Rp. As shown in Fig. 6B, the addition of InsP₆ significantly affected g5Rp cleavage, suggesting

that the InsP₆ can inhibit the decapping activity of g5Rp *in vitro*. To investigate if this inhibitory mechanism of InsP₆ on g5Rp is due to inositol phosphate competitively inhibiting mRNA binding to the g5Rp, we further measured the competition of InsP₆ with nucleic acids for the binding to g5Rp by using EMSA. As expected, the amount of free single-stranded nucleic acids accumulated with an increasing concentration of InsP₆, demonstrating that InsP₆ interrupts the g5Rp-mRNA interaction through the directly binding to g5Rp (Fig. 6C and D). In addition, all residues involved in InsP₆ interaction in g5Rp were mutated to alanine at the same time. The quintuple mutant (Q6A/K8A/K94A/K98A/K133A) of g5Rp lost most of its ability to bind with both InsP₆ and RNA, (Fig. S6F and S8A). These mutations also significantly affected the catalytic ability of g5Rp *in vitro* (Fig. S8C and D), suggesting that InsP₆ inhibits the mRNA decapping activity of g5Rp through competing for the substrate mRNA binding surface in g5Rp.

Transient expression g5Rp decreases levels of mRNA substrates in 293T cells

The above data provide strong *in vitro* evidence for g5Rp-mRNA interaction being a critical step for the decapping enzyme process. To determine whether changes in g5Rp-mRNA interaction was directly related to the stability of cellular mRNAs *in vivo*, representative cellular mRNA (eIF4E, eIF4EA, TP53) levels were tested by RT-qPCR in cells. In 293T cells, the Flag-tagged g5Rp, and the g5Rp mutants (K8E, K94E, K95E, K98E, G132A, K133E, E147Q, K175E, R221D, and K243E) were overexpressed,

respectively. As shown in Figure 7A, the proteins of g5Rp-WT and mutants were detected by western blot. The mRNA levels of target genes (eIF4E, eIF4EA, TP53) were decreased in 293T cells overexpressing g5Rp-WT. There are not obviously mRNA levels changed in catalytic destructive mutants Q132A and E147Q. The cells overexpressed truncated version g5Rp-ΔN and mutant K95E, R221D, which mutants significant lost the RNA binding ability *in vitro*, have no effect on the mRNA levels of target genes in 293T cells. Mutants K8E, K133E which declined the RNA binding *in vitro* have varying degrees of ascent compared with the mRNA levels of g5Rp-WT group. However, the changes in mRNA level of targets genes observed in mutants K94E, K98E, K175E and K243E have not statistical difference with g5Rp-WT (Fig. 7B-D). Taken together, these results suggest that key residues K8, K95, K133, and R221 playing pivotal roles in g5Rp-RNA interaction are also important to the g5Rp related cellular RNA degradation *in vivo*.

Discussion

Given that an ASFV outbreak in China would potentially devastate the world's largest pork producer, significant efforts have been made to determine the structures and functions of essential viral proteins that may be used as targets for new anti-ASFV drugs. Several structures of ASFV-encoded enzymes and associated proteins that are involved in viral transcription and replication have been reported, including AP endonuclease ³⁶, the histone-like protein pA104R ³⁷, pS273R protease ³⁸, DNA ligase ³⁹, and dUTPase ^{40,41}. However, the structures and functions of some critical ASFV proteins remain elusive,

including those of g5Rp, a decapping enzyme that is crucial for viral infection²³. Our structures of g5Rp alone and in complex with InsP₆ provide the molecular basis for g5Rp substrate recognition and reveal that inositol phosphate was involved in the regulation of cellular mRNA degradation through direct interaction with the ASFV decapping enzyme g5Rp. Three potential RNA binding regions are identified, including a novel folding domain located on the N-terminal of g5Rp and Nudix motif on its C-terminus. More importantly, identification of the major nucleic acid-binding surfaces as well as the binding pocket of InsP₆ on g5Rp provides important structural information and a novel strategy for future anti-ASFV drug design.

To explore the nucleic acid binding properties of g5Rp, we conducted a series of nucleic acid binding experiments. Results indicating that an intact dimeric interface is efficient for g5Rp-RNA interaction. Meanwhile, both N-terminus (helical domain) and C-terminal (Nudix domain) of g5Rp are involved in ssRNA interaction. Our EMSA and SPR measurements show that the N-terminal domain of g5Rp can bind with ssRNA with equally high affinity compared with full-length. It composes of six α -helices forming a globin-fold-like domain, which is different from the traditional RNA-binding domain that prefers to adopt the alpha/beta topologies⁴²⁻⁴⁵. According to the g5Rp structure, the surface electrostatic potential characteristics of the N-terminus present a highly positive charged area on the helix α 5. The single point mutations of positively charged residues in N terminus significantly reduced the nucleic acid-binding activity of g5Rp with ssRNA

(Fig. 4*B* and *C*). Furthermore, there are two positive areas located on the C-terminus of g5Rp, including the Nudix motif, participating in the substrate RNA interaction. We mutated the two positively charged regions (K8A/K131A/K133A/K135A and R221A/K225A/R226A/K243A/R247A) located in the Nudix domain, the EMSA data showed that the nucleic acid binding ability of these two mutants was significantly reduced (Fig. S8*B*), and the figure S8*E* and S8*F* data showed that the substantially decline in capacity of K8A/K131A/K133A/K135A remove the m⁷GpppRNA cap. These results predicated that the Nudix motif of g5Rp possesses substrate-selectivity at the step of mRNA binding.

Previously studies revealed that the Nudix motif (residues 132–154) is an essential component of the α - β - α sandwich in catalytic center of g5Rp. Several of the conserved catalytic amino acids and glutamate residues (E¹⁴⁷, E¹⁴⁹, E¹⁵⁰, and E¹⁵¹) located on the α -helix of the Nudix motif of g5Rp have been found to be important for the activity of Nudix hydrolases^{23,28}. However, the function of the loop region within the Nudix motif is exclusive leading us to predicated that the loop region may contribute to binding with mRNA. Therefore, we mutated several residues in this loop region, including the mutations K133E, G132A, and examined the effects on the protein's interaction with single-stranded nucleic acids. It is interesting to find that substituted the conserved residues K¹³³, G¹³² are highly sensitivity to g5Rp-RNA interaction. Comparing with glutamate residues located on the α -helix of the Nudix motif of g5Rp involved in mRNA

cap structure interaction, residues K¹³³, G¹³² are important to bind with RNA moiety on the substrate. In this way, we provided the demonstration that the short loop in the Nudix motif is required for g5Rp–RNA interaction. Including the Nudix motif, three positively charged patches on the g5Rp surface were mapped as mRNA-binding regions. Furthermore, we also investigated the importance of the mRNA–interacting residues in g5Rp mediated decapping. The g5Rp mutants K8E, K94E, K95E, K98E, K175E and R221D, showed a strong reduction in decapping activity, demonstrating the importance of the mRNA binding residues for catalysis. Therefore, further work should be done to elucidate the structure of g5Rp in complex with mRNA.

The other important finding in this study was that InsP₆ is able to inhibit the decapping activity of g5Rp. As we know, InsP₆ is widespread in cells with diverse biological functions⁴⁶⁻⁴⁹. Here, we found that InsP₆ competes with mRNA substrates for binding to g5Rp and inhibits its decapping activity. Previously study reported that the g5Rp is a diphosphoinositol polyphosphate phosphohydrolase (DIPP), which preferentially remove the 5-β-phosphate from the InsP₇ to produce InsP₆ with unclear functional significance²⁵. Later, Susan P. and colleagues identified that g5Rp can hydrolyze the mRNA cap when tethered to an RNA moiety *in vitro*²³. Our results show InsP₆ the product of g5Rp playing the role of DIPP can directly inhibit the mRNA decapping activity of g5Rp. To illustrate the structural basis of the inhibitory mechanism of InsP₆ to the decapping activity of g5Rp, we solved the complex structure of g5Rp with InsP₆, also the enzyme-

product complex in Nudix superfamily. To our surprised, the InsP₆ is located on the mRNA binding region instead in the catalytic center of the g5Rp. Furthermore, we superposed the catalytic domain of g5Rp-InsP₆ complex with the structures of human DIPPI in complex with the substrate InsP₇^{50,51}. The visualizing result showed that the substrate InsP₇ is located in the catalytic center of DIPPI unlike the InsP₆ sitting on the edge of catalytic domain of g5Rp (Fig S9A). Therefore, the structure of g5Rp-InsP₆ complex may represent an intermediate in the release of the product of the enzymatic reaction⁵². This also explain why InsP₆ binding decrease the temperature value (B factor) in the InsP₆ binding sites compared with B factor in the same regions of the g5Rp wild-type structure (Fig. S9B). We also noticed that the InsP₆ was refined with correspondingly high B-factor that exceed the average B-factor of protein in complex. Considering the g5Rp-InsP₆ interaction with a dissociation constant K_d in the 22.5uM range, the ligand only achieves a reasonable occupancy of 70%⁵³. To avoid instance of overenthusiastic interpretation of ligand density, we test the InsP₆ binding site by using single point mutations. Residues involved in the InsP₆ binding surface of g5Rp substituted by alanine (Q6A/K8A/K94A/K98A/K133A) reduced its InsP₆-binding capacity and RNA interaction, indicating destructive InsP₆ binding site has the capability to abolish the substrate RNA binding ability of g5Rp (Fig. S8A, C and D).

Our study raises the possibility that g5Rp hydrolyzes InsP₇ to up-regulation of the level of InsP₆, which is a key regulator of g5Rp-mediated mRNA decapping during the ASFV

infection *in vivo*. Very recently, Soumyadip S. and colleagues reported that InsP₇ regulates the NUDT3-mediated mRNA decapping and also observed the phenomenon that InsP₆ inhibits mRNA decapping by NUDT3⁵⁴. There are emerging signs that the functions of InsP₆ are associated with mRNA transportation and degradation in ASFV infected cells. Further studies into the function of the InsP₆ and the regulation mechanism in the inositol-based cell signaling family during the viral infection is required.

MATERIALS AND METHODS

Cell culture. The human 293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Hyclone) supplemented with 10% fetal bovine serum (FBS) (Gibco), 100 U/mL penicillin and 100 micro g/mL streptomycin at 37 °C under a humidified 5% CO₂ atmosphere (Thermo).

Plasmid construction, protein expression, and purification. The gene encoding ASFV g5Rp (*D250R*) was synthesized and subcloned into the pSMART-1 and pcDNA3.1 respectively. The point mutants and truncation variants of g5Rp (viz., Q6E/K8E, K94E, K95E, K98E, R100D, K101E, K131E/K133E, R146G/E147Q, K175E/K179E/H180D, R221D, K225E, K243E/R247D, Q6A/K8A/K94A/K98A/K133A, K8A/K131A/K133A/K135A, R221A/K225A/R226A/K243A/R247A, I84A/I116A/L200A/I206A/F222A, g5Rp-ΔC (Helical domain, residues 36-124) and g5Rp-ΔN (Nudix domain, connecting residues 1-35 and 125-250 directly)) were generated using the Fast Mutagenesis V2 Kit (Vazyme Biotech, China). The primers used

in this study are listed in Supplementary Table 3. The recombinant plasmids were confirmed by sequencing (Sangon Biotech, China) before being introduced into *E. coli* BL21 (DE3) (Invitrogen, US) or human 293T cells. The bacterial cells were cultured in Luria Broth medium at 35°C until the optical density at 600 nm reached 0.6–0.8. Protein expression was then induced by the addition of isopropyl β -D-1-thiogalactopyranoside for 16 h at 16°C. The g5Rp molecules were purified by Ni-NTA (Qiagen, Germany) affinity chromatography, followed by heparin affinity chromatography (GE Healthcare, US). The peak fractions containing the target proteins were pooled, concentrated to 1 mL, and finally loaded onto a Superdex 75 column (GE Healthcare, US) for further purification and characterization. Selenomethionine-labeled g5Rp (SeMet-g5Rp) was then prepared using a previously described protocol⁵⁵. The purity of all proteins was above 95% on the SDS-PAGE gel.

Protein crystallization and optimization

The prepared SeMet-g5Rp was concentrated to 12 mg/mL for the crystallization trials. The crystals were grown using the hanging-drop vapor diffusion method at 16°C in a reservoir solution containing 0.1 M sodium citrate tribasic dihydrate (pH 5.8), 0.54 M magnesium formate dihydrate, and 10% (v/v) 1,2-butanediol as an additive reagent. The g5Rp–InsP₆ complexes were prepared by mixing g5Rp with InsP₆ at a stoichiometric ratio of 1:3. Then, using the hanging-drop vapor diffusion method, crystals of the complexes were grown from 1 M imidazole (pH 7.0) at 16°C. All crystals were

transferred into solutions containing 20% (v/v) glycerol prior to being frozen and stored in liquid nitrogen.

Data collection, processing, and structure determination

The single wavelength anomalous dispersion (SAD) data were collected using synchrotron radiation of 0.98 Å wavelength under cryogenic conditions (100K) at BL18U1 beamline, Shanghai Synchrotron Radiation Facility. All diffraction data sets including g5Rp-WT and complex with InsP₆ were indexed, integrated and scaled by using HKL-2000 package ⁵⁶. The selenium atoms in the asymmetric unit of SeMet-g5Rp were located and refined, and the SAD data phases were calculated and substantially improved through solvent flattening with the PHENIX program ⁵⁷. A model was built manually into the modified experimental electron density using the model-building tool Coot ⁵⁸ and then further refined in PHENIX. The model geometry was verified using the program MolProbity ⁵⁹. Molecular replacement was used to solve the structure of the g5Rp–InsP₆ complex, using Phaser in the CCP4 program suite with an initial search model of SeMet-g5Rp ⁶⁰. Structural figures were drawn using PyMOL (DeLano Scientific). The data collection and refinement statistics are shown in Table 1.

Surface plasmon resonance analysis

The SPR analyses were carried out using the BIAcore 8K system with a streptavidin-coated (SA) chip (#BR-1005-30, GE Healthcare) at 16°C. To reduce effects attributed to mass transport, low levels of biotin-labeled ssRNA (5'-

GCUUUGAUUUCGUGCAUCAUGGAGC-3'), (5'-GCUUUGAUUUCG-3') ligands (given in relative units [RU]) were immobilized on the SA chip. Given the apparent variation in RU of immobilized ligand used in different binding studies^{61,62}, the average RU of immobilized 26mer RNA and 12mer RNA in this study are approximately 200-400RU. The blank channel served as a negative control. Protein solutions with various concentrations run across the chip at a rate of 30 μ L/min, then dissociated by running buffer (20 mM HEPES, 150 mM NaCl, 3mM EDTA and 0.05% Tween 20 (pH 7.5)) for 300 s at a flow rate of 30 μ L/min. Regeneration of the sensor chips was performed for 30 s using regeneration buffer (0.5%SDS). The data on the binding of the g5Rp molecules to ssRNAs were fitted to a kinetic binding model, which determines association and dissociation constants by fitting the experimental data to a Langmuir model with 1:1 interaction model between analyte A and ligand B: the association (k_a) and dissociation (k_d) rate constants, and the affinity value ($K_D = \frac{k_d}{k_a}$) were determined.

Microscale thermophoresis

The binding between g5Rp and InsP₆ was measured by microscale thermophoresis. In brief, g5Rp was first labeled using the Monolith NT Protein Labeling Kit RED-NHS (NanoTemper Technologies), and the labeled protein was then diluted to 20 nM with buffer containing 50 mM HEPES, 300 mM NaCl, and 0.05% (v/v) Tween 20 (pH 7.0). Then, a series of concentrations of InsP₆ diluted in a buffer composed of 50 mM HEPES, 300 mM NaCl, and 0.05% (v/v) Tween 20 (pH 7.0) was added. The mixtures were loaded

into capillaries and measured at ambient temperature by using 20% LED and medium MST power in a Monolith NT.115 system (NanoTemper Technologies). The data were analyzed using NanoTemper analysis software (v.1.2.101).

Electrophoretic mobility shift assay

EMSAs were performed to determine the nucleic acid affinity of the wild-type and g5Rp mutants. The single-stranded nucleic acids used were 6-carboxyfluorescein-labeled ssRNA (5'-GCUUUGAUUUCGUGCAUCAUGGAGC-3'), 6-carboxyfluorescein-labeled ssRNA (5'-GCUUUGAUUUCG-3') (Sheng Gong, Shanghai, China). Initially, 0.25 μ M of the ssRNA was incubated on ice for 30 min together with different concentrations of wild-type in a buffer composed of 20 mM HEPES, 60 mM KCl, 0.5 mM EDTA, 0.1% Triton-X-100, 4mM DTT, 2mM MgCl₂ and 5% (v/v) glycerol (pH 7.9), the experimental concentration of mutants g5Rp is 2uM. To determine the effect of InsP₆ on nucleic acid-binding ability of g5Rp, we also used the EMSA to test the nucleic acid-binding ability of the enzyme mixed with different concentrations of InsP₆ (0.03mM, 0.06mM, 0.12mM, 0.25mM, 0.5mM, 1mM, 1.50mM, 2.00mM). All samples were incubated on ice for 30 min and then electrophoresed on 4.0 % native-PAGE gels for 45 min at a voltage of 100 V. The results were determined with a Bio-Rad ChemiDoc™ MP Imaging System (Bio-rad, USA).

Capping of the mRNA Body. For decapping assays with g5Rp, uncapped RNA was transcribed using the T7 RNA transcription kit (Vazyme) with the linear DNA (5'-

CATTATTGGCCTGAAAAAGATGATTGACAGCTATAATGATTACTACAACAAC

GAAGTTTTTCGTAAACATAAAAACC-3') as a template. The RNAs were capped in 50ul reaction system typically containing 0.4 nmol RNA, 12ul 3000Ci/mmol [α -³²P] GTP, 0.67 mM Sadenosylmethionine, 50 mM Tris-HCl pH 7.6, 2 mM MgCl₂, 6mM KCl, 1 mM DTT and 75 U Vaccinia Capping Enzyme (NEB) at 37°C for 2 h. Cap-labeled RNAs were then separated from free[α -³²P]GTP nucleotide by chromatography through a G-50 column(sigma).

mRNA Decapping Assays. All decapping experiments were carried out in a buffer containing 50 mM Tris-HCl (pH 7.0), 1 mM DTT and 2 mM MnCl₂. Generally, 100ng of wild-type or g5Rp mutants was used in each 5 μ l reaction. To determine the correlation between InsP₆ and decapping ability of g5Rp, we also test the decapping ability of the enzyme mixed with different concentrations of InsP₆ (0.03mM, 0.06mM, 0.12mM, 0.25mM, 0.5mM, 1mM, 2mM, 4mM). The decapping reaction was carried out at 37°C for 60 min, then stopped by 25 mM EDTA. The products of the reaction were separated by PEI-cellulose thin layer chromatography developed in 0.45M (NH₄)₂SO₄ and detected with autoradiography.

Western Blotting and Analysis. 200,000 cells/well were seeded in 6-well plate (Nest) and cultured for 24h. Then plasmids (2 μ g/well) were transfected into cells using Lipofectamine® 3000 (Invitrogen). After 48 h cells were lysed in RIPA (Biosharp) containing protease inhibitor Cocktail (MCE) and quantified by the Bicinchoninic Acid

(BCA) kit (Sangon). About 20 μ g protein was loaded in 12% SDS-PAGE and transferred to PVDF (0.22 μ M) membrane (Millipore). Membranes were incubated with primary antibodies including anti- β -actin (1:10,000 dilution; HuaBio) and anti-Flag (1:1,000 dilution; CST) at 4 °C overnight. After the membranes were washed three times for 10 min with TBST and incubated with secondary antibodies for 2 h at room temperature, they were washed three times again in TBST. Then the blots were detected with enhanced chemiluminescence reagents (Millipore) using MiniChemi™ (Sagecreation, China)

RNA extraction and quantitative real-time PCR

Briefly, 100,000 cells were seeded in a 12-well plate (Nest) and cultured for 24h before transfection plasmid (1 μ g/well). Total RNA was extracted using RNAiso Reagent (TaKaRa), and 500 ng of total RNA was reversely transcribed using PrimeScript™ RT Reagent Kit (Takara). Subsequently, real-time polymerase chain reaction (PCR) amplification was performed with SYBR Premix ExTag (Takara) on a QuantStudio™ 3 System (Applied Biosystems). The cycling settings were as follows: 95 °C for 30 s, then 40 cycles of 95 °C for 5 s and 60 °C for 34 s. Relative mRNA levels were determined using the $2^{-\Delta\Delta Ct}$ ($\Delta\Delta Ct = \Delta Ct(\text{test}) - \Delta Ct(\text{calibrator})$) method.

Data availability

The structure factors and atomic coordinates of apo g5Rp and g5Rp-InsP6 have been deposited in the Protein Data Bank under the PDB ID codes 7DNT and 7DNU, separately.

Acknowledgements

We thank the staff at the State Key Laboratory of Biotherapy, Sichuan University, who assisted with our research work during the period of the COVID-19 epidemic. The X-ray diffraction experiments were carried out at the Shanghai Synchrotron Radiation Facility (SSRF) at BL18U1. We also thank the beamline staff for their technical help during the data collection. We appreciate helpful advice and discussion with Dr. Xiaoyu Xue in Texas State University. This work was supported by grants from the National Key Research and Development Program of China (2017YFA0505903, 2018YFC1312300), the Special Research Fund on COVID-19 of Sichuan Province (2020YFS0010), and the Key Project on COVID-19 of West China Hospital, Sichuan University (HX-2019-nCoV-044).

Funding

This work was supported by the National Key Research and Development Program of China (2017YFA0505903), the National Natural Science Foundation of China (31370735 and 31670737 to D.S.), the Science and Technology Department of Tianjin Foundation (19YFZCSN00470), the Special Research Fund on COVID-19 of Sichuan Province (2020YFS0010), and the urgent project on COVID-19 of West China Hospital, Sichuan University (HX-2019-nCoV-044).

Conflict and interest disclosure

The authors declare that they have no conflicts of interest surrounding the contents of this article.

References

- 1 Wardley, R. C. *et al.* African Swine Fever virus. Brief review. *Archives of virology* **76**, 73-90, doi:10.1007/bf01311692 (1983).
- 2 Wang, N. *et al.* Architecture of African swine fever virus and implications for viral assembly. *Science* **366**, 640-644, doi:10.1126/science.aaz1439 (2019).
- 3 Cackett, G. *et al.* The African Swine Fever Virus Transcriptome. *Journal of virology* **94**, doi:10.1128/JVI.00119-20 (2020).
- 4 Rodriguez, J. M. & Salas, M. L. African swine fever virus transcription. *Virus research* **173**, 15-28, doi:10.1016/j.virusres.2012.09.014 (2013).
- 5 Netherton, C. L., Connell, S., Benfield, C. T. O. & Dixon, L. K. The Genetics of Life and Death: Virus-Host Interactions Underpinning Resistance to African Swine Fever, a Viral Hemorrhagic Disease. *Frontiers in genetics* **10**, 402, doi:10.3389/fgene.2019.00402 (2019).
- 6 Zhao, D. *et al.* Replication and virulence in pigs of the first African swine fever virus isolated in China. *Emerging microbes & infections* **8**, 438-447, doi:10.1080/22221751.2019.1590128 (2019).
- 7 Dixon, L. K., Stahl, K., Jori, F., Vial, L. & Pfeiffer, D. U. African Swine Fever Epidemiology and Control. *Annual review of animal biosciences*, doi:10.1146/annurev-animal-021419-083741 (2019).
- 8 Teklue, T., Sun, Y., Abid, M., Luo, Y. & Qiu, H. J. Current status and evolving approaches to African swine fever vaccine development. *Transboundary and emerging diseases*, doi:10.1111/tbed.13364 (2019).
- 9 Arabyan, E., Kotsynyan, A., Hakobyan, A. & Zakaryan, H. Antiviral agents against African swine fever virus. *Virus research* **270**, 197669, doi:10.1016/j.virusres.2019.197669 (2019).
- 10 Dixon, L. K., Chapman, D. A., Netherton, C. L. & Upton, C. African swine fever virus replication and genomics. *Virus research* **173**, 3-14, doi:10.1016/j.virusres.2012.10.020 (2013).
- 11 Sanchez, E. G., Quintas, A., Nogal, M., Castello, A. & Revilla, Y. African swine fever virus controls the host transcription and cellular machinery of protein synthesis. *Virus research* **173**, 58-75, doi:10.1016/j.virusres.2012.10.025 (2013).
- 12 Yanez, R. J., Rodriguez, J. M., Bournsnel, M., Rodriguez, J. F. & Vinuela, E. Two putative African swine fever virus helicases similar to yeast 'DEAH' pre-mRNA processing proteins and vaccinia virus ATPases D11L and D6R. *Gene* **134**, 161-174, doi:10.1016/0378-1119(93)90090-p (1993).
- 13 Mazzoni, C. & Falcone, C. mRNA stability and control of cell proliferation. *Biochemical Society transactions* **39**, 1461-1465, doi:10.1042/BST0391461 (2011).
- 14 McLennan, A. G. Decapitation: poxvirus makes RNA lose its head. *Trends in biochemical*

- sciences* **32**, 297-299, doi:10.1016/j.tibs.2007.05.001 (2007).
- 15 Parrish, S. & Moss, B. Characterization of a second vaccinia virus mRNA-decapping enzyme conserved in poxviruses. *Journal of virology* **81**, 12973-12978, doi:10.1128/JVI.01668-07 (2007).
- 16 Parrish, S., Resch, W. & Moss, B. Vaccinia virus D10 protein has mRNA decapping activity, providing a mechanism for control of host and viral gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 2139-2144, doi:10.1073/pnas.0611685104 (2007).
- 17 Grudzien-Nogalska, E. & Kiledjian, M. New insights into decapping enzymes and selective mRNA decay. *Wiley Interdiscip Rev RNA* **8**, doi:10.1002/wrna.1379 (2017).
- 18 Gao, A., Vasilyev, N., Kaushik, A., Duan, W. & Serganov, A. Principles of RNA and nucleotide discrimination by the RNA processing enzyme RppH. *Nucleic acids research* **48**, 3776-3788, doi:10.1093/nar/gkaa024 (2020).
- 19 Song, M. G., Bail, S. & Kiledjian, M. Multiple Nudix family proteins possess mRNA decapping activity. *RNA* **19**, 390-399, doi:10.1261/rna.037309.112 (2013).
- 20 Wang, Z., Jiao, X., Carr-Schmid, A. & Kiledjian, M. The hDcp2 protein is a mammalian mRNA decapping enzyme. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 12663-12668, doi:10.1073/pnas.192445599 (2002).
- 21 Grudzien-Nogalska, E., Jiao, X., Song, M. G., Hart, R. P. & Kiledjian, M. Nudt3 is an mRNA decapping enzyme that modulates cell migration. *RNA* **22**, 773-781, doi:10.1261/rna.055699.115 (2016).
- 22 Charenton, C. *et al.* Structure of the active form of Dcp1-Dcp2 decapping enzyme bound to m(7)GDP and its Edc3 activator. *Nature structural & molecular biology* **23**, 982-986, doi:10.1038/nsmb.3300 (2016).
- 23 Parrish, S., Hurchalla, M., Liu, S. W. & Moss, B. The African swine fever virus g5R protein possesses mRNA decapping activity. *Virology* **393**, 177-182, doi:10.1016/j.virol.2009.07.026 (2009).
- 24 Liu, S. W., Wyatt, L. S., Orandle, M. S., Minai, M. & Moss, B. The D10 decapping enzyme of vaccinia virus contributes to decay of cellular and viral mRNAs and to virulence in mice. *Journal of virology* **88**, 202-211, doi:10.1128/JVI.02426-13 (2014).
- 25 Cartwright, J. L. *et al.* The g5R (D250) gene of African swine fever virus encodes a Nudix hydrolase that preferentially degrades diphosphoinositol polyphosphates. *Journal of virology* **76**, 1415-1421, doi:10.1128/jvi.76.3.1415-1421.2002 (2002).
- 26 Souliere, M. F., Perreault, J. P. & Bisailon, M. Insights into the molecular determinants involved in cap recognition by the vaccinia virus D10 decapping enzyme. *Nucleic acids research* **38**, 7599-7610, doi:10.1093/nar/gkq628 (2010).
- 27 Yang, Q., Gilmartin, G. M. & Double, S. Structural basis of UGUA recognition by the Nudix protein CFI(m)25 and implications for a regulatory role in mRNA 3' processing. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 10062-10067, doi:10.1073/pnas.1000848107 (2010).
- 28 Quintas, A. *et al.* Characterization of the African Swine Fever Virus Decapping Enzyme

- during Infection. *Journal of virology* **91**, doi:10.1128/jvi.00990-17 (2017).
- 29 McLennan, A. G. The Nudix hydrolase superfamily. *Cell Mol Life Sci* **63**, 123-143, doi:10.1007/s00018-005-5386-7 (2006).
- 30 Bessman, M. J. A cryptic activity in the Nudix hydrolase superfamily. *Protein science : a publication of the Protein Society* **28**, 1494-1500, doi:10.1002/pro.3666 (2019).
- 31 Holm, L. DALI and the persistence of protein shape. *Protein Sci* **29**, 128-140, doi:10.1002/pro.3749 (2020).
- 32 Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797, doi:10.1016/j.jmb.2007.05.022 (2007).
- 33 Jeyakanthan, J. *et al.* Free and ATP-bound structures of Ap4A hydrolase from Aquifex aeolicus V5. *Acta Crystallogr D Biol Crystallogr* **66**, 116-124, doi:10.1107/S0907444909047064 (2010).
- 34 Ranatunga, W. *et al.* Structural studies of the Nudix hydrolase DR1025 from Deinococcus radiodurans and its ligand complexes. *Journal of molecular biology* **339**, 103-116, doi:10.1016/j.jmb.2004.01.065 (2004).
- 35 Narwal, M. *et al.* Crystal Structures and Inhibitor Interactions of Mouse and Dog MTH1 Reveal Species-Specific Differences in Affinity. *Biochemistry* **57**, 593-603, doi:10.1021/acs.biochem.7b01163 (2018).
- 36 Chen, Y. *et al.* A unique DNA-binding mode of African swine fever virus AP endonuclease. *Cell Discov* **6**, 13, doi:10.1038/s41421-020-0146-2 (2020).
- 37 Liu, R. *et al.* The structural basis of African swine fever virus pA104R binding to DNA and its inhibition by stilbene derivatives. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 11000-11009, doi:10.1073/pnas.1922523117 (2020).
- 38 Li, G. *et al.* Crystal structure of the African swine fever virus pS273R protease and implications for inhibitor design. *Journal of virology*, doi:10.1128/JVI.02125-19 (2020).
- 39 Chen, Y. *et al.* Structure of the error-prone DNA ligase of African swine fever virus identifies critical active site residues. *Nature communications* **10**, 387, doi:10.1038/s41467-019-08296-w (2019).
- 40 Li, G. *et al.* Structural Insight into African Swine Fever Virus dUTPase Reveals a Novel Folding Pattern in the dUTPase Family. *Journal of virology* **94**, doi:10.1128/JVI.01698-19 (2020).
- 41 Li, C. *et al.* Crystal Structure of African Swine Fever Virus dUTPase Reveals a Potential Drug Target. *mBio* **10**, doi:10.1128/mBio.02483-19 (2019).
- 42 Suryadi, J., Tran, E. J., Maxwell, E. S. & Brown, B. A., 2nd. The crystal structure of the Methanocaldococcus jannaschii multifunctional L7Ae RNA-binding protein reveals an induced-fit interaction with the box C/D RNAs. *Biochemistry* **44**, 9657-9672, doi:10.1021/bi050568q (2005).
- 43 Sidiqi, M. *et al.* Structure and RNA binding of the third KH domain of poly(C)-binding protein 1. *Nucleic acids research* **33**, 1213-1221, doi:10.1093/nar/gki265 (2005).
- 44 Kumarevel, T. *et al.* Crystal structure of activated HutP; an RNA binding protein that regulates transcription of the hut operon in Bacillus subtilis. *Structure* **12**, 1269-1280, doi:10.1016/j.str.2004.05.005 (2004).

- 45 Charron, C. *et al.* The archaeal sRNA binding protein L7Ae has a 3D structure very similar to that of its eukaryal counterpart while having a broader RNA-binding specificity. *Journal of molecular biology* **342**, 757–773, doi:10.1016/j.jmb.2004.07.046 (2004).
- 46 Menniti, F. S., Oliver, K. G., Putney, J. W., Jr. & Shears, S. B. Inositol phosphates and cell signaling: new views of InsP5 and InsP6. *Trends in biochemical sciences* **18**, 53–56, doi:10.1016/0968-0004(93)90053-p (1993).
- 47 Morrison, R. S. *et al.* Inositolhexakisphosphate (InsP6): an antagonist of fibroblast growth factor receptor binding and activity. *In vitro cellular & developmental biology. Animal* **30A**, 783–789, doi:10.1007/BF02631302 (1994).
- 48 Van der Kaay, J., Wesseling, J. & Van Haastert, P. J. Nucleus-associated phosphorylation of Ins(1,4,5)P3 to InsP6 in Dictyostelium. *The Biochemical journal* **312** (Pt 3), 911–917 (1995).
- 49 Verbsky, J. & Majerus, P. W. Increased levels of inositol hexakisphosphate (InsP6) protect HEK293 cells from tumor necrosis factor (alpha)- and Fas-induced apoptosis. *The Journal of biological chemistry* **280**, 29263–29268, doi:10.1074/jbc.M503366200 (2005).
- 50 Dollins, D. E. *et al.* Vip1 is a kinase and pyrophosphatase switch that regulates inositol diphosphate signaling. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 9356–9364, doi:10.1073/pnas.1908875117 (2020).
- 51 Thorsell, A. G. *et al.* Crystal structure of human diphosphoinositol phosphatase 1. *Proteins* **77**, 242–246, doi:10.1002/prot.22489 (2009).
- 52 De Simone, A., Aprile, F. A., Dhulesia, A., Dobson, C. M. & Vendruscolo, M. Structure of a low-population intermediate state in the release of an enzyme product. *Elife* **4**, doi:10.7554/eLife.02777 (2015).
- 53 Pozharski, E., Deller, M. C. & Rupp, B. Validation of Protein-Ligand Crystal Structure Models: Small Molecule and Peptide Ligands. *Methods Mol Biol* **1607**, 611–625, doi:10.1007/978-1-4939-7000-1_25 (2017).
- 54 Sahu, S. *et al.* InsP7 is a small-molecule regulator of NUDT3-mediated mRNA decapping and processing-body dynamics. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 19245–19253, doi:10.1073/pnas.1922284117 (2020).
- 55 Strub, M. P. *et al.* Selenomethionine and selenocysteine double labeling strategy for crystallographic phasing. *Structure* **11**, 1359–1367, doi:10.1016/j.str.2003.09.014 (2003).
- 56 Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–326 (1997).
- 57 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221, doi:10.1107/S0907444909052925 (2010).
- 58 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132, doi:10.1107/S0907444904019158 (2004).
- 59 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography* **66**, 12–21, doi:10.1107/s0907444909042073 (2010).
- 60 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D*

- Biol Crystallogr* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).
- 61 Chen, H. *et al.* Mass spectroscopic characterization of the coronavirus infectious bronchitis virus nucleoprotein and elucidation of the role of phosphorylation in RNA binding by using surface plasmon resonance. *Journal of virology* **79**, 1164-1179, doi:10.1128/JVI.79.2.1164-1179.2005 (2005).
- 62 Katsamba, P. S., Park, S. & Laird-Offringa, I. A. Kinetic studies of RNA-protein interactions using surface plasmon resonance. *Methods* **26**, 95-104, doi:10.1016/S1046-2023(02)00012-9 (2002).

Tables

TABLE 1. Data collection and refinement statistics

	g5Rp	g5Rp–InsP ₆
<i>Data collection</i>		
Resolution (Å)	2.50	2.25
Space group	<i>P</i> 2 ₁ 2 ₁ 2	<i>P</i> 4 ₁ 22
Unit-cell parameters (Å, °)	<i>a</i> = 56.1, <i>b</i> = 105.1, <i>c</i> = 49.5, α = β = γ = 90	<i>a</i> = 48.4, <i>b</i> = 48.4, <i>c</i> = 220.1 α = β = γ = 90
Resolution (Å)	50.00–2.5 (2.54–2.5)	50.00–2.25 (2.29–2.25)
<i>R</i> _{merge} [†] (%)	22.1 (59.3)	6.7 (53.0)
<i>R</i> _{pim} [‡] (%)	4.1 (15.7)	1.7 (14.7)
Average <i>I</i> /σ (<i>I</i>)	20 (2.1)	77.5 (12.4)
No. of observed reflections	17238 (385)	13490 (668)
No. of unique reflections	9721 (248)	13490 (668)
Completeness (%)	91.7 (48.80)	99.90 (99.70)
Matthews coefficient (Å ³ Da ^{−1})	2.32	2.32
Solvent content (%)	47	47
CC _{1/2}	0.97(0.92)	1.00(0.98)
Molecules per asymmetric unit	1	1
<i>Refinement</i>		
Resolution (Å)	50.00–2.50	50.00–2.25
<i>R</i> _{work} / <i>R</i> _{free}	0.2037/0.2744	0.1724/0.2307
Ramachandran favored (%)	93.31	96.73
Ramachandran outliers (%)	0.00	0.00
<i>Average B factor (Å²)</i>		
Protein	59.74	46.61
Water	48.7	46.12
Ligand		148.4
<i>Number of atoms</i>		
Protein	2026	2111
Water	28	106
Wilson B-value	43.01	37.45
<i>RMSDs</i>		
Bond lengths (Å)	0.01	0.008
Bond angles (°)	1.186	0.925

*Numbers in the brackets are for the highest resolution shell.

† $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is an individual intensity measurement and $\langle I(hkl) \rangle$ is the average intensity for all i reflections.

‡ R_{pim} is approximated by multiplying the R_{merge} value by the factor $[1/(N - 1)]^{1/2}$, where N is the overall redundancy of the data set.

RMSDs, root-mean-square deviations.

FIGURES

FIG 1 Characterization of the African swine fever virus decapping enzyme g5Rp. (A-B) The binding abilities of g5Rp to 12-mer and 26-mer ssRNA were determined by EMSA, in the reaction, the protein concentration is marked on the figure, and the nucleic acid concentration is 0.25 μM . **(C)** The decapping activity of g5Rp. **(D-E)** The binding abilities of g5Rp to 12-mer and 26-mer ssRNA were determined by surface plasmon resonance.

FIG 2 Structure of the African swine fever virus decapping enzyme g5Rp. (A) The dimeric structure of the African swine fever virus decapping enzyme g5Rp. The dimer consists of protomers A and B in a back-to-back orientation. Protomer A is colored orange and protomer B is colored pale green. (B) “Boxing glove” arrangement of the overall structure of g5Rp. The division of the domain is based on the g5Rp structure, where each

domain is color coded, with the helical domain in orange and the Nudix domain in light blue (containing Nudix motif, in green). The N-terminus and C-terminus domains are marked by black box. **(C)** The details structure information of g5Rp. The 11 α -helices, 4 β -sheets, and loops in the structure. The loops are colored in white, except for linker I (in magenta) and linker II (in cyan). **(D)** The details of Nudix motif in g5Rp. The Nudix motif is colored by green.

FIG 3 Structure of the g5Rp–InsP₆ complex. (A) The binding ability between g5Rp with InsP₆ is measured by MST. The dissociation constant was calculated from three independent replicates (shown as mean ± s.d.) (B) The dimer of g5Rp in complex with InsP₆. (C) Details of g5Rp in complex with InsP₆. The electron density of InsP₆ is shown as the 2Fo–Fc map contoured at 0.8 σ and generated with InsP₆ omitted. (D) The binding ability between various g5Rp mutants and InsP₆ is measured by MST. The dissociation constant between g5Rp mutants with InsP₆ were calculated from three independent replicates (shown as mean ± s.d.). InsP₆ is shown as a stick model.

FIG 4 The residues involved in g5Rp-mRNA interfaces and decapping activity.

(A) The mutation sites are located in the positive charged regions of g5Rp. (B-C) The binding abilities of g5Rp mutants to 12-mer and 26-mer ssRNA were determined by EMSA, in the reaction, the protein concentration is 2.00uM, and the nucleic acid concentration is 0.25uM, ☆ represents the precipitation in the gel (D) The decapping activity of g5Rp mutants. (E) The Semi-quantitative of m⁷GDP by Image J and drawing by graphpad prism8. (mean ± SD, n ≥ 3, *P < 0.05, **P < 0.01, unpaired t test).

FIG 5 The Nudix motif of g5Rp involved in mRNA binding and decapping. (A) The details of nudix motif in g5Rp. The Nudix motif is colored by green. The mutant residues are shown as sticks. (B) The decapping activity of g5Rp mutants. (C) The Semi-quantitative of m⁷GDP by Image J and drawing by GraphPad prism8. (mean ± SD, n ≥ 3, *P < 0.05, ***P < 0.001, unpaired t test). (D-E) The binding abilities of g5Rp mutants to 12-mer and 26-mer ssRNAs were determined by EMSA, in this reaction, the protein experimental concentration is 2.00uM, and the nucleic acid experimental concentration is 0.25uM, ☆ represents the precipitation in the gel.

FIG 6 The influence of InsP₆ on mRNA binding and decapping activity of g5Rp. (A)

The surface charge of g5Rp in complex with InsP₆. **(B)** The influence of InsP₆ on decapping

activity of g5Rp. **(C-D)** The influence of InsP₆ on binding abilities of g5Rp with 12-mer and

26-mer ssRNA were determined by EMSA, respectively. The different concentrations of

InsP₆ with g5Rp in the reaction system is indicated above of the gel, the concentration of

protein and the nucleic acid is 2.00 μM and 0.25 μM, respectively.

FIG 7 The influence of g5Rp and mutants on mRNA substrates level in 293T cells. (A)

The expression of various g5Rp mutants in 293T cells were analyzed by western blot. **(B)**

The mRNA levels of eIF4E in 293T cells when a series of g5Rp mutants overexpressed.

(C)The mRNA level of eIF4A in 293T cells when a series of g5Rp mutants overexpressed.

(D) The mRNA level of TP53 in 293T cells when a series of g5Rp mutants overexpressed.

(mean ± SD, n ≥ 3, *P < 0.05, **P < 0.01, unpaired t test).