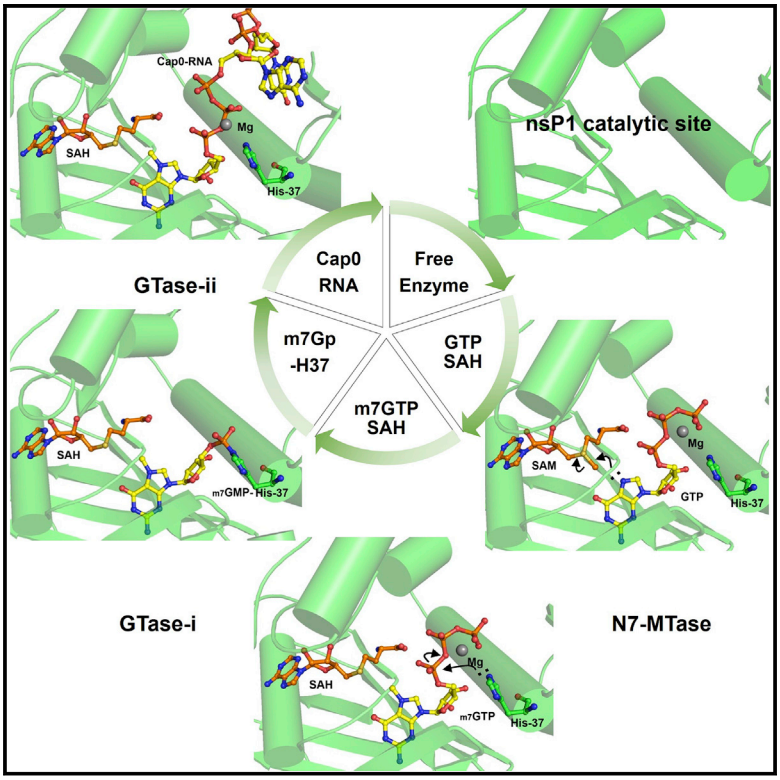
Cell Reports

Article

Molecular basis of specific viral RNA recognition and 5′-end capping by the Chikungunya virus nsP1

Graphical abstract



Highlights

- Multiple structures of CHIKV nsP1 during viral RNA 5′-cap formation are presented
- Molecular basis of viral RNA 5′-end capping by nsP1 is elucidated
- nsP1 recognizes the 5′-end A_1U_2 sequence viral RNA of preference
- The conserved 5′-end A_1U_2 is essential for viral RNA replication

Authors

Kuo Zhang, Michelle Cheok Yien Law, Trinh Mai Nguyen, ..., Yee-Song Law, Lak Shin Jeong, Dahai Luo

Correspondence

luodahai@ntu.edu.sg

In brief

Zhang et al. report the high-resolution structures of nsP1 in complex with m7GTP/SAH, m7GMP, and Cap0-viral RNA representing the key steps during viral RNA 5′-cap formation by cryo-EM. These structures elucidate the molecular basis of specific viral RNA 5′-end capping by the Chikungunya virus nsP1.

Zhang et al., 2022, Cell Reports *40*, 111133
July 26, 2022
https://doi.org/10.1016/j.celrep.2022.111133

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Molecular basis of specific viral RNA recognition and 5′-end capping by the Chikungunya virus nsP1

Kuo Zhang,[1,2,5,6] Michelle Cheok Yien Law,[1,2,6] Trinh Mai Nguyen,[1,2] Yaw Bia Tan,[1,2] Melissa Wirawan,[1,2] Yee-Song Law,[1,2] Lak Shin Jeong,[4] and Dahai Luo[1,2,3,7,*]

[1]Lee Kong Chian School of Medicine, Nanyang Technological University, EMB 03-07, 59 Nanyang Drive, Singapore 636921, Singapore
[2]NTU Institute of Structural Biology, Nanyang Technological University, EMB 06-01, 59 Nanyang Drive, Singapore 636921, Singapore
[3]School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 636921, Singapore
[4]College of Pharmacy, Seoul National University, Seoul, South Korea
[5]School of Medicine, Southern University of Science and Technology, Shenzhen, Guangdong 518055, China
[6]These authors contributed equally
[7]Lead contact
*Correspondence: luodahai@ntu.edu.sg
https://doi.org/10.1016/j.celrep.2022.111133

SUMMARY

Many viruses encode RNA-modifying enzymes to edit the 5′ end of viral RNA to mimic the cellular mRNA for effective protein translation, genome replication, and evasion of the host defense mechanisms. Alphavirus nsP1 synthesizes the 5′ end Cap-0 structure of viral RNAs. However, the molecular basis of the capping process remains unclear. We determine high-resolution cryoelectron microscopy (cryo-EM) structures of Chikungunya virus nsP1 in complex with m7GTP/SAH, covalently attached m7GMP, and Cap-0 viral RNA. These structures reveal details of viral-RNA-capping reactions and uncover a sequence-specific virus RNA-recognition pattern that, in turn, regulates viral-RNA-capping efficiency to ensure optimal genome replication and subgenomic RNA transcription. This sequence-specific enzyme-RNA pairing is conserved across all alphaviruses.

INTRODUCTION

In eukaryotic cells, 5′-cap structure is one of the most conserved modifications of mRNA during transcription (Shatkin, 1976). The cap structure at the 5′ end of mRNA plays many essential roles during mRNA biogenesis and functioning, such as guiding and promoting correct pre-mRNA splicing, exporting mature mRNA, facilitating efficient translation, and regulating mRNA stability (Ramanathan et al., 2016; Shuman, 2001). The most common mRNA cap is N7-methylated guanosine linked to the first nucleotide of the RNA via a 5′–5′ triphosphate linker. This 5′ m7G cap is also referred to as Cap-0, as the mRNA can be further methylated at the 2′-OH positions of the +1 nucleotide and additionally the +2 nucleotide to generate the hypermethylated cap structures, named Cap-1 and Cap-2, respectively (Banerjee, 1980; Ghosh and Lima, 2010; Ramanathan et al., 2016). To generate the 5′ m7G cap, pre-mRNA undergoes sequential enzymatic reactions: (1) the RNA triphosphatase removes 5′ γ-phosphate of pre-mRNA; (2) the RNA guanylyltransferase (GTase) prepares a covalent enzyme-guanylate intermediate using a GTP as the substrate; (3) the GTase then transfers the 5′-phosphoguanosine (GMP) to the 5′-diphosphate mRNA to form the G0pppRNA structure; and (4) the RNA (guanine-N7)-methyltransferase (N7-MTase) then methylates the 0-guanine at its N7 position using S-adenosyl-L-methionine (SAM) as the methyl donor. Finally, RNA 2′-O-ribose methyltransferases (2′-O-MTases) catalyze further methylation on the ribose 2′-hydroxyl (2′-O) of the first and the second nucleotides (Ghosh and Lima, 2010; Kyrieleis et al., 2014; Shuman, 2001).

Being obligatory intracellular parasites, eukaryotic viruses have evolved many strategies to produce viral RNA with the same cap structures as cellular mRNA for effective protein translation, genome replication, and evasion of the host defense mechanisms (Picard-Jean et al., 2013). Decroly et al. reviewed the diverse viral-RNA-capping mechanisms and showed that many viruses encode RNA-capping enzymes to synthesize the 5′ cap of their transcripts following either "conventional" or "unconventional" mRNA-capping pathways (Decroly et al., 2011). Unconventional mRNA-capping mechanisms commonly require different viral enzymes that are unique to specific virus families. Therefore, the viral-RNA-capping process is an attractive target for antiviral drug development. Recently, it was discovered that severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) makes use of the nidovirus RdRP-associated nucleotidyltransferase (NiRAN) domain of RNA polymerase nsp12 as the RNA GTase for genome and viral mRNA capping (Walker et al., 2021; Yan et al., 2021). It has been shown that remdesivir triphosphate inhibits both the GTase as well as RNA polymerase

activities (Walker et al., 2021). Alphavirus contains a single-strand (+) RNA genome that encodes two reading frames: the first reading frame encodes four non-structural proteins, nsP1–nsP4, to form the RNA-replication complex, and the second transcript is made through a subgenomic promoter to produce the structural proteins. Both genomic and subgenomic RNAs contain a 5′ Cap-0 structure that is produced by the viral nsP1–nsP2 capping machinery. nsP1 has both N7-guanine-methyltransferase (MTase) and GTase activities (Ahola and Kaariainen, 1995; Ahola et al., 1997), and nsP2 hydrolyses the β–γ-phosphate bond at the 5′ triphosphate end through the N-terminal RTPase/helicase domain (Law et al., 2019; Vasiljeva et al., 2000). Cryoelectron microscopy (cryo-EM) structures of Chikungunya virus (CHIKV) nsP1 have been reported recently (Jones et al., 2021; Zhang et al., 2021). Twelve copies of nsP1 organize into a dodecameric ring structure. The dual-functional MTase/GTase catalytic domain forms the upper portion of the nsP1 ring (upper ring). The lower ring is responsible for oligomerization and membrane association. nsP1 catalyzes the Cap-0 formation differently from the eukaryotic mRNA-capping machinery: the N7-MTase first adds the methyl group to the N7 of GTP, and the GTase then transfers the m7GMP moiety to the 5′ diphosphate viral RNA (Ahola and Kaariainen, 1995; Ahola and Karlin, 2015; Ahola et al., 1997; Decroly et al., 2011). However, a large tunnel through the central region of nsP1 contains the putative catalytic pocket, and the molecular basis of the capping process by nsP1 remains unknown (Jones et al., 2021; Zhang et al., 2021). The structure and the catalytic mechanism of alphaviral nsP1 GTase are different from the functionally equivalent GTases in humans (Chu et al., 2011), vaccinia virus (Kyrieleis et al., 2014), flavivirus (Issur et al., 2009), and SARS-CoV-2 (Yan et al., 2020).

Here, we report a series of the high-resolution cryo-EM structures of CHIKV nsP1 in complex with m7GTP/SAH, covalently attached m7GMP, and Cap-0 viral RNA as the snapshots of the viral-RNA-capping process. We discover that nsP1 recognizes the viral RNA 5′ end via sequence-specific interactions, which are essential for virus replication.

RESULTS

Structural basis of the N7-MTase activity of CHIKV nsP1

To reveal the molecular mechanism of the methyl-transfer reaction to the N7 position of GTP, we determined the structure of CHIKV nsP1 in complex with the products of the reaction—m7GTP and S-adenosyl homocysteine (SAH) at 2.2 Å by cryo-EM (Figures 1A–1D, S1, and S2; Table S1). The overall structure is similar to that we reported previously (root-mean-square deviation [RMSD] 1.22 Å) (Jones et al., 2021; Zhang et al., 2021). All twelve copies of the N7MTase active sites in the nsP1 ring are occupied by one m7GTP and one SAH. The high-quality experimental cryo-EM density maps allow model building with little ambiguity (Figures S1 and S2). The m7G-binding pocket is located deeply inside the catalytic center (Figure 1C). The N7-methyl group enhances stacking interactions between m7G with Y153 and Y248 from both sides of the base. The sugar moiety of the m7G is supported by V243 from the bottom. The Waston-Crick edge of the guanine base forms the hydrogen bonds with E250 of nsP1 (Leontis and Westhof, 2001). The ribose is further stabilized by the H bond between 2′OH and the side chain of Y285. The triphosphates form extensive direct and water-mediated polar contact with the positively charged arginine residues (R41, R70, and R92) and the catalytic residue H37. The Mg^{2+} ion neutralizes the negative charges by directly binding to the β- and γ-phosphates. SAH, the by-product of the methyl-transfer reaction, sits above the N7-methyl group of m7GTP (Figure 1D). The adenine base of SAH is also sandwiched between P83-R85 (3.5 Å distance) from the top and V156 (3.9 Å distance) from the bottom. The ribose of SAH forms an H bond with D89. The homocysteine group of SAH is further stabilized by the H bonds with G65, R70, and Q151. The adenosyl ribose edge of SAH is solvent accessible, indicating a route for its release from, and SAM's entry into, nsP1. The distance between the sulfur group of SAH and the methyl group of the m7GTP base is about 3.5 Å, compatible with a model of transferring a methyl group from a SAM methyl donor to the N7 position of the GTP with no major conformational change.

Formation of the covalent intermediate nsP1-m7GMP

We captured the intermediate nsP1-m7GMP during the guanylyl-transfer process (Figures 1E and 1F). The covalent bond that formed between H37 and the α-phosphate of the m7GTP is visible in the 2.2 Å resolution density map (Figures 1E, S1, and S2; Table S1). The α-phosphate further forms an H bond with N35. The catalytic H37 forms a water-mediated bridge with D36. The precisely conserved N35-D36-H37 and the bridging water molecule help to maintain the configuration between H37 and the α-phosphate. The m7GMP moiety is located at the same position as m7GTP (Figure 1F), while the side chain of Y248 swings a bit and establishes a slightly stronger stacking interaction with the m7G base. The distance between H37 and the m7GTP α-phosphate is about 5.1 Å. The triphosphate moiety then rotates relative to the sugar base so that the α-phosphate can effectively get close to the H37 and form the covalent intermediate. The Mg^{2+} coordinates the β,γ-diphosphate, helps break the phosphoanhydride bond between α and β phosphates, and leaves with the pyrophosphate by-product. The nsP1-m7GMP intermediate is therefore a relatively stable species once formed. Interestingly, the nsP1-m7GMP structure was obtained when we incubated nsP1 with a 12-mer RNA from the 5′ end of the CHIKV viral genome with a Cap-0 structure at the 5′ end (m7GpppAUGGCUGCGUGA, hereafter named m7GpppAU-10). This indicates that wild-type (WT) nsP1 can cleave off the m7GMP from the capped RNA.

Cap-0-viral RNA recognition by nsP1

To study how a capped viral RNA is produced, we introduced an H37A mutation to nsP1 and determined the structure of nsP1 (H37A) in complex with the RNA m7GpppAU-10 at a resolution of 2.5 Å (Figures 2, S1, and S2; Table S1; Video S1). The ordered part of the RNA substrate consists of m7GpppAU, with some density in the map accounting for the phosphate moiety of G3 of the RNA. The well-resolved Cap-0 RNA tri-nucleotides occupy the single-stranded RNA (ssRNA) binding groove that is formed at the interface between the neighbor nsP1 molecules (Figures 2A–2C; Video S2). Both the m7G cap of the

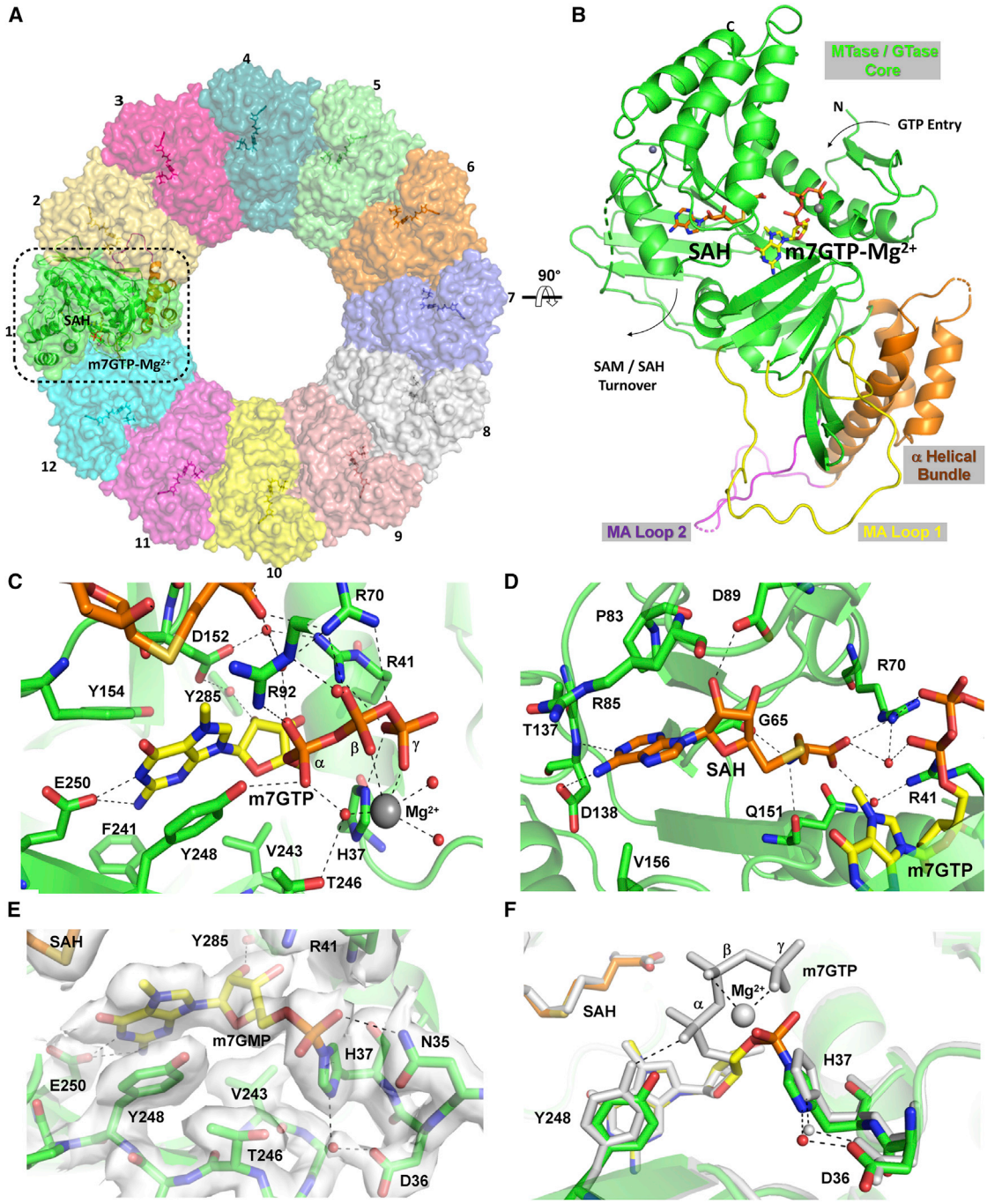


Figure 1. Cryo-EM structure of nsP1-m7GTP-Mg^{2+} and nsP1-m7GMP complexes

(A–F) Cryo-EM structure of the (A–D) nsP1-m7GTP-Mg^{2+} and (E and F) nsP1-m7GMP complexes.

(A) The nsP1 ring is shown as a semi-transparent surface view. M7GTP and SAH are displayed as sticks in each of the catalytic sites.

(B) Cartoon representation of nsP1 monomer colored and labeled according to the functional domains and secondary structures of nsP1 (Zhang et al., 2021).

(C) Close-up view of the m7GTP-binding site. Protein residues binding to SAH are represented as sticks and are labeled.

(D) Close-up view of the SAH-binding site. Protein residues binding to m7GTP-Mg^{2+} are represented as sticks and are labeled. Dashed lines indicate polar interactions.

(E) EM map for m7GMP and the surrounding residues. Dashed lines indicate polar interactions.

(F) Superimposing the complexes of nsP1-m7GTP-Mg^{2+} in gray and nsP1-m7GMP (colored according to the atom types).

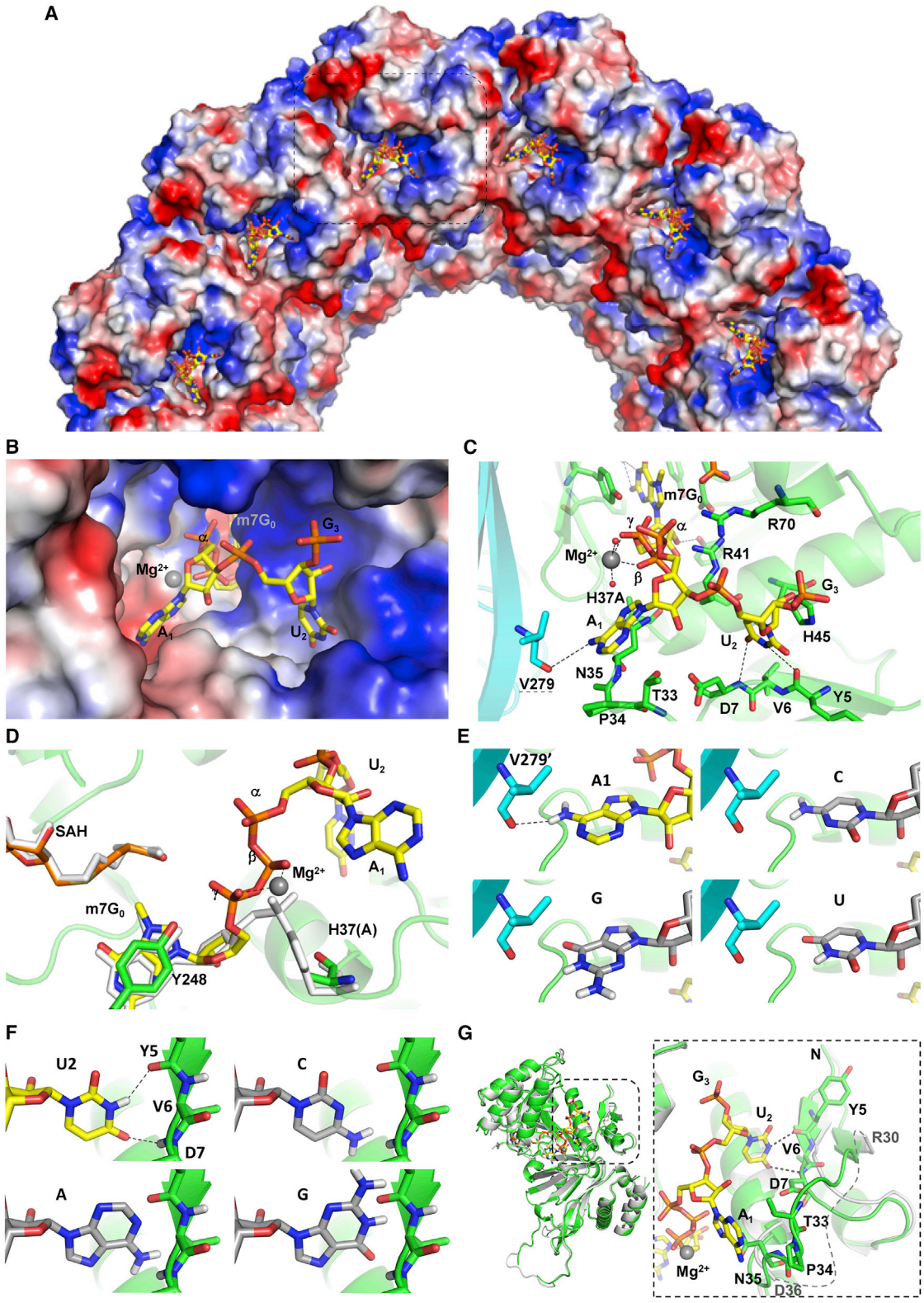


Figure 2. The structure of nsP1 in complex with m7GpppAU-RNA10

(A) Top electrostatic surface view of the nsP1(H37A)with Cap-0-AU10 complex, calculated by PyMol-molecular visualization system. Positive charges are colored in blue, neutral in white, and negative in red. The RNA is shown as sticks. Half of the nsP1 ring is shown for simplicity purposes.

(legend continued on next page)

RNA and SAH are located at the same position as that of the nsP1-m7GMP structure (Figure 2D). The negative charges of the triphosphate group that links m7G and A1 of the RNA are neutralized by the Mg^{2+} ion, which mainly coordinates with the γ- and β-phosphates (names assigned based on the relative position to the RNA 5′ end). When the two structures nsP1-m7GMP and nsP1(H37A)-m7GpppAU-10 are compared and aligned based on the m7G cap, the distance between γ- and β-phosphates of the capped RNA and H37 of the nsP1-m7GMP structure is about 3.9 Å (Figure 2D). Similar rearrangement of the triphosphates of the capped RNA to that of the bound m7GTP (Figure 1C) allows a guanylyl-transfer reaction to occur.

The residues A1 and U2 from the 5′ end of the RNA are not stacked on top of each other. Instead, both are engaged in base-specific interactions: the A1 base is stabilized by the base stacking interacting with a loop formed by T33, P34, and N35; the A1 base also forms an H bond with the backbone carbonyl group of V279 from the neighboring nsP1. The U2 base forms a stacking interaction with H45 that is on the same a helix as the catalytic H37; the Waston-Crick edge of the U2 base forms two specific H bonds with the backbone of the strand Y5-V6-D7 close to the N terminus of nsP1 (Figures 2C and S3). Replacing A1 and U2 with other ribonucleotides causes either loss of the base-specific H bonds, the introduction of potential clashes, or both (Figures 2E and 2F). Overall, the capped RNA induces a closer movement of the N-terminal region of nsP1 and stabilizes it via the stacking interactions and base-specific H bonds (Figure 2G).

nsP1 is a versatile capping enzyme *in vitro*

To uncover the enzymatic capability of nsP1, we sampled different nucleotides and RNA templates as cap donors and recipients, respectively. Cap-0 formation is detected with the commonly used cap donors such as GTP with SAM, m7GTP, m7GTP with SAM, and m7GpppA (Figure 3B). In agreement with the structural observations (Figures 1 and 2), divalent cation Mg^{2+} is essential for the capping reaction, and guanyl transfer is unable to proceed in the absence of Mg^{2+} (Figure 3B, lane 6). When nsP1 is incubated with either a 5′ tri- or di-phosphorylated RNA, there appears to be no significant difference in cap formation between the two cap recipients (Figure 3E). The RNA 5′-end binding pocket can accommodate both 5′ppp or 5′ppRNA (Figure 2C). In the case of 5′pppRNA, the γ-phosphate will be hydrolyzed during the guanyl-transfer reaction.

To investigate whether RNA 5′AU sequence-specific recognition by nsP1 affects the capping activity (Figure 2), the reactions were set up with 1 μM 12-mer recipient RNAs of the sequence from either the CHIKV genome 5′ end (labeled as pppAU-10) or the Dengue virus genome 5′ end (labeled pppAG-10). Capping activity was observed in all three lanes of decreasing concentrations (0.5–0.1 μM) of CHIKV nsP1 in the presence of pppAU-10 (Figure 3C, lanes 4–6) whereas there was little detectable capped product when pppAG-10 was used as the cap recipient (Figure 3C, lanes 7–9). nsP1 may exhibit decapping activity, as it can cleave off the m7GMP from the Cap-0 RNA (Figures 1E and 1F).

Multiple concentrations of substrates were tested to evaluate the capping activity as below, which demonstrates the minimum concentration for each donor varies between the substrates. Both GTP-SAM and m7GpppA can be used as cap donors at concentrations above 100 μM, while m7GTP requires a minimum concentration of 1 mM (Figure 3D). These data suggest that capping is the most preferred enzymatic reaction with GTP/SAM and m7GpppA acting as cap donors successfully at lower concentrations compared with m7GTP. We also show that the enzyme can not only remove the m7GMP cap from other RNAs but also revert from the nsP1-m7GMP in the presence of ATP or ADP (Figure 3F).

Conserved 5′AU in CHIKV nsP1 recognition regulates alphavirus RNA replication and transcription

Many alphaviruses, including CHIKV, start with the AU nucleotides on both genomic and subgenomic RNA (sgRNA) sequences (Figure 4A). This may act as a regulatory function for nsP1 to distinguish its (+) RNAs from the anti-sense intermediate as well as host or other cellular RNAs. CHIKV nsP1 demonstrates preferential binding to ppAU (K_d = 166.5 ± 24.1), pppAU (K_d = 1,920 ± 557.2), and m7GpppAU (K_d = 2,329 ± 550.4). There was minimal to no binding to pppAG and m7GpppAG, which may be a function of both recognition of self RNAs as well as regulating replication and transcription whereby capped RNAs are released to allow continued capping of RNA transcripts (Figure 3G). To determine the functional role of the 5′-end A1 and U2 sequence in CHIKV viral replication and transcription, we used a previously established *trans*-replicase system to perform the mutagenesis on both the genomic 5′AU and the subgenomic 5′AU (Figure 4B) (Bartholomeeusen et al., 2018; Utt et al., 2016). In this assay, the expression levels of firefly luciferase (Fluc) and *Gaussia* luciferase (Gluc) correspond, respectively, to the viral genomic RNA replication and the subgenomic RNA transcription levels.

Substitutions in A1 and U2 in the 5′ UTR have detrimental effects on both genomic and subgenomic expression activities. Fluc activities were severely reduced in all the 5′ UTR mutants (Figure 4C) compared with WT 5′ UTR AU. A similar trend was observed in Gluc activities as well (Figure 4D): the 5′ UTR

(B) Close-up view of one RNA-binding pocket.
(C) Interaction network between the Cap-0 RNA and nsP1. V279 from the neighbor nsP1 (colored in cyan) establishes a base-specific interaction with A1 of the RNA. U2 base establishes polar interactions with the backbone region of Y5-V6-D7 loop region of nsP1.
(D) Superimposing the complexes of nsP1-m7GMP in gray and nsP1(H37A) with Cap-0-AU10-Mg^{2+} (colored according to the atom types).
(E) Close-up view of the RNA A1-specific interaction with the neighbor nsP1 in the top left panel. Modeled C1, G1, and U1 at the same position of the RNA are shown in the top-right, bottom-left, and bottom-right panels.
(F) Close-up view of the RNA U2-specific interaction with the nsP1 Y5-V6-D7 loop region in the top left panel. Modeled C2, A2, and G2 at the same position of the RNA are shown in the top-right, bottom-left, and bottom-right panels.
(G) Superposition RNA-bound nsP1 (in green) with the free enzyme (in gray).

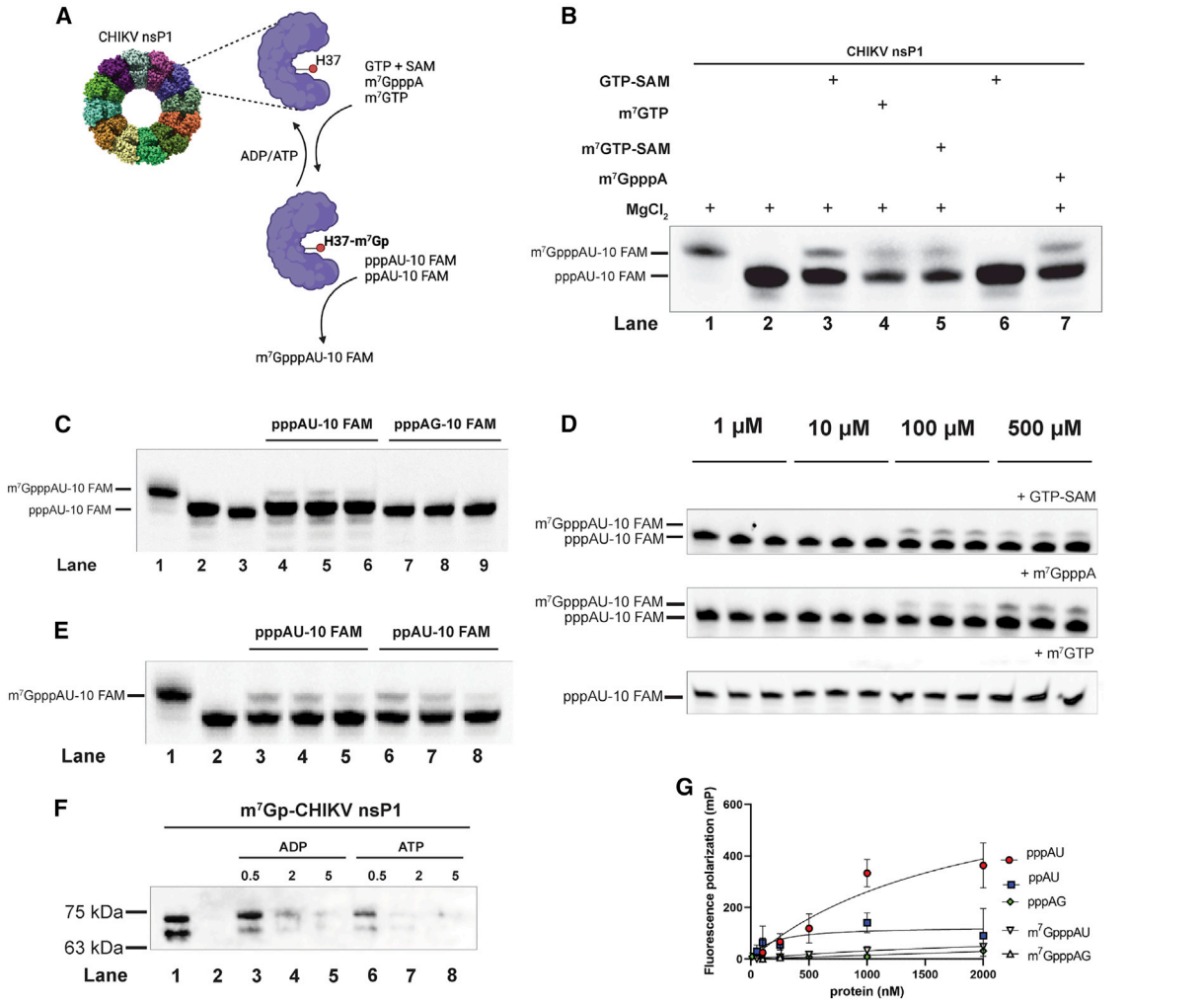


Figure 3. Enzymatic activities of CHIKV nsP1

(A) The flowchart summarizes the RNA-capping experiments.

(B) Capping activity of recombinant nsP1 (concentration 0.75 μM) with pppAU-10 FAM as the RNA-cap recipient using various substrates for cap formation: GTP (1 mM), SAM (0.5 mM), m7GTP (1 mM), m7GpppA (1 mM).

(C) Capping activity of recombinant nsP1 with both pppAU-10 FAM in lanes 4–6 (1 μM) and pppAG-10 FAM in lanes 7–9 (1 μM) as cap-recipient substrate in decreasing concentrations of nsP1 0.5, 0.2, and 0.1 μM. Lanes 2 and 3 are references for pppAU-10 FAM and pppAG-10 FAM without nsP1, respectively.

(D) Evaluation of capping activity in a concentration gradient with GTP-SAM, m7GpppA, and m7GTP.

(E) Above capping activity of recombinant nsP1 with either pppAU-FAM (lane 3–5) ppAU-10 FAM (lane 6–8) as cap-recipient substrate with GTP (1 mM) and SAM (0.5 mM).

(F) Loss of covalent intermediate m7Gp-CHIKV nsP1 when incubated with increasing concentrations of ATP or ADP from 0.5 to 5 mM. Lanes 1 and 2 are references for nsP1 incubated with and without GTP/SAM, respectively.

(G) RNA binding assessed by fluorescence polarization of nsP1 incubated with pppAU-10, ppAU-10, pppAG-10, and capped m7GpppAU-10 and m7GpppAG-10 with N = 3, data are represented as mean ± SEM.

mutants were diminished compared with WT. Collectively, this suggests that both genomic replication and subgenomic transcription activities were impaired significantly. In contrast, mutations in the subgenomic RNA 5′ end had no significant change in Fluc or Gluc activity (Figures 4E and 4F), indicating that the genomic replication is unimpaired (unpaired t test $p > 0.05$). Overall, these results suggest that the conserved AU sequence is crucial in the viral 5′ UTR and that there is some flexibility in the subgenomic promoter.

Structural basis of an antiviral compound targeting the SAM-binding site in CHIKV nsP1

6′-β-Fluoro-homoaristeromycin (FHA) is a potent inhibitor of CHIKV replication with an effective concentration (EC) value of 0.12 ± 0.04 uM and was suggested to target the MTase activity of nsP1 (Kovacikova et al., 2020, 2021). FHA was predicted to bind to two sites in nsP1, but there are no experimental data to support the docking results (Kovacikova et al., 2020, 2021). Here, we determined the cryo-EM structure of nsP1 in complex

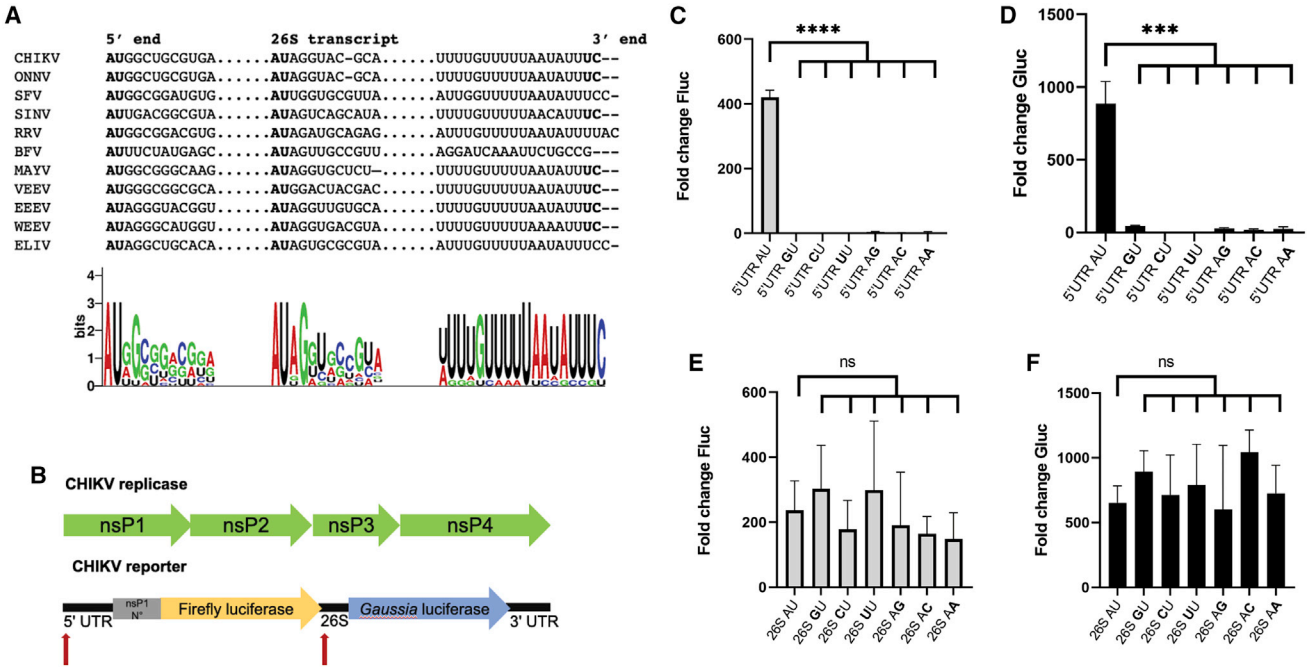


Figure 4. Functional demonstration of the RNA-sequence-specific recognition

(A) The multisequence alignment was generated in CLUSTAL format by MAFFT FTT-NS-i v.7.471 (accessed on June 2, 2021). The web logo was generated by WebLogo (https://weblogo.berkeley.edu/logo.cgi, accessed on June 3, 2021). Accession codes are listed in the supplemental material.

(B) The *trans*-replicase system is comprised of two plasmids encoding the CHIKV replication complex CMV-P1234 and a plasmid HSPoll CHIKV-FG. The reporter plasmid CHIKV-FG to produce replication-competent RNA containing CHIKV 5′ UTR followed by firefly luciferase (Fluc) reporter, the subgenomic promoter followed by *Gaussia* luciferase reporter, and CHIKV 3′ UTR with the sites for substitution denoted by the red arrow. In this assay, the expression levels of Fluc and Gluc correspond, respectively, to the production of full-length genomic and subgenomic RNA. The reporter plasmid or its mutant 5′ UTR or subgenomic promoter sequence plasmids were co-transfected with CMV-P1234. Fluc and Gluc activities were normalized to the negative control CMV-P1234-GAA containing CHIKV nsP4 (RNA-dependent RNA polymerase) that lacks RNA polymerase activity.

(C–F) Effects of substitutions in A1 and U2 in the 5′ UTR (C) Fluc (unpaired t test $p < 0.0001$) and (D) Gluc activity (unpaired t test $p < 0.001$) and 26S promoter on (E) Fluc and (F) Gluc activity. All data are represented as mean ± SEM.

The mean and data points of three independent experiments (N = 3) are shown in (C)–(F).

with FHA (Figure 5). In this structure, FHA occupies all 12 SAM/SAH-binding sites on the nsP1 ring and causes no structural changes to the protein. FHA binds at the SAM-binding site, and the adenosine ribonucleosides overlap with each other very well (Figure 5B), explaining the direct inhibitory activity of FHA to nsP1 MTase activity and virus replication. Our data also show that the two point mutations in nsP1 (K299E and G230R) required to confer resistance to the compound may function by blocking FHA (but not SAM) entry/binding at the SAM-binding site (Figure 5C). These structure data help understand the mechanism of inhibition and guide the further development of the inhibitor.

DISCUSSION

Here, we elucidate the molecular basis of the CHIKV RNA 5′-capping mechanism using combinatory structural and enzymatic approaches (Figures 2 and 3). Essentially, N7-methyltransfer and guanylyl-transfer reactions occur sequentially in the same catalytic tunnel, and no major conformational rearrangement has been observed to complete this multistep enzymatic reaction chain overall. Given the C12 symmetry of the nsP1 ring, all twelve active sites are identical and equally functional. We have also observed that nsP1 recognizes the 5′AU sequence preferentially and demonstrated its functional relevance to capping efficiency and viral RNA replication and transcription in cells as well as in mammalian-expressed recombinant enzyme. This effect is more pronounced when mutations are introduced at the 5′ end compared with the start of sgRNA in the *trans*-replicase assay, reiterating the importance of the AUG initiation codon nearest to the cap structure, which was found to be functional, but not the internal AUG present in sgRNA in Sindbis virus, another alphavirus closely related to CHIKV (Carrasco et al., 2018).

Recently, nsP4 has been implicated in viral RNA template recognition and replication efficiency (Lello et al., 2021). The recently elucidated alphavirus replication complex shows that replication is tightly protected in the spherule, allowing for deliberate coupling of ssRNA capping by nsP1 and nsP2 before exporting it into the cytosol (Tan et al., 2022). The dynamics of alphaviral RNA replication and transcription require the interplay between viral RNA and all four essential members of the alphavirus replication enzymes (Ahola and Merits, 2016; Pietila et al., 2017; Rupp et al., 2015). Similarly, the protein-RNA coevolution

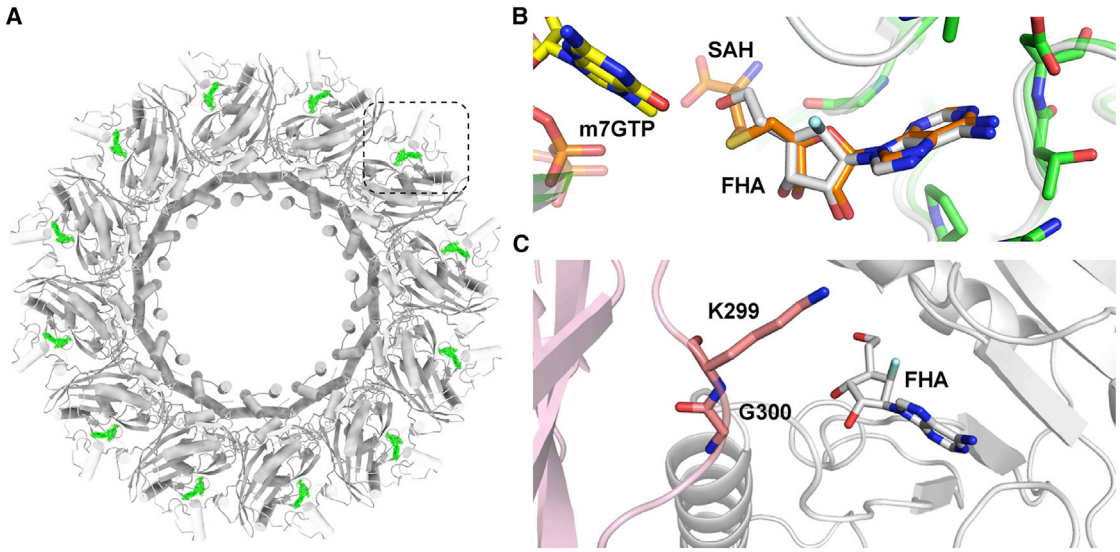


Figure 5. Cryo-EM structure of CHIKV nsP1 and FHA

(A) FHA presence in all twelve catalytic sites of CHIKV nsP1 dodecamer. The difference map between FHA dataset and the free nsP1 dataset is colored in green.

(B) Superposition of FHA and SAH. The two overlap at the same binding site.

(C) Residues K299 and G230 from the neighbor nsP1 that is part of the SAM-binding pocket are highlighted relative to FHA. Point mutations K299E and G230R that give rise to FHA resistance may interfere with FHA by forming a stereo clash with the 6′-β-fluoro moiety.

during viral RNA replication also occurs in both flavivirus and coronavirus (Issur et al., 2009; Swarbrick et al., 2017; Yan et al., 2020; Zhao et al., 2015), suggesting a common molecular principle behind viral RNA biogenesis.

We also demonstrated that nsP1 is a versatile capping enzyme that can utilize various sources for capping including GTP/SAM, m7GTP, and m7GpppA. In the case of m7GpppA, it was shown in VEEV nsP1, another alphavirus, that the excess of diphosphate prevents the regeneration of GTP leading to the methylation of GTP and results in the formation of m7GpppA, allowing us to speculate that nsP1 can switch to alternative cap donors in the absence of GTP for continued formation of the m7GMP-nsP1 complex (Ortega Granda et al., 2021).

Given the strict sequence conservation, all alphaviral nsP1 are likely to share the same structure and function (Figure S3). As many members of alphavirus are known human pathogens causing (re)emerging epidemics across the world (Queyriaux et al., 2008; Schwartz and Albert, 2010; Strauss and Strauss, 1994), our work, together with the recently published alphavirus replication complex (Tan et al., 2022) structure, will guide the development of specific antiviral drugs targeting nsP1, the viral-RNA-capping process, and the alphavirus replication complex (Delang et al., 2016; Feibelman et al., 2018; Ferron et al., 2012).

Limitations of the study

Our data suggest (but do not directly prove) that during viral infection, nsP1 may prefer to recognize its genome or subgenomic sequences bearing the 5′-end AU sequences for capping. The main limitation of this method used to evaluate RNA-binding and -capping activity is that it may not reflect viral infection and replication, as these data are derived from *in vitro* experiments. It disregards the role of the other non-structural proteins and host factors during viral replication after infection. Future work may be to systemically dissect the host immune response against CHIKV infection and map out the contribution from each virus-host interaction pathway in a spatial-temporal manner. For RNA recognization by nsP1, one could expand the RNA repertoire to include RNA length, conserved RNA secondary structures, and different 5′ conserved sequence elements (CSEs) from different alphaviruses.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Material availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Protein expression and purification
  - Cryo-EM grid preparation and microscopy
  - Cryo-EM image processing and model building
  - RNA 5′ end capping assay
  - Immunoblotting of CHIKV nsP1 guanylylation reaction
  - Generation of 5′ UTR and 26S RNA transcript mutant reporter plasmids
  - Analysis of CHIKV replicase activity using a trans-replicase assay
  - Multisequence alignment of alphaviral 5′UTR and 26S RNA transcript

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.111133.

ACKNOWLEDGMENTS

This research is supported by the Singapore National Research Foundation grant NRF2016NRF-CRP001-063 and the Singapore Ministry of Education under its MOE AcRF Tier 2 Award MOE-T2EP30220-0009 and MOE AcRF Tier 1 Award 2021-T1-002-021 to D.L. We thank Dr. Andres Merits for providing the CHIKV cDNA and for insightful discussions. We thank members of the D.L. lab for their support.

AUTHOR CONTRIBUTIONS

K.Z., M.C.Y.L., and D.L. designed the study; K.Z., M.C.Y.L., T.M.N., M.W., Y.-S.L., Y.B.T., L.S.J., and D.L. performed the experiments and analyzed the data; K.Z., M.C.Y.L., and D.L. wrote the manuscript with input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 9, 2021
Revised: June 15, 2022
Accepted: July 1, 2022
Published: July 26, 2022

REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., and Zwart, P.H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221. https://doi.org/10.1107/S0907444909052925.

Ahola, T., and Kääriäinen, L. (1995). Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. Proc. Natl. Acad. Sci. USA *92*, 507–511. https://doi.org/10.1073/pnas.92.2.507.

Ahola, T., and Karlin, D.G. (2015). Sequence analysis reveals a conserved extension in the capping enzyme of the alphavirus supergroup, and a homologous domain in nodaviruses. Biol. Direct *10*, 16. https://doi.org/10.1186/s13062-015-0050-0.

Ahola, T., Laakkonen, P., Vihinen, H., and Kääriäinen, L. (1997). Critical residues of Semliki Forest virus RNA capping enzyme involved in methyltransferase and guanylyltransferase-like activities. J. Virol. *71*, 392–397. https://doi.org/10.1128/JVI.71.1.392-397.1997.

Ahola, T., and Merits, A. (2016). Functions of chikungunya virus nonstructural proteins. In Chikungunya Virus, C.M. Okeoma, ed. (Springer), pp. 75–98.

Banerjee, A.K. (1980). 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. Microbiol. Rev. *44*, 175–205. https://doi.org/10.1128/mr.44.2.175-205.1980.

Bartholomeeusen, K., Utt, A., Coppens, S., Rausalu, K., Vereecken, K., Ariën, K.K., and Merits, A. (2018). A chikungunya virus trans-replicase system reveals the importance of delayed nonstructural polyprotein processing for efficient replication complex formation in mosquito cells. J. Virol. *92*, e00152-18. https://doi.org/10.1128/JVI.00152-18.

Carrasco, L., Sanz, M.A., and González-Almela, E. (2018). The regulation of translation in alphavirus-infected cells. Viruses *10*, E70. https://doi.org/10.3390/v10020070.

Chu, C., Das, K., Tyminski, J.R., Bauman, J.D., Guan, R., Qiu, W., Montelione, G.T., Arnold, E., and Shatkin, A.J. (2011). Structure of the guanylyltransferase domain of human mRNA capping enzyme. Proc. Natl. Acad. Sci. USA *108*, 10104–10108. https://doi.org/10.1073/pnas.1106610108.

Decroly, E., Ferron, F., Lescar, J., and Canard, B. (2011). Conventional and unconventional mechanisms for capping viral mRNA. Nat. Rev. Microbiol. *10*, 51–65. https://doi.org/10.1038/nrmicro2675.

Delang, L., Li, C., Tas, A., Quérat, G., Albulescu, I.C., De Burghgraeve, T., Guerrero, N.A.S., Gigante, A., Piorkowski, G., Decroly, E., and Neyts, J. (2016). The viral capping enzyme nsP1: a novel target for the inhibition of chikungunya virus infection. Sci. Rep. *6*, 31819. https://doi.org/10.1038/srep31819.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132. Retrieved from. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrievedb=PubMeddopt=Citationlist_uids=15572765.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501. https://doi.org/10.1107/S0907444910007493.

Feibelman, K.M., Fuller, B.P., Li, L., LaBarbera, D.V., and Geiss, B.J. (2018). Identification of small molecule inhibitors of the Chikungunya virus nsP1 RNA capping enzyme. Antiviral Res. *154*, 124–131. https://doi.org/10.1016/j.antiviral.2018.03.013.

Ferron, F., Decroly, E., Selisko, B., and Canard, B. (2012). The viral RNA capping machinery as a target for antiviral drugs. Antiviral Res. *96*, 21–31. https://doi.org/10.1016/j.antiviral.2012.07.007.

Ghosh, A., and Lima, C.D. (2010). Enzymology of RNA cap synthesis. Wiley Interdiscip. Rev. RNA *1*, 152–172. https://doi.org/10.1002/wrna.19.

Issur, M., Geiss, B.J., Bougie, I., Picard-Jean, F., Despins, S., Mayette, J., Hobdey, S.E., and Bisaillon, M. (2009). The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure. RNA *15*, 2340–2350. https://doi.org/10.1261/rna.1609709.

Jones, R., Bragagnolo, G., Arranz, R., and Reguera, J. (2021). Capping pores of alphavirus nsP1 gate membranous viral replication factories. Nature *589*, 615–619. https://doi.org/10.1038/s41586-020-3036-8.

Kovacikova, K., Gorostiola González, M., Jones, R., Reguera, J., Gigante, A., Pérez-Pérez, M.J., Pürstinger, G., Moesslacher, J., Langer, T., Jeong, L.S., et al. (2021). Structural insights into the mechanisms of action of functionally distinct classes of chikungunya virus nonstructural protein 1 inhibitors. Antimicrob. Agents Chemother. *65*, e0256620. https://doi.org/10.1128/AAC.02566-20.

Kovacikova, K., Morren, B.M., Tas, A., Albulescu, I.C., van Rijswijk, R., Jarhad, D.B., Shin, Y.S., Jang, M.H., Kim, G., Lee, H.W., et al. (2020). 6'-beta-Fluoro-Homoaristeromycin and 6'-fluoro-homoneplanocin A are potent inhibitors of chikungunya virus replication through their direct effect on viral nonstructural protein 1. Antimicrob. Agents Chemother. *64*, e02532-19. https://doi.org/10.1128/AAC.02532-19.

Kyrieleis, O.J.P., Chang, J., de la Peña, M., Shuman, S., and Cusack, S. (2014). Crystal structure of vaccinia virus mRNA capping enzyme provides insights into the mechanism and evolution of the capping apparatus. Structure *22*, 452–465. https://doi.org/10.1016/j.str.2013.12.014.

Law, Y.S., Utt, A., Tan, Y.B., Zheng, J., Wang, S., Chen, M.W., Griffin, P.R., Merits, A., and Luo, D. (2019). Structural insights into RNA recognition by the Chikungunya virus nsP2 helicase. Proc. Natl. Acad. Sci. USA *116*, 9558–9567. https://doi.org/10.1073/pnas.1900656116.

Lello, L.S., Bartholomeeusen, K., Wang, S., Coppens, S., Fragkoudis, R., Alphey, L., Ariën, K.K., Merits, A., and Utt, A. (2021). nsP4 is a major determinant of alphavirus replicase activity and template selectivity. J. Virol. *95*, JVI0035521. https://doi.org/10.1128/JVI.00355-21.

Leontis, N.B., and Westhof, E. (2001). Geometric nomenclature and classification of RNA base pairs. RNA *7*, 499–512.

Li, C., Guillén, J., Rabah, N., Blanjoie, A., Debart, F., Vasseur, J.J., Canard, B., Decroly, E., and Coutard, B. (2015). mRNA capping by Venezuelan equine

encephalitis virus nsP1: functional characterization and implications for antiviral research. J. Virol. *89*, 8292–8303. https://doi.org/10.1128/JVI.00599-15.

Naydenova, K., Peet, M.J., and Russo, C.J. (2019). Multifunctional graphene supports for electron cryomicroscopy. Proc. Natl. Acad. Sci. USA *116*, 11718–11724. https://doi.org/10.1073/pnas.1904766116.

Ortega Granda, O., Valle, C., Shannon, A., Decroly, E., Canard, B., Coutard, B., and Rabah, N. (2021). Structure and sequence requirements for RNA capping at the Venezuelan equine encephalitis virus RNA 5' end. J. Virol. *95*, e0077721. https://doi.org/10.1128/JVI.00777-21.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612. https://doi.org/10.1002/jcc.20084.

Picard-Jean, F., Tremblay-Ltourneau, M., Serra EDimech, C., Schulz HAnselin, M., and Dutilly, V.B.,M. (2013). RNA 5′-end maturation: a crucial step in the replication of viral genomes. In Current Issues in Molecular Virology - Viral Genetics and Biotechnological Applications, V. Romanowski, ed.

Pietilä, M.K., Hellström, K., and Ahola, T. (2017). Alphavirus polymerase and RNA replication. Virus Res. *234*, 44–57. https://doi.org/10.1016/j.virusres.2017.01.007.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296. https://doi.org/10.1038/nmeth.4169.

Queyriaux, B., Simon, F., Grandadam, M., Michel, R., Tolou, H., and Boutin, J.P. (2008). Clinical burden of chikungunya virus infection. Lancet Infect. Dis. *8*, 2–3. https://doi.org/10.1016/S1473-3099(07)70294-3.

Ramanathan, A., Robb, G.B., and Chan, S.H. (2016). mRNA capping: biological functions and applications. Nucleic Acids Res. *44*, 7511–7526. https://doi.org/10.1093/nar/gkw551.

Rupp, J.C., Sokoloski, K.J., Gebhart, N.N., and Hardy, R.W. (2015). Alphavirus RNA synthesis and non-structural protein functions. J. Gen. Virol. *96*, 2483–2500. https://doi.org/10.1099/jgv.0.000249.

Schwartz, O., and Albert, M.L. (2010). Biology and pathogenesis of chikungunya virus. Nat. Rev. Microbiol. *8*, 491–500. https://doi.org/10.1038/nrmicro2368.

Shatkin, A.J. (1976). Capping of eucaryotic mRNAs. Cell *9*, 645–653. https://doi.org/10.1016/0092-8674(76)90128-8.

Shuman, S. (2001). Structure, mechanism, and evolution of the mRNA capping apparatus. Prog. Nucleic Acid Res. Mol. Biol. *66*, 1–40. https://doi.org/10.1016/s0079-6603(00)66025-7.

Strauss, J.H., and Strauss, E.G. (1994). The alphaviruses: gene expression, replication, and evolution. Microbiol. Rev. *58*, 491–562. Retrieved from. http://www.ncbi.nlm.nih.gov/pubmed/7968923.

Swarbrick, C.M.D., Basavannacharya, C., Chan, K.W.K., Chan, S.A., Singh, D., Wei, N., Phoo, W.W., Luo, D., Lescar, J., and Vasudevan, S.G. (2017). NS3 helicase from dengue virus specifically recognizes viral RNA sequence to ensure optimal replication. Nucleic Acids Res. *45*, 12904–12920. https://doi.org/10.1093/nar/gkx1127.

Tan, Y.B., Chmielewski, D., Law, M.C.Y., Zhang, K., He, Y., Chen, M., and Chiu, W. (2022). Molecular architecture of the chikungunya virus replication complex. Preprint at bioRxiv. https://doi.org/10.1101/2022.04.08.487651.

Utt, A., Quirin, T., Saul, S., Hellström, K., Ahola, T., and Merits, A. (2016). Versatile trans-replication systems for chikungunya virus allow functional analysis and tagging of every replicase protein. PLoS One *11*, e0151616. https://doi.org/10.1371/journal.pone.0151616.

Vasiljeva, L., Merits, A., Auvinen, P., and Kääriäinen, L. (2000). Identification of a novel function of the alphavirus capping apparatus. RNA 5'-triphosphatase activity of Nsp2. J. Biol. Chem. *275*, 17281–17287. https://doi.org/10.1074/jbc.M910340199.

Walker, A.P., Fan, H., Keown, J.R., Grimes, J.M., and Fodor, E. (2021). Identification of guanylyltransferase activity in the SARS-CoV-2 RNA polymerase. Preprint at bioRxiv. https://doi.org/10.1101/2021.03.17.435913.

Yan, L., Ge, J., Zheng, L., Zhang, Y., Gao, Y., Wang, T., Huang, Y., Yang, Y., Gao, S., Li, M., and Lou, Z. (2020). Cryo-EM structure of an extended SARS-CoV-2 replication and transcription complex reveals an intermediate state in cap synthesis. Cell *184*, 184–193.e10. https://doi.org/10.1016/j.cell.2020.11.016.

Yan, L., Ge, J., Zheng, L., Zhang, Y., Gao, Y., Wang, T., Huang, Y., Yang, Y., Gao, S., Li, M., and Lou, Z. (2021). Cryo-EM structure of an extended SARS-CoV-2 replication and transcription complex reveals an intermediate state in cap synthesis. Cell *184*, 184–193.e10. https://doi.org/10.1016/j.cell.2020.11.016.

Zhang, K., Law, Y.S., Law, M.C.Y., Tan, Y.B., Wirawan, M., and Luo, D. (2021). Structural insights into viral RNA capping and plasma membrane targeting by Chikungunya virus nonstructural protein 1. Cell Host Microbe *29*, 757–764.e3. https://doi.org/10.1016/j.chom.2021.02.018.

Zhao, Y., Soh, T.S., Lim, S.P., Chung, K.Y., Swaminathan, K., Vasudevan, S.G., Shi, P.Y., Lescar, J., and Luo, D. (2015). Molecular basis for specific viral RNA recognition and 2'-O-ribose methylation by the dengue virus nonstructural protein 5 (NS5). Proc. Natl. Acad. Sci. USA *112*, 14834–14839. https://doi.org/10.1073/pnas.1514978112.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332. https://doi.org/10.1038/nmeth.4193.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Expi293™ Expression Medium	Thermo Fisher Scientific	Cat# A1435102
Dulbecco's Modified Eagle's Medium (DMEM)	Gibco	Cat# 11965092
Fetal Bovine Serum	Gibco	Cat# 26140079
Phosphate Buffered Saline	Gibco	Cat# 10010002
Opti-MEM™ Reduced Serum Medium	Gibco	Cat# 31985062
Polyethylenimine (PEI 40K)	Polysciences	Cat# 24765
n-Dodecyl-β-D-Maltopyranoside (DDM)	Anatrace	Cat# D310
GDN	Anatrace	Cat# GDN101
Deposited data		
Cryo-EM map of CHIKV nsP1 with m7GTP	This study	EMD-31580
Coordinates of CHIKV nsP1 with m7GTP	This study	7FGG
Cryo-EM map of CHIKV nsP1 with m7GMP	This study	EMD-31581
Coordinates of CHIKV nsP1 with m7GMP	This study	7FGH
Cryo-EM map of CHIKV nsP1 with m7GpppAU-10RNA	This study	EMD-31582
Coordinates of CHIKV nsP1 with m7GpppAU-10RNA	This study	7FGI
Cryo-EM map of CHIKV nsP1 with 6′-β-Fluoro-Homoaristeromycin (FHA)	This study	EMD-32914
Coordinates of CHIKV nsP1 with 6′-β-Fluoro-Homoaristeromycin (FHA)	This study	7X01
Experimental models: Cell lines		
Expi293F™ Cells	Thermo Fisher Scientific	Cat# 100044202
HEK293T	ATCC	Cat# CRL-3216
Recombinant DNA		
pEGFP-C1-CHIKV nsP1 (1-516-SF)	This study	N/A
pEGFP-C1-CHIKV nsP1 (1-516-SF) H37A	This study	N/A
pEGFP-C1-CHIKV nsP1 (1-516-SF) H37E	This study	N/A
pUC57KAN-HPolI-CHIKV-FG 26S GU	This study	N/A
pUC57KAN-HPolI-CHIKV-FG 26S CU	This study	N/A
pUC57KAN-HPolI-CHIKV-FG 26S UU	This study	N/A
pUC57KAN-HPolI-CHIKV-FG 26S AG	This study	N/A
pUC57KAN-HPolI-CHIKV-FG 26S AC	This study	N/A
pUC57KAN-HPolI-CHIKV-FG 26S AA	This study	N/A
pUC57KAN-HPolI-CHIKV-FG 5′ UTR GU	This study	N/A
pUC57KAN-HPolI-CHIKV-FG 5′ UTR CU	This study	N/A
pUC57KAN-HPolI-CHIKV-FG 5′ UTR UU	This study	N/A
pUC57KAN-HPolI-CHIKV-FG 5′ UTR AG	This study	N/A
pUC57KAN-HPolI-CHIKV-FG 5′ UTR AC	This study	N/A
pUC57KAN-HPolI-CHIKV-FG 5′ UTR AA	This study	N/A
Software and algorithms		
PyMOL	Molecular Graphics System	https://pymol.org/2/
PHENIX	Adams et al. (2010)	https://www.phenix-online.org
COOT	Emsley et al. (2010)	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
UCSF Chimera	Pettersen et al. (2004)	http://www.cgl.ucsf.edu/chimera

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
cryoSPARC	Punjani et al. (2017)	https://cryosparc.com
Other		
Quantifoil R1.2/1.3 300 mesh Au holey carbon grids	Quantifoil	Cat #N1-C14nAu30-01
Vitrobot Mark IV	Thermo Fisher Scientific	https://www.thermofisher.com/cn/zh/home.html
300 kV Titan Krios	Thermo Fisher Scientific	https://www.thermofisher.com/us/en/home.html
Strep-Tactin® Sepharose beads	IBA Lifesciences	Cat# 2-1201-010
D-biotin	IBA Lifesciences	Cat# 2-1016-005
BioLock	IBA Lifesciences	Cat# 2-0205-250
Superose® 6 increased 10/300 GL	GE Healthcare	Cat# 17-5172-01
HiLoad®16/600 Superdex® 200	GE Healthcare	Cat# 28-9893-35
HisPur™ Ni-NTA Resin	Thermo Fisher Scientific	Cat# 88223
Anti-7-methylguanoside (m7G)-Cap mAb (monoclonal antibody)	MBL International Corporation	Cat# RN016M; RRID:AB_2921296
Anti-mouse IgG, HRP-linked Antibody	Cell Signalling Technology	Cat# RN016M; RRID:AB_2921296
PVDF Western Blotting Membranes	Sigma Aldrich	Cat# 03010040001
Immobilon Western Chemiluminescent HRP substrate	Millipore	Cat# WBKLS0500

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dahai Luo (luodahai@ntu.edu.sg).

Material availability

Plasmids generated in this study are available from the lead contact with a completed Material Transfer Agreement.

Data and code availability

- Cryo-EM reconstructions and atomic models generated during this study are available at wwPDB and EMBD (https://www.rcsb.org; http://emsearch.rutgers.edu) under the accession codes PDB IDs 7FGG, 7FGI, 7FGH, 7X01, and EMDB IDs 31580, 31581, 31582, and 32914.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Expi293F™ cells (Thermo Fisher Scientific) were incubated a humidified 37°C incubator with 8% CO_2. Cells were incubated in Expi293™ Expression Medium (Thermo Fisher Scientific) with agitation at 120 rpm. Plasmids were transiently transfected into cells using PEI MAX® (Polysciences).

Human embryonic kidney HEK293T cells (ATCC) were maintained in Dubecco's Modified Eagle Medium (Gibco) containing 10% fetal bovine serum (Gibco) at 37°C in a 5% CO_2 incubator. Cells grown on 12-well plates were transfected using PEI MAX® (Polysciences).

METHOD DETAILS

Protein expression and purification

The method for preparing recombinant CHIKV nsP1 was adapted from Zhang et al. (Zhang et al., 2021). The CHIKV nsP1 gene was cloned into the vector pEGFP-C1 (Clonetech) to replace the eGFP gene. Subsequently, a tag composed of strep and 3XFLAG tag was inserted after A516 residue. This plasmid was transfected into Expi293F cells by PEI MAX® (Polysciences) for transient expression. Cell pellets were harvested after 96 h and lysed in lysis buffer 50 mM Tris-HCl pH 8.0 150 mM NaCl 0.5 mM TCEP with protease inhibitor, 1% n-dodecyl-β-D-maltoside (DDM), and 70 mU/mL of BioLock solution (IBA Lifesciences). The pellet was incubated at 4°C for 2 h on a rotator with occasional vortex to ensure complete suspension of the pellet in the lysis buffer. It was subsequently

sonicated for 10 min, and centrifuged for 1 h at 18,000 g at 4°C. The supernatant was applied to an Econo-Pac® chromatography column (Biorad) containing Strep-Tactin® Sepharose beads (IBA Lifesciences). After serial washing with 5, 25, and 50 column volumes with the wash buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 0.5 mM TCEP, 0.01% GDN, 10% glycerol), the recombinant proteins were eluted using wash buffer with added 10 mM D-biotin (IBA Lifesciences). The elutants were further purified through Superose® 6 increased 10/300 GL (GE Healthcare Life Sciences) or concentrated to the desired concentration for RNA 5' end capping assay.

Cryo-EM grid preparation and microscopy

To prepare samples of nsP1 with m7GTP, m7GMP, and FHA, purified wild type (WT) nsP1 protein were incubated with SAH-M7GTP, SAH-m7GpppAU-10RNA, and FHA for 12hrs respectively. For the sample of nsP1 with m7GpppAU-10RNA, nsP1 with H37A mutation was used to incubate with SAH-m7GpppAU-10RNA. Before preparing Cryo-EM grids, the carbon side of Quantifoil R1.2/1.3 gold 300 mesh grid was covered with one layer of graphene following a published protocol (Naydenova et al., 2019). After glow-discharging for 10 s at low energy with Plasma cleaner, 3 μL of CHIKV nsP1 (concentration about 0.2 mg/mL) with different ligands or RNA was applied to the grids, blotted for 2.5 s with blot force −2, and plunge-frozen in liquid ethane using Vitrobot (Thermo Fisher Scientific). Cryo-EM grids were imaged on a Thermo Fisher Scientific 300 kV TEM Titan Krios using Gatan K2 direct electron detector or Falcon III detector. The detailed information of data collection was summarized in Table S1.

Cryo-EM image processing and model building

The collected movies of different samples were motion-corrected and dose weighted using MotionCor2 (Zheng et al., 2017). The output micrographs were imported to Cryosparc v3.2.0 (Punjani et al., 2017). After CTF Estimation, micrographs with CTF-estimated maximum resolution better than 4 Å were selected to pick particles using Template Picker. After one round of 2D classification, the selected good particles were subjected to Heterogeneous Refinement using the previously published map of CHIKV nsP1 as a template. After several rounds of Heterogeneous Refinement, the best classes of each sample were performed one round of Homogeneous Refinement with no symmetry (C1) and resulted in 2.9–3.6 Å maps. After CTF refinement and one round of Non-uniform Refinement with C12 symmetry, the maps were improved to 2.2–2.6 Å. The maps were further sharped with different B-factors as shown in Table S1. Figures representing the map features were prepared with UCSF Chimera (Pettersen et al., 2004). The details of the data collection and processing refer to the extended data Figures S1, S2 and Table S1.

The final map allows *de novo* model building for 80% of the structure using Phenix.map-to-Model (Adams et al., 2010). Our recently deposited nsP1 structure (PDB: 7DOP) was used as a reference for further model building and refinement. The model was further improved through cycles of real-space refinement (with Ramachandran restraints and secondary structure restraints) in Phenix and following manual corrections by Coot (Emsley and Cowtan, 2004; Emsley et al., 2010). The refinement statistics of the model are summarized in Table S1. Figures representing the structural features were prepared with UCSF Chimera and PyMOL (http://pymol.org).

RNA 5′ end capping assay

The method previously described by Li et al. was used with modifications. CHIKV 5′ UTR sequence is used as the template to synthesize the substrate RNA (Li et al., 2015). The 5′ triphosphorylated (ppp) single-stranded 12-mer RNA AUGGCUGCGUGA labeled with a FAM dye was named "pppAU-10 FAM". A 5′ diphosphorylated (pp) single-stranded 12-mer RNA AUGGCUGCGUGA was labeled with a FAM dye at the 3′ end and named as "ppAU-10 FAM". A Cap-0 RNA was synthesized based on the above sequence, named as "m^7GpppAU-10" or labeled with FAM dye was named "m7GpppAU-10 FAM" respectively. In addition, 5′ triphosphorylated (ppp) single-stranded 12-mer RNA AGUUGUUAGUCU was based on the dengue virus 5′ UTR labeled with FAM dye was named "pppAG-10 FAM" and a Cap-0 RNA "m^7GpppAG-10 FAM". The RNAs were synthesised by Trilink Biotechnologies and Bio-Synthesis Inc. The capping reaction is in two steps, first, a covalent nsP1-m7GMP intermediate reaction was prepared in a 20 μL containing 20 μM nsP1, 50 mM Tris-HCl pH 7.5, 2 mM DTT, 10 mM KCl, and 2 mM $MgCl_2$ and 0.5 mM S-adenosylmethionine (SAM) and 1 mM GTP or an alternative substrate as cap donor was incubated at 30°C for 2 h. Second, after the preparation of the covalent m7GMP-nsP1 is used for the transfer to RNA recipient in a 20 μL mixture containing 5 μL covalent intermediate, 50 mM Tris-HCl pH 7.5, 2 mM DTT, 2 mM $MgCl_2$, and 1 μM synthetic RNA, and 20 U Murine RNase Inhibitor (New England Biolabs). The reactions were incubated at 30°C for 12 h and terminated by adding the stop solution 2X RNA loading dye (95% formamide 0.02% bromophenol blue 0.01% xylene cyanol 0.02% SDS 1 mM EDTA). The capped RNA products were separated by 20% denaturing RNA 8M urea PAGE gel and visualized using ChemiDoc™ MP imaging system (Biorad).

Immunoblotting of CHIKV nsP1 guanylylation reaction

Reactions were prepared in 20 μL containing 1.5 μM nsP1, 50 mM Tris-HCl pH 7.5, 2 mM DTT, 10 mM KCl, and 2 mM $MgCl_2$ and 0.5 mM S-adenosylmethionine (SAM) and 1 mM GTP or alternative cap donors such as mRNA or total RNA extracts and incubated for 2 h at 30°C. It was stopped with the addition of Laemmli sample buffer, and products were loaded onto a 12% SDS-PAGE gel. The 12% SDS-PAGE gel was stained with Coomassie blue or prepared for immunoblotting and transferred on to a polyvinylidene difluoride (PVDF) membrane (Sigma Aldrich) with 1X Transfer buffer (48 mM Tris 39 mM glycine and 20% methanol) by Trans-Blot Turbo™ (Biorad) for 20 min at max 2.5 A and 25 V. The membrane was blocked with 5% BSA 1X PBS for 30 min at room temperature and

subsequently incubated with the primary antibody anti-7-methylguanosine (m7G) cap (RN016M, MBL International Corporation) 1 μg/mL overnight at 4°C and washed thrice with washing buffer 1% BSA 1X PBS 0.1% Tween 20 for 5 min with agitation. The secondary antibody (anti-mouse IgG HRP-linked #7076, Cell Signaling Technology) is added to the membrane for 1 h at room temperature and washed thrice again with washing buffer before the addition of the chemiluminescence reagent (Immobilon Western Chemiluminescent HRP Substrate, Millipore). The extra reagent is removed from the membrane and visualised on the ChemiDoc™ MP imaging system (Biorad).

Generation of 5′ UTR and 26S RNA transcript mutant reporter plasmids

The reporter plasmid HPol-FG-CHIKV was constructed containing the full length 5′ UTR of CHIKV before the firefly luciferase (Fluc) as well as the subgenomic promoter preceding the *Gaussia* luciferase (Gluc) under the CMV promoter as described in Utt et al. (Utt et al., 2016). To generate mutant reporter plasmids to substitute the first nucleotide (1) A to either G, C, or U/T and the second nucleotide (2) from U/T to G, C or A primers were designed to target nucleotides 613 and 614 in the region corresponding to the first two nucleotides in the 5′ UTR with the sequence CGCCGGGTT12TATGGCTGCGT. Similarly, for mutations in the subgenomic promoter region, primers targeting nucleotides 2594 (1) and 2595 (2) where the first or second nucleotide of the subgenomic promoter is substituted similarly in the resulting 26S RNA transcript with the sequence GCCGTGTAATCCGGT12AATTAATTAACATGGCC. Successful mutants were identified via Sanger sequencing. The mutant plasmids in the 5′ UTR with the first nucleotide substitution correspondingly referred to as 5′UTR G, 5′UTR C, and 5′UTR U, with the second nucleotide substitution as 5′UTR AG, 5′UTR AC, and 5′UTR AA. Likewise, mutant plasmids where mutations in the subgenomic promoter with the first nucleotide substituted resulting in substitutions in the 26S RNA transcripts are referred to as 26S GU, 26S CU, and 26S UU, and with the second nucleotide substituted as 26SAG, 26S AC and 26S AA. The replicase activity of these plasmids was evaluated by trans-replicase assay.

Analysis of CHIKV replicase activity using a trans-replicase assay

HEK293T cells were maintained in Dubecco's Modified Eagle Medium (Gibco) containing 10% FBS at 37°C in a 5% CO_2 incubator. Cells grown on 12-well plates were cotransfected with 1 μg of each plasmid encoding CHIKV replicase (CMV P1234 Utt et al., 2016) with the reporter HPol-CHIKV-FG or one of the following reporter mutants; 5′UTR G, 5′UTR C, 5′UTR U, 5′UTR AC, 5′UTR AG, 5′ UTR AA, 26S GU, 26S CU, 26S UU, 26S AG, 26S AC or 26S AA by PEI MAX® (Polysciences). Transfected cells were incubated for 24 h, transfections were performed in triplicate, and experiments were replicated. After incubation, media was aspirated, cells were washed once in Phosphate Buffered Saline (Gibco), lysed and activities of Fluc and Gluc were measured using Pierce™ Gaussia-Firefly Luciferase Dual assay kit (Thermo Fisher) onto a 96 well plate (Corning) with a microplate reader (Synergy™ H1 BioTek). The activity of the replicase with the reporter mutants was normalised to control cells transfected with HPol CHIKV FG wt reporter and CMV P1234-GAA the CHIKV replicon with non-functional RNA dependent RNA polymerase. Each experiment was performed in triplicate and replicated. The results were analyzed with Unpaired t test on Prism 9.3.1 (GraphPad).

Multisequence alignment of alphaviral 5′UTR and 26S RNA transcript

The multisequence alignment was generated in CLUSTAL format by MAFFT FTT-NS-i v7.471 (accessed on June 2, 2021). The web logo was generated by WebLogo (https://weblogo.berkeley.edu/logo.cgi accessed on June 3, 2021). Accession codes used are as follows: Chikungunya virus (CHIKV) NC_004162, O'nyong nyong virus (ONNV) AF079456.1, Semliki Forest virus (SFV) NC_003215.1, Sindbis virus (SINV) NC_001547.1, Ross River virus (RRV) GQ433354.1, Barmah Forest virus (BFV) MN689034.1, Mayaro virus (MYV) NC_003417.1, Venezuelan equine encephalitis virus (VEEV) L01442.2, Eastern equine encephalitis virus (EEEV) EF151502.1, Western equine encephalitis virus (WEEV) MN477208.1, Eliat virus (ELIV) NC_018615.1.

Recombinant CHIKV nsP1 RNA-binding assay

100 mM pppAU-10 FAM synthetic RNA was diluted in TE buffer 10 mM Tris-HCl pH 7.5 1 mM EDTA to 100 nM. Reactions were prepared in 30 μL containing increasing concentration of recombinant CHIKV nsP1 or its mutant was incubated in 50 mM Tris-HCl pH 7.5, 2 mM DTT, 2 mM $MgCl_2$ with 50 nM pppAU-10 FAM synthetic RNA and 20 U Murine RNase Inhibitor (New England Biolabs) in 384-well half-area black flat-bottom microplate plate (Corning) and equilibrated at room temperature for 20 min. The measurement was performed by a Synergy H1 plate reader (BioTek) at an excitation wavelength of 485 nM and an emission wavelength of 528 nm. The data were analyzed using nonlinear regression with Prism 9.3.1 (GraphPad).

QUANTIFICATION AND STATISTICAL ANALYSIS

Cryo-EM data were processed and analyzed using cryoSPARC. Cryo-EM structural statistics were analyzed with Phenix. Statistical details of experiments are described in method details or Figure Legends.

All data shown in Figures 3 and 4 represent the average ± standard deviation (SD) from at least three technical replicates. Statistical analyses were performed on Prism 9.3.1 (GraphPad).