

Review

## A Structural Overview of RNA-Dependent RNA Polymerases from the *Flaviviridae* Family

Jiqin Wu <sup>1,2,3</sup>, Weichi Liu <sup>1,2,3</sup> and Peng Gong <sup>1,2,\*</sup>

<sup>1</sup> Key Laboratory of Special Pathogens and Biosafety, Wuhan Institute of Virology, Chinese Academy of Sciences, No. 44 Xiao Hong Shan, Wuchang District, Wuhan 430071, China; E-Mails: wujiqin2010@163.com (J.W.); liuweichi1990@163.com (W.L.)

<sup>2</sup> State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, China

<sup>3</sup> University of Chinese Academy of Sciences, Beijing 100049, China

\* Author to whom correspondence should be addressed; E-Mail: gongpeng2001@yahoo.com or gongpeng@wh.iov.cn; Tel./Fax: +86-27-8719-7578.

Academic Editor: Charles A. Collyer

Received: 1 April 2015 / Accepted: 28 May 2015 / Published: 8 June 2015

---

**Abstract:** RNA-dependent RNA polymerases (RdRPs) from the *Flaviviridae* family are representatives of viral polymerases that carry out RNA synthesis through a *de novo* initiation mechanism. They share a  $\approx$  600-residue polymerase core that displays a canonical viral RdRP architecture resembling an encircled right hand with palm, fingers, and thumb domains surrounding the active site. Polymerase catalytic motifs A–E in the palm and motifs F/G in the fingers are shared by all viral RdRPs with sequence and/or structural conservations regardless of the mechanism of initiation. Different from RdRPs carrying out primer-dependent initiation, *Flaviviridae* and other *de novo* RdRPs utilize a priming element often integrated in the thumb domain to facilitate primer-independent initiation. Upon the transition to the elongation phase, this priming element needs to undergo currently unresolved conformational rearrangements to accommodate the growth of the template-product RNA duplex. In the genera of *Flavivirus* and *Pestivirus*, the polymerase module in the C-terminal part of the RdRP protein may be regulated *in cis* by the N-terminal region of the same polypeptide. Either being a methyltransferase in *Flavivirus* or a functionally unclarified module in *Pestivirus*, this region could play auxiliary roles for the canonical folding and/or the catalysis of the polymerase, through defined intra-molecular interactions.

**Keywords:** *Flaviviridae*; RNA-dependent RNA polymerase; catalytic motif; *de novo* initiation; elongation; *in cis* regulation

---

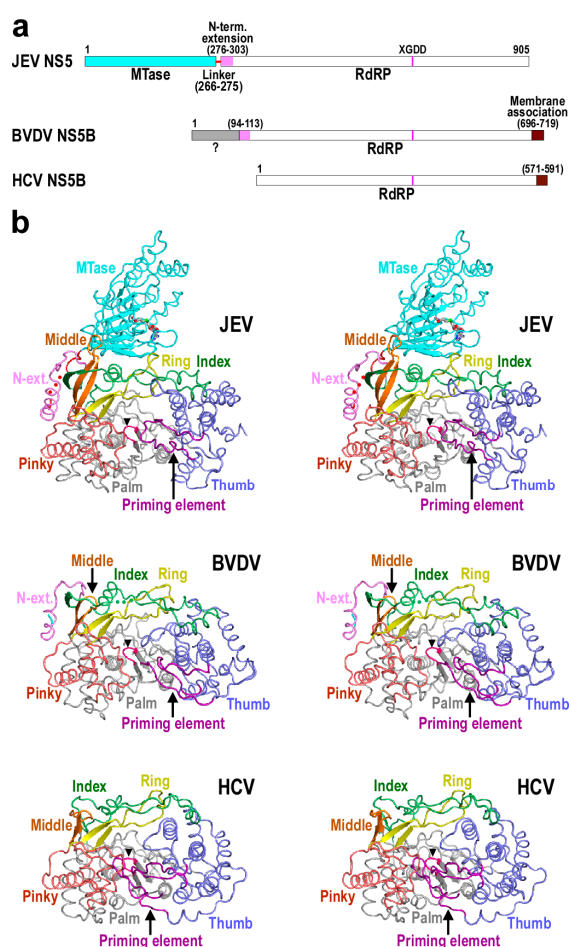
## 1. Introduction

The viruses of the *Flaviviridae* family include a large number of important human and animal pathogens with notable members including hepatitis C virus (HCV) from the genus *Hepacivirus*, dengue virus (DENV), Japanese encephalitis virus (JEV), West Nile virus (WNV), and tick-borne encephalitis virus (TBEV) from the genus *Flavivirus*, and bovine viral diarrhea virus (BVDV) and classical swine fever virus (CSFV) from the genus *Pestivirus*. The non-segmented single-stranded RNA genome of these viruses is positive sense and has a typical length of 9.5–12.3 kilo-bases [1–3]. The genome contains a large open reading frame (ORF) flanked by 5′ and 3′ nontranslated regions (NTRs) that usually contain structured element regulating viral genome replication and viral protein translation [4–7]. The 5′ end of the *Flavivirus* genome bears a type 1 cap structure (cap 1) [8], while the genome of other members of *Flaviviridae* is not 5′ capped and rather contains an internal ribosome entry site (IRES) in the 5′ NTR for cap-independent translation [9,10]. The 3′ end of the genome for all *Flaviviridae* members is unexceptionally not poly-adenylated. The single ORF encodes a ≈3000–3900-residue polyprotein that is processed into ≈10–12 structural and non-structural proteins by viral and host proteases [11]. While the structural proteins are key components of viral capsid and envelop, the non-structural proteins all participate in genome replication that occurs in membrane-associated sub-structures derived from and connected to the endoplasmic reticulum [12]. Lying in the heart of the genome replication machinery (also termed the replication complex) is the virally encoded RNA-dependent RNA polymerase (RdRP) that governs the catalysis in synthesizing genome-length RNA. In this review, we present our current understanding of *Flaviviridae* RdRPs primarily from structural perspective with focuses on polymerase catalysis and regulation.

## 2. The Architecture of *Flaviviridae* RdRP Protein and Important Components for Polymerase Catalysis

The viral proteins that carry out RdRP function in family *Flaviviridae* vary in size, with about 600–900 residues encoded. Among these, the nonstructural protein 5 (NS5) of *Flavivirus* is the largest, having a ≈260-residue *S*-adenosyl-L-methionine (SAM)-dependent methyltransferase (MTase) fused to the N-terminus of the polymerase module through a short flexible linker (Figure 1a). Based on consensus folding of the MTase and RdRP in numerous crystal structures [13–16] and the first full-length NS5 crystal structure in JEV [17], the linker spans about ten residues following a highly consensus GTR sequence in the C-terminus of the MTase. It exhibits tendency to be disordered or to adopt an extended conformation in the majority of structures. As an exception, a helical fold was observed in this region in a recent report of the full-length DENV NS5 structure [18], suggesting its potentials to coordinate domain motions of the regions it is tethering. The interactions and regulations between the MTase and RdRP had remained enigmatic until the first full-length NS5 crystal structure was reported in 2013, starting to unravel the rather complicated nature of the MTase-RdRP interplay [17].

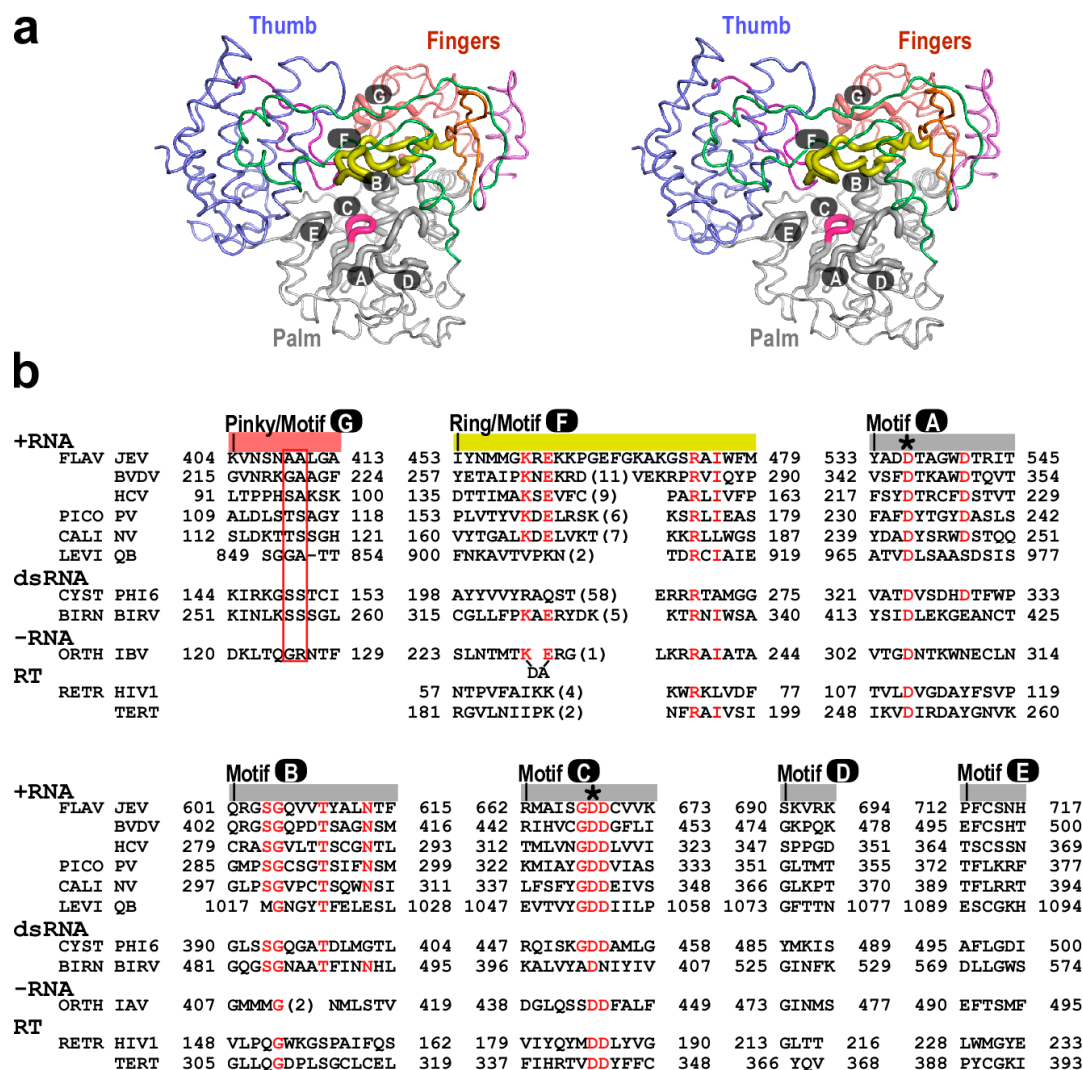
This JEV NS5 structure defines a medium size MTase-RdRP interface with key residues highly conserved in *Flavivirus* and recently proven to be functionally important [19,20]. In contrast to *Flavivirus* NS5, the nonstructural protein 5B (NS5B) of HCV is essentially the polymerase module with a C-terminal 21-residue “membrane anchor” (Figure 1a). The *Pestivirus* NS5B appears to be a “hybrid” with its C-terminal hydrophobic region resembling that of HCV NS5B and its N-terminal 90 residues reminiscent of the *Flavivirus* NS5 MTase (Figure 1a). The *Flavivirus* NS5 and *Pestivirus* NS5B also share a 20–28-residue region first defined in the full-length JEV NS5 structure as “the N-terminal extension” to the core polymerase and recently shown to be non-dispensable to regular polymerase catalysis in JEV NS5 (Figure 1a) [19]. Although sequence similarity of this N-terminal extension is quite low between the two genera, they adopt a similar fold, and if properly presented, were structurally integrated into the core polymerase to form one entirety (Figure 1b) [14,15,17,21].



**Figure 1.** Structural comparison of representative *Flaviviridae* RdRPs. (a) A schematic of *Flaviviridae* RdRPs defining functional regions; (b) Stereo-pair images of *Flaviviridae* RdRP structures (pdb entries: 4K6M, 1S4F and 1NB4) viewing down into the polymerase active site. The RdRP signature sequence XGDD is shown in magenta (also indicated by a black triangle). Red dots in the JEV structure indicates disordered residues 271–273 in the MTase-RdRP linker and the *S*-adenosyl-L-homocysteine (SAH) bound in the MTase domain is shown as sticks. In the BVDV structure, residues 92–127 from a second molecule are connected to the C-terminal part of the protein by the green dots to present the likely canonical fold. Structures were superimposed using program THESEUS [22].

Viral RdRPs adopt a unique encircled right hand architecture with palm, fingers, and thumb domains surrounding the active site. The fingers domain has been further divided into index, middle, ring, and pinky finger subdomains to better elucidate the RdRP function [17,23] and the encirclement of the active site is achieved through interactions between index finger and thumb (Figure 1b). As a consequence, large-scale rotational conformational changes of the fingers domain typically observed in the nucleotide addition cycle of Pol I family polymerases are not feasible for viral RdRPs [24,25], and they instead utilize small-scale rearrangements in the palm domain to achieve active site closure necessary for the phosphoryl transfer reaction [26], thus providing a structural basis for rational design of active site inhibitors specifically acting on viral RdRPs. With respect to the initiation mechanism of RNA synthesis, viral RdRPs can be classified into two major classes. The *Flaviviridae* core polymerase represents *de novo* RdRPs that utilize two initiating NTPs to form the first phosphodiester bond of the product strand [27,28]. The other class is represented by *Picornaviridae* (e.g., poliovirus, or PV) RdRPs that take advantage of a primer (a virally encoded peptide in PV) in the early stages of RNA synthesis [29]. One primary structural difference between *de novo* and primer-dependent RdRPs is that the former often have a priming element integrated into their thumb domain. This priming element penetrates into the active site from the upstream direction (Figure 1b), allows only the 3' portion of the template strand to approach the active site from the downstream for initiating NTP binding, and facilitates the first few steps of RNA synthesis [28].

For all viral RdRPs, the palm domain harbors polymerase motifs A–E with the most conserved feature embedded in motifs A–C [30,31]. Motifs A and C each contains an aspartic acid residue that is universally conserved for all single-subunit processive nucleic acid polymerases (Figure 2b), playing central roles in the two-metal ion catalytic mechanism [17,32]. Motif B includes a highly conserved serine (JEV NS5 residue 604 or equivalent) known to play key roles in recognizing the 2'-hydroxyl group of the NTP ribose [26,33]. It also contains a glycine residue (JEV NS5 residue 605 or equivalent) that is conserved in all RNA-dependent polymerases. This glycine not only provides flexibility for the Ser-Gly peptide bond flip to coordinate the side-chain rotamer change of the adjacent serine in active site closure [26], but also may play a key role in polymerase translocation. In a recent study, a loop region (JEV NS5 residues 603–609 or equivalent) of PV RdRP centering around this glycine was shown to adopt two distinct backbone conformations and was proposed to mediate the movement of the template-product duplex toward the upstream in the post-catalysis translocation event [34]. Motifs D and E are less conserved in sequence. Structurally, motif D is associated with motif A and they undergo coordinated conformational changes during the closure and reopening of the active site [26]. A motif D lysine residue (K359 in PV RdRP) that is conserved in primer-dependent RdRPs has been proposed to play critical roles in catalysis [35]. However, the region spanning this residue does not have equivalents in *de novo* RdRPs, either by sequence or by structural homology. Motif E folds with motif C, and primarily interacts with the backbone of the –2 and –3 positions (*i.e.*, the 2nd and the 3rd nucleotide upstream of the active site) of the product RNA.



**Figure 2.** Catalytic motifs of *Flaviviridae* RdRPs. (a) Stereo-pair images of spatial organization of the JEV RdRP catalytic motifs A–G (pdb entry: 4K6M). Seven RdRP motifs are shown as thick noodles. The color-coding is as in Figure 1b; (b) A structure-based sequence alignment depicting the conservation of RdRP motifs (pdb entries: 4K6M, 1S4F, 1NB4, 3OL6, 3BSO, 3AVT [36], 1HI0, 2PGG [37], 4WRT [38], 1RTD [39], and 3DU6 [40]). Three RdRPs from other positive-strand RNA viruses (PICO/PV: *Picornaviridae*/poliovirus; CALI/NV: *Caliciviridae*/norovirus; LEVI/QB: *Leviviridae*/bacteriophage Q $\beta$ ), two from double-stranded RNA viruses (CYST/PHI6: *Cystoviridae*/bacteriophage  $\phi$ 6; BIRN/BIRV: *Birnaviridae*/birnavirus), one from negative-stranded RNA viruses (ORTH/IBV: *Orthomyxoviridae*/influenza virus B), and two reverse transcriptases (RTs) (RETR/HIV1: *Retroviridae*/human immunodeficiency virus 1; TERT: Telomerase RT) were chosen as representatives for alignment and/or comparison. Only the structurally conserved segment of motif D is included in this alignment. Some important RdRP consensus residues are highlighted in red. Two motif G residues interacting with the +1/+2 junction of the RNA template are indicated by a red box. The two universal aspartic acid residues are indicated by asterisks. Numbers in parenthesis indicate the number of residues not shown. The alignment of IBV polymerase motif G is of lower confidence due to a lower level of structural similarity to motif G in other structures.

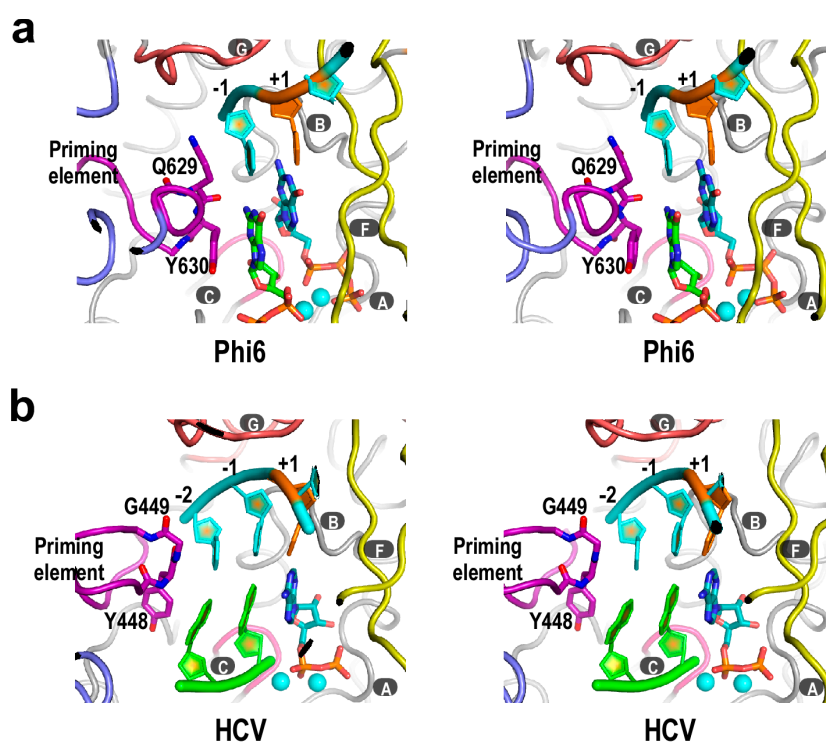
Motifs F and G reside in the fingers domain, and both play critical roles for polymerase function. Motif F is common for all RNA-dependent polymerases and its functions are yet to be clarified [41]. When properly folded, it forms the roof of the NTP entry channel and adopts an antiparallel  $\beta$ -type structure with key residues (JEV NS5 K459 and E461 or equivalent) in the N-terminal half approaching the nascent base pair through the hoogsteen face and those (JEV NS5 R474 and I476 or equivalent) in the C-terminal half interacting with the base pair plane and the NTP triphosphate moiety [26,28,42]. The N-terminal half of motif F was also proposed to mediate stem-loop A (SLA)-promoted RNA synthesis in DENV, suggesting possible different roles of motif F in different stages of RNA synthesis [43]. Note that not all RdRPs motifs exhibit high level of sequence conservation, even if they were originally identified based on sequence homology among a collection of viral species [30]. Based on known RdRP-RNA structures [26,44–46], Motif G runs approximately parallel to the template strand with two residues (JEV NS5 A409-A410 or equivalent) running vertically to the +1/+2 backbone kink of the RNA template. Although sequence conservation is quite low across the representative viral species for motif G, in particular, between primer-dependent and *de novo* RdRPs, its critical location implies potential key roles in binding and/or translocating the RNA template strand.

### 3. Structural Basis for Polymerase Catalysis

Analogous to DNA-dependent RNA polymerases [47–49], an initiation complex (IC) of *Flaviviridae* RdRP is unstable, and tends to release short RNA products and re-initiate [27]. However, precise and relatively efficient initiation can be achieved with the help of the priming element. Note that the priming element is structurally quite diverse across different viruses (Figure 1b). It could comprise solely an extended loop structure, as in *Flavivirus* NS5 [14,15,17], or two peptide segments that contain a  $\beta$ -hairpin structure, as in HCV NS5B [50], or even a helical module, as in bacteriophage  $\phi$ 6 polymerase [28]. Crystal structures of the *de novo* RdRP IC from  $\phi$ 6 and HCV have been reported, providing the structural basis for how the early stages of RNA synthesis is generally achieved with the assistance of the priming element (Figure 3). In the  $\phi$ 6 polymerase IC structure, residues Q629 and Y630 of the priming element were placed at the immediate upstream position of the initiation site, stacking onto the  $-1$  templating nucleotide and the priming NTP, respectively (Figure 3a) [28]. Such an arrangement likely ensures terminal initiation for faithful replication, and may also stabilize the priming NTP for catalysis. In the recently reported set of HCV NS5B IC structures, different dinucleotides were used as short primers to mimic the situation in *de novo* initiation [42]. Similar to the  $\phi$ 6 polymerase IC, priming element residues Y448 and G449 interact with the upstream end (now position  $-2$  vs.  $-1$  in  $\phi$ 6 structure) of the template and product strands, respectively (Figure 3b). Note that the priming element could contribute to efficient initiation, or precise start site selection, or both. In an earlier report, when the  $\beta$ -hairpin of the HCV NS5B priming element was truncated for eight residues (residues 444–447 and 450–453 were removed, and residues 448–449 were replaced by two glycines to accommodate the truncation), the *de novo* synthesis activity was not much affected, but this mutant NS5B lost the specificity of terminal initiation and could initiate from internal sites of a template [51]. In a very recent report, mutation of hydrophobic and charged residues within the priming element of HCV NS5B resulted in a substantial decrease of *de novo* initiation activities [52]. These data indicate that the priming element does play essential roles in *de novo* initiation, although the mechanistic details remain to be further clarified. Similar initiation



mechanisms may be taken by RdRPs from *Flavivirus* and *Pestivirus* genera, and several residues within the priming element of DENV NS5 has been proposed to contribute to the initiation-related processes [53]. However, how the initiation platform is set up in these RdRPs still awaits clarification from relevant IC structures. In contrast to *de novo* viral RdRPs that take advantage of the priming element at initiation, primer-dependent viral RdRPs represented by PV 3D<sup>pol</sup> utilize a protein primer (termed VPg for “viral protein genome-linked”) or its uridylylated forms to initiate RNA synthesis [54,55]. To date, RdRP-VPg complex crystal structures have been reported in foot-and-mouth disease virus (FMDV), coxsackievirus B3 (CV B3), and enterovirus 71 (EV71), each displaying a distinct mode of interactions between the polymerase and VPg [56–58]. However, how VPg or uridylylated VPg is utilized in the template-directed polymerase initiation remains elusive, as none of the three polymerase-VPg structures includes an RNA template or provides clear clues for how such an IC is spatially arranged.



**Figure 3.** The crystal structures of *de novo* RdRP initiation complex (IC). (a) Bacteriophage  $\phi 6$  polymerase IC structure (pdb entry: 1HI0). Template RNA is in cyan. Initiation NTPs and residues Q629 and Y630 are shown as sticks, and magnesium ions are shown as cyan spheres; (b) HCV NS5B IC structure (pdb entry: 4WTJ). Template RNA is in cyan and dinucleotide primer is in green. ADP and residues Y448 and G449 are shown as sticks, and manganese ions are shown as cyan spheres. Structure superimpositions were carried out using motif C residues and the least-square method. For both panels, the +1 templating nucleotide is shown in orange. Capital letters with dark grey background indicate RdRP catalytic motifs.

The  $\phi 6$  and HCV polymerase IC structures could represent the situation in the first and second nucleotide addition cycles of the initiation process, respectively. From these one could deduce the possible outcome as the template-product RNA duplex grows to a length of about 7–8 base pairs and

occupies the entire front channel of the polymerase, a status observed in multiple *Picornaviridae* RdRP elongation complex (EC) crystal structures [26,44]. Presumably, the priming element would withdraw from the active site to accommodate the growth of RNA duplex, leading to a substantial but currently unresolved conformational change in the thumb domain. Whether the priming element adopts a certain fold in an RdRP EC, and whether it establishes new interactions with other regions of the EC or contributes to the processivity of the EC, are yet to be clarified. To date, solving a crystal structure of an EC with intact priming element remains a challenge [42,59], and if fulfilled, would be critical for answering these questions.

Among other basic aspects of RdRP catalysis, the details of elongation catalytic cycle are largely clarified in PV 3D<sup>pol</sup>, a representative of primer-dependent RdRP [26]. In this multi-step catalytic cycle, a pre-positioned +1 templating nucleotide directs the incoming NTP to bind in a site very close to its catalytic position. Next, the 2' and 3' hydroxyl groups of the NTP ribose likely trigger the active site closure featuring the movement of motifs A and D to allow a critical aspartic acid residue in motif A (JEV NS5 residue 536 or equivalent) to participate in the coordination of two catalytic magnesium ions. This closure of the active site leads to the phosphoryl transfer reaction, and is generally accepted as one of the slow steps in the catalytic cycle. Important post-catalysis events include the release of the pyrophosphate byproduct, the re-opening of the active site, and polymerase translocation to realign the active site to the next register on the template RNA. It is worth mentioning that the mechanism of translocation of viral RdRPs may be the most important part of the cycle that remains elusive [26]. Translocation of the single subunit Pol I family polymerases is coupled with active site reopening. In such a reopening process, the “O helix” in the fingers domain rotates significantly, and utilizes an aromatic side chain to push the nascent base pair toward the upstream [24,25]. As mentioned above, viral RdRPs close and reopen the active site primarily through the movement of motifs A and D in the palm, and there is no O helix equivalent in these proteins. Therefore, a different translocation mechanism has to be taken by viral RdRPs. Structurally, motif B intimately interacting with the nascent base pair on one side of the template strand and motif G that holds the other side of the template strand and makes close contact with the backbone of the template +1/+2 kink are the best candidates to participate in translocation [17,34]. However, more biochemical evidences as well as structural biology efforts in capturing possible low energy translocation intermediates are necessary to advance our understanding of this missing link in the RdRP catalytic cycle.

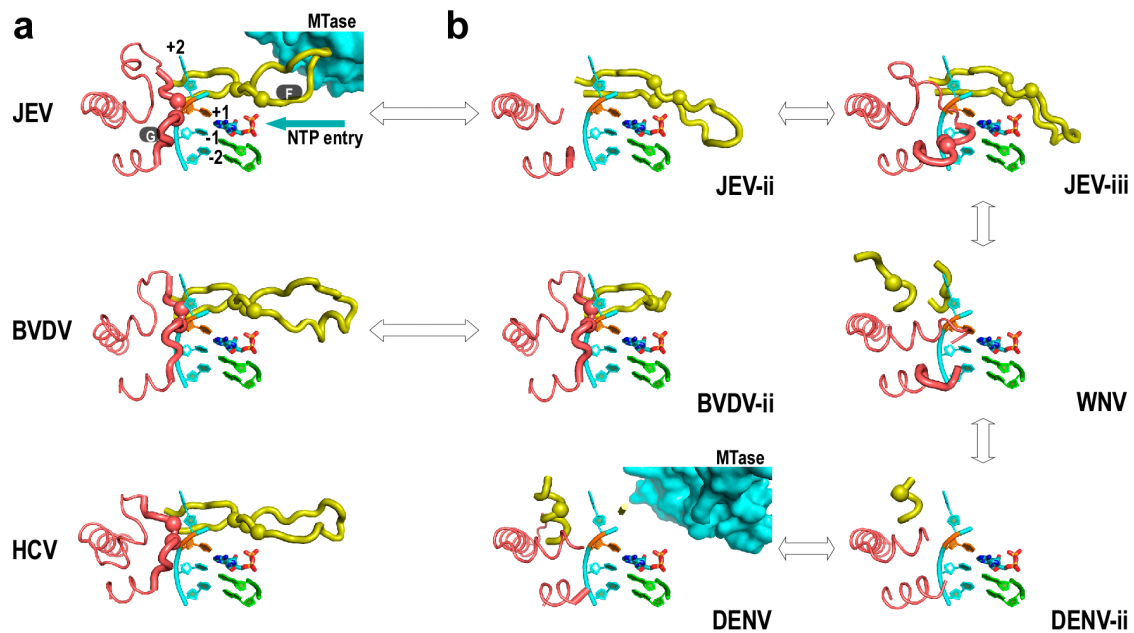
#### 4. *Flaviviridae* RdRP Regulation in *Cis*

In contrast to HCV NS5B, JEV NS5 and BVDV NS5B each has an N-terminal region that may provide *in cis* regulation of the core polymerase. This difference is reflected in the crystal structures of these RdRPs to some extent (Figure 4). The structure of HCV NS5B exhibits a canonical RdRP fold (Figure 4a, bottom panel) with all seven catalytic motifs properly arranged around the active site [50]. By contrast, the crystal structures of the *Flavivirus* NS5 in the absence of the N-terminal MTase consistently exhibit fingers domain disorder or non-canonical folding, particularly for motifs F and G (Figure 4b, JEV-ii, JEV-iii, WNV, and DENV-ii structures) [14,15,60]. It was until the report of the first full-length NS5 crystal structure that the intact and canonically folded *Flavivirus* polymerase was observed (Figure 4a, top panel). In this full-length JEV NS5 structure, motif F participates in



an intra-molecular interface between MTase and polymerase and contributes Phe467 to the formation of a six-residue hydrophobic network [17]. These hydrophobic residues are highly conserved in genus *Flavivirus*, implying the functional significance of the interface. When polar or charged residue mutations were introduced into these sites, virus replication levels were significantly affected [20]. Furthermore, the introduction of negatively charged Asp mutation into these sites could affect *in vitro* polymerase activities at both initiation and elongation [19]. Taken together, the MTase-polymerase interface observed in the full-length JEV NS5 structure play auxiliary roles to fingers domain folding and polymerase catalysis. In DENV NS5 studies, the removal of the MTase domain and/or further removal of the linker residues were shown to alter polymerase activities [61,62]. Very recently, a full-length DENV NS5 crystal structure was reported, and a different interface between MTase and polymerase was observed with less hydrophobic features [18]. In such a structure, motif F no longer participates in the interface and both motif F and motif G are largely disordered (Figure 5b, DENV structure) as with those MTase-less NS5 structures. As a natural fusion of MTase and polymerase, the full-length *Flavivirus* NS5 may have the MTase interact with the polymerase in different ways. However, the conformational state observed in the full-length JEV NS5 crystal structure with canonically folded motifs F and G may be regarded as the ground state for polymerase catalysis, as the highly conserved K/E/R residues in motif F (JEV NS5 residues 459, 461, 474 or equivalent) and the structurally critical residue pair in motif G (JEV NS5 409–410 or equivalent) are in-line for NTP/+1 templating nucleotide interactions and for template RNA binding, respectively (Figures 2b and 4a). Although the binding of NTP or an inhibitor molecule could affect the folding of motif F in JEV RdRP structures [60] or motif G in a DENV RdRP structure [63], respectively, the conformations of the NTP/inhibitor-induced regions apparently deviate from the canonical folds (see Figure 4b and JEV-iii for the conformation of the JEV RdRP-GTP structure).

The N-terminal region of *Pestivirus* NS5B contains  $\approx 90$  residues, and its function remains elusive. It is less likely to form a relatively independent enzyme module due to its limit in size, but could regulate the function of polymerase through intra-molecular interactions as observed in *Flavivirus* NS5. Two basic conformational states have been observed in reported BVDV NS5B crystal structures obtained using N-terminal truncated proteins. In one conformational state, the majority of the polymerase (residue 134 and beyond) adopts the canonical fold (Figure 4a, middle panel). However, residues 92–133 exist in a domain swapped mode and fold with another NS5B molecule [64]. The other conformational state was obtained using an NS5B carrying an Asn438 duplication [21]. In such a structure, residues 92–133 now folds with the rest of the same polypeptide chain. Interestingly, motif F and index finger in the Asn438 duplication structure were partially disordered (Figure 4b, BVDV-ii structure, index finger not shown), reminiscent of what is observed in *Flavivirus* polymerase structure in the absence of the N-terminal MTase. Therefore, it is also possible that the N-terminal region of *Pestivirus* NS5B could contribute to the proper folding and/or the catalytic regulation of the polymerase through yet unidentified intra-molecular interactions. Indeed, deletions within the N-terminal region could reduce or abolish *de novo* RNA synthesis activities in BVDV and CSFV systems [65,66].



**Figure 4.** Conformational heterogeneity of *Flaviviridae* polymerase structure may be related to *in cis* regulations. (a) Representative *Flaviviridae* polymerase structure adopting the canonical conformations (pdb entries: JEV/4K6M, BVDV/1S4F, HCV/1NB4); (b) Disorder and/or alternative folding of motifs F and G observed in *Flaviviridae* polymerases (pdb entries: JEV-ii/4MTP [60], JEV-iii/4HDG, WNV/2HFZ, DENV/4V0Q, DENV-ii/2J7U, BVDV-ii/2CJQ). Motifs F and G are shown as thick noodles and the pinky finger residues flanking motif G are shown as thin noodles. Note that in the apo JEV RdRP structure (JEV-ii) and GTP-bound JEV RdRP structure, the NTP entry channel is blocked by the non-canonically folded motif F. Color coding is as in Figures 1b and 3b. A 4 nt RNA template, a dinucleotide primer, and an ADP molecule taken from an HCV IC structure (pdb entry: 4WTJ) were modeled into all structures for comparison. The  $\alpha$ -carbons of JEV NS5 residues 409–410 in motif G and 459, 461, 474 in motif F and their equivalents in other polymerases are shown as spheres to help distinguish canonical and alternative folding of these two motifs. Double arrows are used to connect polymerases from the same viral species or from the same genus. The MTase in the full-length *Flavivirus* NS5 structures are shown as surface representations. Structure superimpositions were carried out as in Figure 1b.

## 5. Perspectives

As ideal systems to study the function of *de novo* viral polymerases, *Flaviviridae* RdRPs share a polymerase module that combines common RdRP motifs A through G with a priming element essential for primer-independent initiation. Major challenges remain in elucidating the mechanism of polymerase translocation on the RNA template and the conformational rearrangements primarily involving the priming element in transition from polymerase initiation to elongation. Within the two motifs most probable in participating the translocation process, motif B is shared by viral RdRPs and reverse transcriptases (RTs) and motif G is RdRP specific. Whether or not the primary translocation mechanism is shared by RdRPs and RTs, it probably differs from what is taken by the Pol I family polymerases based on structural divergences. In practice, the mystery of the priming element conformational changes

can be uncovered in any *de novo* viral RdRPs systems including those from *Cystoviridae*. With the low level of sequence conservation of the priming element considered, the details of the rearrangement during the transition process could be very different. However, the general strategy to complete these conformational changes is likely to be similar.

The *in cis* regulation on some of the *Flaviviridae* RdRPs by their N-terminal domain has brought in interactions and functions beyond the fundamental processes of RdRP catalysis. In *Flavivirus* NS5, crystal structures of the full-length protein have shed light on how the regulation could take place, and also provide valuable clues to investigate the function of NS5 beyond polymerase catalysis. *Pestivirus* NS5B still awaits a high-resolution full-length structure to reveal the true face of the N-terminal domain and its relationship with the polymerase module.

### Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant 31370198), the Open Research Fund Program of the Key Laboratory of Special Pathogens and Biosafety, Wuhan Institute of Virology, Chinese Academy of Sciences, China (grant 2015SPCAS003), and the “One-Three-Five” Strategic Programs, Wuhan Institute of Virology, Chinese Academy of Sciences, China.

### Author Contributions

Analyzed the previously reported structural data: Peng Gong; Wrote the paper: Peng Gong, Jiqin Wu, and Weichi Liu.

### Conflicts of Interest

The authors declare no conflict of interest.

### References

1. Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimotohno, K. Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9524–9528.
2. Westaway, E.G.; Brinton, M.A.; Gaidamovich, S.; Horzinek, M.C.; Igarashi, A.; Kaariainen, L.; Lvov, D.K.; Porterfield, J.S.; Russell, P.K.; Trent, D.W. *Flaviviridae*. *Intervirology* **1985**, *24*, 183–192.
3. Westaway, E.G.; Brinton, M.A.; Gaidamovich, S.; Horzinek, M.C.; Igarashi, A.; Kaariainen, L.; Lvov, D.K.; Porterfield, J.S.; Russell, P.K.; Trent, D.W. *Togaviridae*. *Intervirology* **1985**, *24*, 125–139.
4. Ito, T.; Tahara, S.M.; Lai, M.M. The 3'-untranslated region of hepatitis C virus RNA enhances translation from an internal ribosomal entry site. *J. Virol.* **1998**, *72*, 8789–8796.
5. Filomatori, C.V.; Lodeiro, M.F.; Alvarez, D.E.; Samsa, M.M.; Pietrasanta, L.; Gamarnik, A.V. A 5' RNA element promotes dengue virus RNA synthesis on a circular genome. *Genes Dev.* **2006**, *20*, 2238–2249.

6. Zhang, B.; Dong, H.; Zhou, Y.; Shi, P.Y. Genetic interactions among the West Nile virus methyltransferase, the RNA-dependent RNA polymerase, and the 5' stem-loop of genomic RNA. *J. Virol.* **2008**, *82*, 7047–7058.
7. Friebe, P.; Lohmann, V.; Krieger, N.; Bartenschlager, R. Sequences in the 5' nontranslated region of hepatitis C virus required for RNA replication. *J. Virol.* **2001**, *75*, 12047–12057.
8. Bisailon, M.; Lemay, G. Viral and cellular enzymes involved in synthesis of mRNA cap structure. *Virology* **1997**, *236*, 1–7.
9. Tsukiyama-Kohara, K.; Iizuka, N.; Kohara, M.; Nomoto, A. Internal ribosome entry site within hepatitis C virus RNA. *J. Virol.* **1992**, *66*, 1476–1483.
10. Poole, T.L.; Wang, C.; Popp, R.A.; Potgieter, L.N.; Siddiqui, A.; Collett, M.S. Pestivirus translation initiation occurs by internal ribosome entry. *Virology* **1995**, *206*, 750–754.
11. Chambers, T.J.; Hahn, C.S.; Galler, R.; Rice, C.M. Flavivirus genome organization, expression, and replication. *Annu. Rev. Microbiol.* **1990**, *44*, 649–688.
12. Romero-Brey, I.; Bartenschlager, R. Membranous replication factories induced by plus-strand RNA viruses. *Viruses* **2014**, *6*, 2826–2857.
13. Egloff, M.P.; Benarroch, D.; Selisko, B.; Romette, J.L.; Canard, B. An RNA cap (nucleoside-2'-O)-methyltransferase in the flavivirus RNA polymerase NS5: Crystal structure and functional characterization. *EMBO J.* **2002**, *21*, 2757–2768.
14. Malet, H.; Egloff, M.P.; Selisko, B.; Butcher, R.E.; Wright, P.J.; Roberts, M.; Gruez, A.; Sulzenbacher, G.; Vonnrhein, C.; Bricogne, G.; *et al.* Crystal structure of the RNA polymerase domain of the West Nile virus non-structural protein 5. *J. Biol. Chem.* **2007**, *282*, 10678–10689.
15. Yap, T.L.; Xu, T.; Chen, Y.L.; Malet, H.; Egloff, M.P.; Canard, B.; Vasudevan, S.G.; Lescar, J. Crystal structure of the dengue virus RNA-dependent RNA polymerase catalytic domain at 1.85-angstrom resolution. *J. Virol.* **2007**, *81*, 4753–4765.
16. Geiss, B.J.; Thompson, A.A.; Andrews, A.J.; Sons, R.L.; Gari, H.H.; Keenan, S.M.; Peersen, O.B. Analysis of flavivirus NS5 methyltransferase cap binding. *J. Mol. Biol.* **2009**, *385*, 1643–1654.
17. Lu, G.; Gong, P. Crystal structure of the full-length Japanese Encephalitis virus NS5 reveals a conserved methyltransferase-polymerase interface. *PLoS Pathog.* **2013**, *9*, e1003549.
18. Zhao, Y.; Soh, T.S.; Zheng, J.; Chan, K.W.; Phoo, W.W.; Lee, C.C.; Tay, M.Y.; Swaminathan, K.; Cornvik, T.C.; Lim, S.P.; *et al.* A crystal structure of the dengue virus NS5 protein reveals a novel inter-domain interface essential for protein flexibility and virus replication. *PLoS Pathog.* **2015**, *11*, e1004682.
19. Wu, J.; Lu, G.; Zhang, B.; Gong, P. Perturbation in the conserved methyltransferase-polymerase interface of flavivirus NS5 differentially affects polymerase initiation and elongation. *J. Virol.* **2015**, *89*, 249–261.
20. Li, X.D.; Shan, C.; Deng, C.L.; Ye, H.Q.; Shi, P.Y.; Yuan, Z.M.; Gong, P.; Zhang, B. The interface between methyltransferase and polymerase of NS5 is essential for flavivirus replication. *PLoS Negl. Trop. Dis.* **2014**, *8*, e2891.
21. Choi, K.H.; Gallei, A.; Becher, P.; Rossmann, M.G. The structure of bovine viral diarrhea virus RNA-dependent RNA polymerase and its amino-terminal domain. *Structure* **2006**, *14*, 1107–1113.
22. Theobald, D.L.; Wuttke, D.S. THESEUS: Maximum likelihood superpositioning and analysis of macromolecular structures. *Bioinformatics* **2006**, *22*, 2171–2172.

23. Thompson, A.A.; Peersen, O.B. Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. *EMBO J.* **2004**, *23*, 3462–3471.
24. Yin, Y.W.; Steitz, T.A. The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell* **2004**, *116*, 393–404.
25. Li, Y.; Korolev, S.; Waksman, G. Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: Structural basis for nucleotide incorporation. *EMBO J.* **1998**, *17*, 7514–7525.
26. Gong, P.; Peersen, O.B. Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 22505–22510.
27. Sun, J.H.; Kao, C.C. Characterization of RNA products associated with or aborted by a viral RNA-dependent RNA polymerase. *Virology* **1997**, *236*, 348–353.
28. Butcher, S.J.; Grimes, J.M.; Makeyev, E.V.; Bamford, D.H.; Stuart, D.I. A mechanism for initiating RNA-dependent RNA polymerization. *Nature* **2001**, *410*, 235–240.
29. Nomoto, A.; Detjen, B.; Pozzatti, R.; Wimmer, E. The location of the polio genome protein in viral RNAs and its implication for RNA synthesis. *Nature* **1977**, *268*, 208–213.
30. Gorbalenya, A.E.; Pringle, F.M.; Zeddam, J.L.; Luke, B.T.; Cameron, C.E.; Kalkmakoff, J.; Hanzlik, T.N.; Gordon, K.H.; Ward, V.K. The palm subdomain-based active site is internally permuted in viral RNA-dependent RNA polymerases of an ancient lineage. *J. Mol. Biol.* **2002**, *324*, 47–62.
31. Poch, O.; Sauvaget, I.; Delarue, M.; Tordo, N. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J.* **1989**, *8*, 3867–3874.
32. Beese, L.S.; Steitz, T.A. Structural basis for the 3′–5′ exonuclease activity of *Escherichia coli* DNA polymerase I: A two metal ion mechanism. *EMBO J.* **1991**, *10*, 25–33.
33. Crotty, S.; Gohara, D.; Gilligan, D.K.; Karelsky, S.; Cameron, C.E.; Andino, R. Manganese-dependent polioviruses caused by mutations within the viral polymerase. *J. Virol.* **2003**, *77*, 5378–5388.
34. Sholders, A.J.; Peersen, O.B. Distinct conformations of a putative translocation element in poliovirus polymerase. *J. Mol. Biol.* **2014**, *426*, 1407–1419.
35. Castro, C.; Smidansky, E.D.; Arnold, J.J.; Maksimchuk, K.R.; Moustafa, I.; Uchida, A.; Gotte, M.; Konigsberg, W.; Cameron, C.E. Nucleic acid polymerases use a general acid for nucleotidyl transfer. *Nat. Struct. Mol. Biol.* **2009**, *16*, 212–218.
36. Takeshita, D.; Tomita, K. Molecular basis for RNA polymerization by Qbeta replicase. *Nat. Struct. Mol. Biol.* **2012**, *19*, 229–237.
37. Pan, J.; Vakharia, V.N.; Tao, Y.J. The structure of a birnavirus polymerase reveals a distinct active site topology. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7385–7390.
38. Reich, S.; Guilligay, D.; Pflug, A.; Malet, H.; Berger, I.; Crepin, T.; Hart, D.; Lunardi, T.; Nanao, M.; Ruigrok, R.W.; *et al.* Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature* **2014**, *516*, 361–366.
39. Huang, H.; Chopra, R.; Verdine, G.L.; Harrison, S.C. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: Implications for drug resistance. *Science* **1998**, *282*, 1669–1675.
40. Gillis, A.J.; Schuller, A.P.; Skordalakes, E. Structure of the *Tribolium castaneum* telomerase catalytic subunit TERT. *Nature* **2008**, *455*, 633–637.

41. Bruenn, J.A. A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. *Nucleic Acids Res.* **2003**, *31*, 1821–1829.
42. Appleby, T.C.; Perry, J.K.; Murakami, E.; Barauskas, O.; Feng, J.; Cho, A.; Fox, D., 3rd; Wetmore, D.R.; McGrath, M.E.; Ray, A.S.; *et al.* Viral replication. structural basis for RNA replication by the hepatitis C virus polymerase. *Science* **2015**, *347*, 771–775.
43. Iglesias, N.G.; Filomatori, C.V.; Gamarnik, A.V. The F1 motif of dengue virus polymerase NS5 is involved in promoter-dependent RNA synthesis. *J. Virol.* **2011**, *85*, 5745–5756.
44. Gong, P.; Kortus, M.G.; Nix, J.C.; Davis, R.E.; Peersen, O.B. Structures of coxsackievirus, rhinovirus, and poliovirus polymerase elongation complexes solved by engineering RNA mediated crystal contacts. *PLoS ONE* **2013**, *8*, e60272.
45. Zamyatkin, D.F.; Parra, F.; Alonso, J.M.; Harki, D.A.; Peterson, B.R.; Grochulski, P.; Ng, K.K. Structural insights into mechanisms of catalysis and inhibition in Norwalk virus polymerase. *J. Biol. Chem.* **2008**, *283*, 7705–7712.
46. Ferrer-Orta, C.; Arias, A.; Perez-Luque, R.; Escarmis, C.; Domingo, E.; Verdaguer, N. Structure of foot-and-mouth disease virus RNA-dependent RNA polymerase and its complex with a template-primer RNA. *J. Biol. Chem.* **2004**, *279*, 47212–47221.
47. Carpousis, A.J.; Gralla, J.D. Interaction of RNA polymerase with *lacUV5* promoter DNA during mRNA initiation and elongation. Footprinting, methylation, and rifampicin-sensitivity changes accompanying transcription initiation. *J. Mol. Biol.* **1985**, *183*, 165–177.
48. Martin, C.T.; Muller, D.K.; Coleman, J.E. Processivity in early stages of transcription by T7 RNA polymerase. *Biochemistry* **1988**, *27*, 3966–3974.
49. Gong, P.; Martin, C.T. Mechanism of instability in abortive cycling by T7 RNA polymerase. *J. Biol. Chem.* **2006**, *281*, 23533–23544.
50. Lesburg, C.A.; Cable, M.B.; Ferrari, E.; Hong, Z.; Mannarino, A.F.; Weber, P.C. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat. Struct. Biol.* **1999**, *6*, 937–943.
51. Hong, Z.; Cameron, C.E.; Walker, M.P.; Castro, C.; Yao, N.; Lau, J.Y.; Zhong, W. A novel mechanism to ensure terminal initiation by hepatitis C virus NS5B polymerase. *Virology* **2001**, *285*, 6–11.
52. Cherry, A.L.; Dennis, C.A.; Baron, A.; Eisele, L.E.; Thommes, P.A.; Jaeger, J. Hydrophobic and charged residues in the C-terminal arm of hepatitis C virus RNA-dependent RNA polymerase regulate initiation and elongation. *J. Virol.* **2015**, *89*, 2052–2063.
53. Selisko, B.; Potisopon, S.; Agred, R.; Priet, S.; Varlet, I.; Thillier, Y.; Sallamand, C.; Debart, F.; Vasseur, J.J.; Canard, B. Molecular basis for nucleotide conservation at the ends of the dengue virus genome. *PLoS Pathog.* **2012**, *8*, e1002912.
54. Takeda, N.; Kuhn, R.J.; Yang, C.F.; Takegami, T.; Wimmer, E. Initiation of poliovirus plus-strand RNA synthesis in a membrane complex of infected HeLa cells. *J. Virol.* **1986**, *60*, 43–53.
55. Paul, A.V.; Wimmer, E. Initiation of protein-primed picornavirus RNA synthesis. *Virus Res.* **2015**, doi:10.1016/j.virusres.2014.12.028.
56. Ferrer-Orta, C.; Arias, A.; Agudo, R.; Perez-Luque, R.; Escarmis, C.; Domingo, E.; Verdaguer, N. The structure of a protein primer-polymerase complex in the initiation of genome replication. *EMBO J.* **2006**, *25*, 880–888.

57. Gruez, A.; Selisko, B.; Roberts, M.; Bricogne, G.; Bussetta, C.; Jabafi, I.; Coutard, B.; de Palma, A.M.; Neyts, J.; Canard, B. The crystal structure of coxsackievirus B3 RNA-dependent RNA polymerase in complex with its protein primer VPg confirms the existence of a second VPg binding site on *Picornaviridae* polymerases. *J. Virol.* **2008**, *82*, 9577–9590.
58. Chen, C.; Wang, Y.; Shan, C.; Sun, Y.; Xu, P.; Zhou, H.; Yang, C.; Shi, P.Y.; Rao, Z.; Zhang, B.; *et al.* Crystal structure of enterovirus 71 RNA-dependent RNA polymerase complexed with its protein primer VPg: Implication for a *trans* mechanism of VPg uridylylation. *J. Virol.* **2013**, *87*, 5755–5768.
59. Mosley, R.T.; Edwards, T.E.; Murakami, E.; Lam, A.M.; Grice, R.L.; Du, J.; Sofia, M.J.; Furman, P.A.; Otto, M.J. Structure of hepatitis C virus polymerase in complex with primer-template RNA. *J. Virol.* **2012**, *86*, 6503–6511.
60. Surana, P.; Satchidanandam, V.; Nair, D.T. RNA-dependent RNA polymerase of Japanese encephalitis virus binds the initiator nucleotide GTP to form a mechanistically important pre-initiation state. *Nucleic Acids Res.* **2014**, *42*, 2758–2773.
61. Lim, S.P.; Koh, J.H.; Seh, C.C.; Liew, C.W.; Davidson, A.D.; Chua, L.S.; Chandrasekaran, R.; Cornvik, T.C.; Shi, P.Y.; Lescar, J. A crystal structure of the dengue virus non-structural protein 5 (NS5) polymerase delineates interdomain amino acid residues that enhance its thermostability and *de novo* initiation activities. *J. Biol. Chem.* **2013**, *288*, 31105–31114.
62. Potisopon, S.; Priet, S.; Collet, A.; Decroly, E.; Canard, B.; Selisko, B. The methyltransferase domain of dengue virus protein NS5 ensures efficient RNA synthesis initiation and elongation by the polymerase domain. *Nucleic Acids Res.* **2014**, *42*, 11642–11656.
63. Noble, C.G.; Lim, S.P.; Chen, Y.L.; Liew, C.W.; Yap, L.; Lescar, J.; Shi, P.Y. Conformational Flexibility of the dengue virus RNA-dependent RNA polymerase revealed by a complex with an inhibitor. *J. Virol.* **2013**, *87*, 5291–5295.
64. Choi, K.H.; Groarke, J.M.; Young, D.C.; Rossmann, M.G.; Pevear, D.C.; Kuhn, R.J.; Smith, J.L. Design, expression, and purification of a *Flaviviridae* polymerase using a high-throughput approach to facilitate crystal structure determination. *Protein Sci.* **2004**, *13*, 2685–2692.
65. Xiao, M.; Li, H.; Wang, Y.; Wang, X.; Wang, W.; Peng, J.; Chen, J.; Li, B. Characterization of the N-terminal domain of classical swine fever virus RNA-dependent RNA polymerase. *J. Gen. Virol.* **2006**, *87*, 347–356.
66. Lai, V.C.; Kao, C.C.; Ferrari, E.; Park, J.; Uss, A.S.; Wright-Minogue, J.; Hong, Z.; Lau, J.Y. Mutational analysis of bovine viral diarrhea virus RNA-dependent RNA polymerase. *J. Virol.* **1999**, *73*, 10129–10136.