



**HAL**  
open science

## Molecular Basis for Nucleotide Conservation at the Ends of the Dengue Virus Genome

Barbara Selisko, Supanee Potisopon, Rym Agred, Stephane Priet, Isabelle Varlet, Yann Thillier, Corinne Sallamand, Françoise Debart, Jean-Jacques Vasseur, Bruno Canard

► **To cite this version:**

Barbara Selisko, Supanee Potisopon, Rym Agred, Stephane Priet, Isabelle Varlet, et al.. Molecular Basis for Nucleotide Conservation at the Ends of the Dengue Virus Genome. PLoS Pathogens, 2012, 8 (9), pp.e1002912. 10.1371/journal.ppat.1002912 . hal-00733199

**HAL Id: hal-00733199**

**<https://hal.science/hal-00733199>**

Submitted on 5 Oct 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Molecular Basis for Nucleotide Conservation at the Ends of the Dengue Virus Genome

Barbara Selisko<sup>1\*</sup>, Supanee Potisopon<sup>1</sup>, Rym Agred<sup>1</sup>, Stéphane Priet<sup>1</sup>, Isabelle Varlet<sup>1</sup>, Yann Thillier<sup>2</sup>, Corinne Sallamand<sup>2</sup>, Françoise Debart<sup>2</sup>, Jean-Jacques Vasseur<sup>2</sup>, Bruno Canard<sup>1\*</sup>

<sup>1</sup> Aix-Marseille Université, CNRS, AFMB UMR 7257, 163, Marseille, France, <sup>2</sup> IBMM, UMR 5247 CNRS-UM1-UM2, cc 1704 Université Montpellier 2, Montpellier, France

## Abstract

The dengue virus (DV) is an important human pathogen from the *Flavivirus* genus, whose genome- and antigenome RNAs start with the strictly conserved sequence pppAG. The RNA-dependent RNA polymerase (RdRp), a product of the NS5 gene, initiates RNA synthesis *de novo*, *i.e.*, without the use of a pre-existing primer. Very little is known about the mechanism of this *de novo* initiation and how conservation of the starting adenosine is achieved. The polymerase domain NS5Pol<sub>DV</sub> of NS5, upon initiation on viral RNA templates, synthesizes mainly dinucleotide primers that are then elongated in a processive manner. We show here that NS5Pol<sub>DV</sub> contains a specific priming site for adenosine 5'-triphosphate as the first transcribed nucleotide. Remarkably, in the absence of any RNA template the enzyme is able to selectively synthesize the dinucleotide pppAG when Mn<sup>2+</sup> is present as catalytic ion. The T794 to A799 priming loop is essential for initiation and provides at least part of the ATP-specific priming site. The H798 loop residue is of central importance for the ATP-specific initiation step. In addition to ATP selection, NS5Pol<sub>DV</sub> ensures the conservation of the 5'-adenosine by strongly discriminating against viral templates containing an erroneous 3'-end nucleotide in the presence of Mg<sup>2+</sup>. In the presence of Mn<sup>2+</sup>, NS5Pol<sub>DV</sub> is remarkably able to generate and elongate the correct pppAG primer on these erroneous templates. This can be regarded as a genomic/antigenomic RNA end repair mechanism. These conservational mechanisms, mediated by the polymerase alone, may extend to other RNA virus families having RdRps initiating RNA synthesis *de novo*.

**Citation:** Selisko B, Potisopon S, Agred R, Priet S, Varlet I, et al. (2012) Molecular Basis for Nucleotide Conservation at the Ends of the Dengue Virus Genome. *PLoS Pathog* 8(9): e1002912. doi:10.1371/journal.ppat.1002912

**Editor:** Michael S. Diamond, Washington University School of Medicine, United States of America

**Received:** July 6, 2012; **Accepted:** August 3, 2012; **Published:** September 13, 2012

**Copyright:** © 2012 Selisko et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work is supported by the Fondation pour la Recherche Médicale (Programme Équipe FRM to B. Canard), by the Infectiopôle-Sud (PhD scholarship to S. Potisopon) and by the European Union Seventh Framework Programme (FP7/2007-20013) through the SILVER project (Small inhibitor leads against emerging RNA viruses, grant agreement 260644). Y. Thillier thanks the Ministère pour la Recherche et la Technologie for the award of a research studentship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: Barbara.Selisko@afmb.univ-mrs.fr (BS); Bruno.Canard@afmb.univ-mrs.fr (BC)

## Introduction

Most RNA viruses maintain the specific sequences present at the ends of their genomes. The 5' genome end may carry a cap structure to ensure both genome stability and efficient translation [1]. The 3'-end may carry a poly(A) tail or adopt specific 3'-end sequences required for viral replication [2,3]. They are generally copied exactly to avoid loss of genetic information, and have supposedly evolved towards optimum replication efficiency. Terminal genome damage can be caused by errors introduced by the viral polymerase during initiation and termination, or by cellular ribonucleases [4]. In addition to special mechanisms to ensure efficient initiation of RNA synthesis, viruses have evolved mechanisms to repair or correct damaged extremities such as the use of abortive transcripts as primers, the generation and use of non-templated primers, and the addition of one or few non-templated nucleotides to the 3'-end by a terminal transferase activity [4]. However, our knowledge about these mechanisms is still very limited. Many RNA virus polymerases, which do not use a primer and thus initiate RNA synthesis *de novo*, generate abortive transcripts during the initiation phase of RNA synthesis [5,6,7]. Primer-mediated repair of template extremities was so far only demonstrated for the positive-strand RNA (+RNA) turnip crinkle virus (TCV) [8]. Non-templated primer synthesis by the viral

polymerase might be involved in the repair mechanism of TCV [9]. Such mechanism was also proposed as the molecular basis of the reconstitution of 5'-ends of negative-strand RNA (-RNA) respiratory syncytial virus (RSV) replicons [10]. In this study we demonstrate how the dengue virus (DV) RNA-dependent RNA polymerase (RdRp), which starts RNA synthesis *de novo*, plays a decisive role in the nucleotide conservation of viral RNA ends.

DV belongs to the *Flavivirus* genus within the +RNA virus family of *Flaviviridae* together with viruses of the genera *Hepacivirus* and *Pestivirus* [11]. The *Flavivirus* genus comprises around 50 virus species [12] including major human pathogens such as DV, yellow fever virus (YFV), West Nile virus (WNV) and Japanese encephalitis virus (JEV). Flaviviruses harbour the RdRp activity in the C-terminal domain (amino acids 272–900) of non-structural protein NS5 [13,14,15,16,17]. The N-terminal domain contains methyltransferase activities involved in RNA capping [18,19]. Evidence has been presented that the N-terminal domain of NS5 also harbours the central RNA capping guanylyltransferase activity [20]. The structure of full-length NS5 is not known but several structures of methyltransferase domains have been determined (for review see [21]). Likewise, crystal structures of *Flavivirus* NS5 RdRp domains have been determined for DV [16] and WNV [22]. All structurally characterized viral RdRps so far adopt the basic fold of the SCOP superfamily of DNA/RNA

## Author Summary

The 5'- and 3'-ends of RNA virus genomes have evolved towards efficient replication, translation, and escape from defense mechanisms of the host cell. Little is known about how RNA viruses conserve or restore the correct ends of their genomes. The *Flavivirus* genus of positive-strand RNA viruses contains important human pathogens such as yellow fever virus, West Nile virus, Japanese encephalitis virus and dengue virus (DV). The *Flavivirus* genome ends are strictly conserved as 5'-AG...CU-3'. We demonstrate here the primary role of the DV polymerase in the conservation of the first and last genomic residue. We show that DV polymerase contains an ATP-specific priming site, which imposes a strong preference for the *de novo* synthesis of a dinucleotide primer starting with an ATP. Furthermore, the polymerase is able to indirectly correct erroneous sequences by producing the correct primer in the absence of template and on templates containing incorrect nucleotides at the 3'-end. The correct primer is productively elongated on either correct or incorrect templates. Our findings provide a direct demonstration of the implication of a viral RNA polymerase in the conservation and repair of genome ends. Other polymerases from other RNA virus families are likely to employ similar mechanisms.

polymerases. As the other subgroups of this superfamily, DNA-dependent DNA polymerases (DdDp, prototype Klenow fragment of the *E.coli* DdDp I), RNA-dependent DNA polymerase (prototype HIV reverse transcriptase) and DNA-dependent RNA polymerases (DdRp, prototype bacteriophage T7 DdRp), their apo-structure is usually likened to a right hand comprising fingers, palm and thumb subdomains. Viral RdRps contain an encircled active site having connecting elements between the fingers and thumb subdomains. Active sites of viral RdRps performing *de novo* RNA synthesis are additionally closed in their initiation conformation due to the existence of structural elements allowing the stable positioning of the first NTP into a priming site [23,24]. All *Flaviviridae* RdRps studied so far initiate RNA synthesis *de novo*. Accordingly, *Flavivirus* RdRp domain structures contain a "priming loop" in the thumb subdomain closing the catalytic site [16,22]. The putative priming loop of DV RdRp was defined as comprising residues 792 to 804. Of particular interest are two aromatic residues near the tip of the loop, W795 and H798, which are conserved in all *Flavivirus* RdRps. They might play the role of an initiation platform to which the base of the priming NTP stacks as it was shown for bacteriophage  $\phi 6$  [23] and proposed for HCV and BVDV RdRps [25,26]. Structures of DV RdRp in complex with 3'dGTP as well as two models of *de novo* initiation complexes of DV and WNV RdRps favor Trp795 in the role of the initiation platform [16,22].

Genomes of *Flaviviridae* lack a poly(A) tail at the 3'-end. A remarkable trait of *Flavivirus* genomes is the strict conservation of the 5'- and 3'-end dinucleotides as 5' AG...CU 3'. The molecular basis for this strict conservation of the 5'- and 3'-end dinucleotides and/or the use of the same starting nucleotide for +RNA and -RNA strand synthesis by the viral polymerases is not known. Its *Hepacivirus* and *Pestivirus* counterparts have to display higher nucleotide tolerance. They are able to initiate with (A/G)C and G(G/U), respectively, since the 5'- and 3'-ends of *Hepacivirus* genomes of different genotypes correspond to 5' (A/G)C...GU 3' and the genomes of *pestiviruses* to 5' GU...CC 3'. Interestingly,

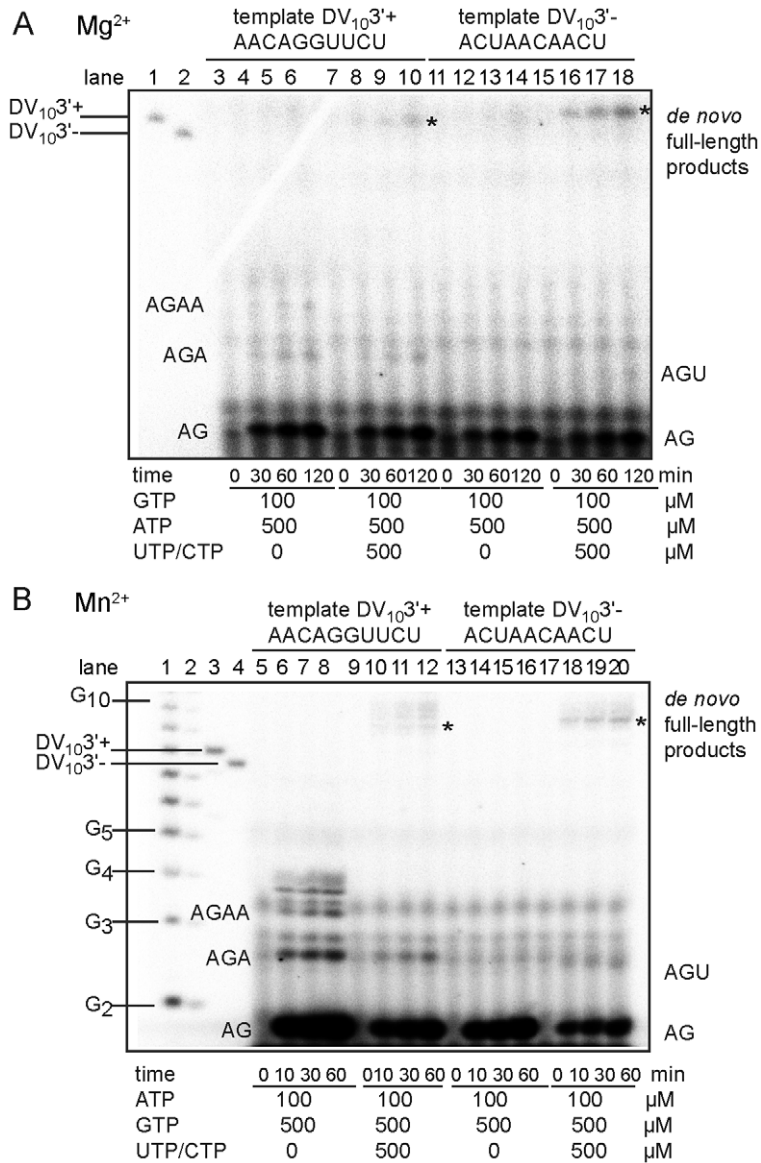
genomes and antigenomes of non-segmented -RNA (ns-RNA) paramyxoviruses, whose RdRps perform *de novo* RNA synthesis, start with a conserved 5'-AC [10].

Here we show that the strict sequence conservation of *Flavivirus* genome ends is entirely polymerase-encoded. We demonstrate ATP-specific *de novo* initiation using the RdRp domain of DV protein NS5 (NS5Pol<sub>DV</sub>) and specific 10-mer oligonucleotidic RNA templates corresponding to the 3'-end of genomic +RNA and -RNA. We document the existence of a built-in ATP-specific priming site of NS5Pol<sub>DV</sub>. This specific site is one of the means by which NS5Pol<sub>DV</sub> ensures that the DV genome and antigenome start with an A, the others being several correction mechanisms including the generation of non-templated pppAG primers as well as the preferential formation and elongation of pppAG even on templates with non-cognate 3'-ends. Finally, we show that the ATP-specific priming site is part of the putative priming loop coming from the thumb subdomain. There, residue H798, and not W795, is essential for *de novo* initiation and may act as a priming platform stabilizing the ATP priming nucleotide. DV RdRp is actively involved in the conservation of the correct ends of the genome proving thus a direct example of how RNA viruses maintain the integrity of their genomes. The mechanisms described here may more broadly apply to other RNA viruses having viral RdRps able to initiate RNA synthesis *de novo*.

## Results

### NS5Pol<sub>DV</sub> generates pppAG by abortive *de novo* initiation on short RNA templates

We set out to study primer synthesis by the RdRp domain of dengue virus protein NS5 (NS5Pol<sub>DV</sub>) using small specific templates corresponding to the 3'-ends of the genome (+RNA) and the antigenome (-RNA). Templates are comprised of 10 nucleotides and are predicted to be devoid of stable secondary structure (see Materials and Methods). Both templates end with the dinucleotide 5'-CU-3'. Product formation over time was followed using either ATP and GTP, or all NTPs needed to form a full-length product when synthesis is precisely started at the 3'-end of the template. Figure 1 shows reaction kinetics of RNA synthesis on DV<sub>103'+</sub> corresponding to the 3'-end of the RNA genome 5'-AACAGGUUCU-3' (left) and on DV<sub>103'-</sub> corresponding to that of the antigenome 5'-ACUACAACU-3' (right). We used either [ $\alpha$ -<sup>32</sup>P]-GTP ( $\alpha$ GTP, panel A) or [ $\gamma$ -<sup>32</sup>P]-ATP ( $\gamma$ ATP, panel B) as the radioactive nucleotide. For the catalytic ion, either Mg<sup>2+</sup> (panel A) or Mg<sup>2+</sup> supplemented with Mn<sup>2+</sup> (panel B) were used at their optimum concentrations 5 mM for Mg<sup>2+</sup> and 2 mM for Mn<sup>2+</sup> [14]. Reactions with ATP and GTP render time-dependent accumulation of a short product migrating below the marker G2 (see panel B). Comparison with authentic unlabeled pppAG (see Materials and Methods) visualized using UV-shadowing indicated that it indeed corresponds to pppAG (not shown), the expected product of the first step of *de novo* RNA synthesis. When DV<sub>103'+</sub> is used as a template, pppAG is formed as well as pppAGA and pppAGAA. When all NTPs are used, pppAG accumulates with time as does pppAGA in the case of DV<sub>103'+</sub> and pppAGU in the case of DV<sub>103'-</sub>. After the synthesis of trinucleotides NSPol<sub>DV</sub> adopts a processive RNA synthesis elongation mode to continue synthesis up to full-length products (labeled by asterisks in Figure 1). As we had observed before [14], when using Mn<sup>2+</sup> the reaction is much more efficient and allows for the use of [ $\gamma$ -<sup>32</sup>P]-ATP ( $\gamma$ ATP) as radiolabeled nucleotide in order to visualize exclusively *de novo* RNA synthesis products starting with ATP. The pattern observed with Mg<sup>2+</sup> is reproduced when Mn<sup>2+</sup> is present (Figure 1B). One difference is that the use of Mn<sup>2+</sup>



**Figure 1. De novo initiation by NS5Pol<sub>DV</sub> on oligonucleotides corresponding to the 3'-ends of dengue virus genome and antigenome.** (A) De novo initiation in the presence of  $Mg^{2+}$  ions. Reaction mixtures were prepared as described in Materials and Methods plus 5 mM  $MgCl_2$ , 10  $\mu M$  template, 500 nM NS5Pol<sub>DV</sub> and the given concentrations of NTPs. Radiolabeled GTP ( $\alpha$ -GTP) was used. Reactions were started by the addition of  $MgCl_2$  and reactions incubated for indicated time periods. Samples were analyzed by PAGE and autoradiography. Markers in lanes 1 and 2 are labeled DV<sub>10</sub>3'+ and DV<sub>10</sub>3'-, respectively. DV<sub>10</sub>3'+ (lanes 3 to 10) and DV<sub>10</sub>3'- (lanes 11 to 18) were used as templates. Nucleotide sequences are given above the panel. Identities of labeled product bands are given on the right and left side of the gel. Full-length products are labeled by an asterisk. (B) De novo initiation in the presence of  $Mn^{2+}$  ions. Reaction mixtures were prepared as indicated in Materials and Methods plus 5 mM  $MgCl_2$ , 2 mM  $MnCl_2$ , 1  $\mu M$  template, 500 nM NS5Pol<sub>DV</sub> and the given concentrations of NTPs. Radiolabeled ATP ( $\gamma$ -ATP) was used. Reactions were started by the addition of  $MnCl_2$ , incubated for given time periods, and analyzed by PAGE and autoradiography. Markers in lanes 1 to 4 include an oligoG-ladder (lanes 1 and 2), labeled DV<sub>10</sub>3'+ (lane 3) and DV<sub>10</sub>3'- (lane 4). DV<sub>10</sub>3'+ (lanes 5 to 12) and DV<sub>10</sub>3'- (lanes 13 to 20) were used as templates. Identities of labeled product bands are given on the right and left side of the reaction lanes. Full-length products are labeled by an asterisk.  
 doi:10.1371/journal.ppat.1002912.g001

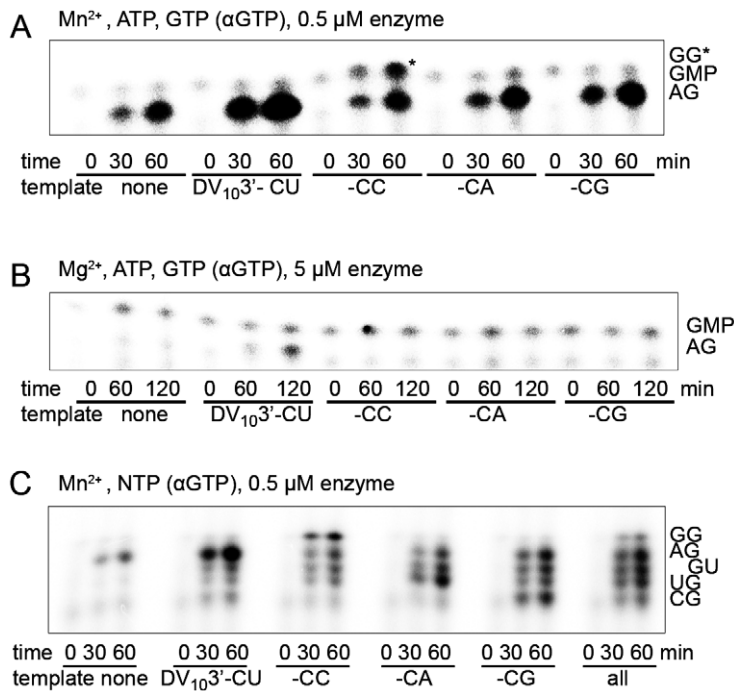
results in longer full-length products, which might be caused by an alteration of the terminal nucleotide transferase activity of NS5Pol<sub>DV</sub> [14,27,28].

In conclusion, using RNA templates mimicking viral sequences, dinucleotide and trinucleotide products are formed during initiation and before processive RNA elongation, the most abundant being the dinucleotide pppAG.

#### NS5Pol<sub>DV</sub> contains a built-in ATP-specific priming site for de novo RNA synthesis initiation

The first nucleotide of *Flavivirus* genomes is an adenosine, followed by a guanosine. This 5'-pppAG sequence is strictly conserved along the *Flavivirus* genus. In order to answer the question whether the polymerase (and/or the correct template) is at the origin of the conservation of the first nucleotide, we tested a





**Figure 3. pppAG-formation on the correct  $DV_{10}3'$ - and on variant templates having an incorrect last nucleotide.** Dinucleotide formation by NS5Pol<sub>DV</sub> on  $DV_{10}3'$ - templates (ACUAACAA-CN) varying the last nucleotide (correct -CU versus -CC, -CA and -CG). Control reactions were included without template. (A) pppAG dinucleotide formation in the presence of  $Mn^{2+}$ . Only ATP (500  $\mu$ M) and GTP (100  $\mu$ M containing  $\alpha$ GTP) were used as substrates. Reaction mixtures were prepared as given in Materials and Methods plus 2 mM  $MnCl_2$ , 500 nM NS5Pol<sub>DV</sub>, and 1  $\mu$ M template. Reactions were started by the addition of  $MnCl_2$  and samples were taken at given time points. Samples were analyzed by PAGE and autoradiography. pppGG is marked by an asterisk for clarity. (B) pppAG dinucleotide formation in the presence of  $Mg^{2+}$ . Only ATP (500  $\mu$ M) and GTP (100  $\mu$ M containing  $\alpha$ GTP) were used as substrates. Reaction mixtures were prepared as given in Materials and Methods plus 5 mM  $MgCl_2$ , 5  $\mu$ M NS5Pol<sub>DV</sub>, and 1  $\mu$ M template. Reactions were started by addition of  $MgCl_2$  and samples were taken at given time points. Samples were analyzed as under A. (C) pppNG dinucleotide formation in the presence of  $Mn^{2+}$ . All NTPs were given as substrates at equal concentration (100  $\mu$ M). Reaction mixtures were prepared as given in Materials and Methods plus 2 mM  $MnCl_2$ , 500 nM NS5Pol<sub>DV</sub>, and 1  $\mu$ M template. 1  $\mu$ M overall RNA concentration was used when all templates were present. Radiolabelled  $\alpha$ GTP was used. Reactions were started by addition of  $MnCl_2$  and samples were taken at given time points and analyzed as in A.  
 doi:10.1371/journal.ppat.1002912.g003

elongates pppAG even on templates with incorrect 3'-nucleotides thus enabling to repair incorrect 3'-ends.

First, pppAG formation was tested on the four  $DV_{10}3'$ - variants in the presence of only ATP and GTP. Figure 3A shows that NS5Pol<sub>DV</sub> is indeed able to form pppAG in the presence of templates with any 3'-nucleotide and  $Mn^{2+}$ . In contrast, in the presence of  $Mg^{2+}$  only the natural  $DV_{10}3'$ -CU template supports pppAG formation even in the presence of an increased concentration of NS5Pol<sub>DV</sub> (Figure 3B). We then tested pppAG formation exclusively in the presence of  $Mn^{2+}$  on all  $DV_{10}3'$ -variants in the presence of all nucleotides, a scenario putatively mimicking the situation within the replication complex. Figure 3C shows that pppAG is always formed in parallel to the dinucleotide, which corresponds to the template. In the case of the template variant with a -CG 3'-end, pppAG is produced with even higher efficiency than the base-paired dinucleotide. Note that the dinucleotide pppGU is also produced on all templates by internal initiation. For the reaction in the presence of all templates and all nucleotides, we quantified all products, which were initiated *de novo* over the very 3'-end, and found that pppAG is formed as the prominent product ( $32.3 \pm 1.5\%$ , three independent reactions). Note that all templates are present at the same concentration, which should not correspond to the situation *in vivo*. We conclude that in the presence of incorrect templates and  $Mg^{2+}$ , NS5Pol<sub>DV</sub> discriminates against these templates and forms pppAG only on the correct template (see also Figure 2A). In contrast,  $Mn^{2+}$  ions

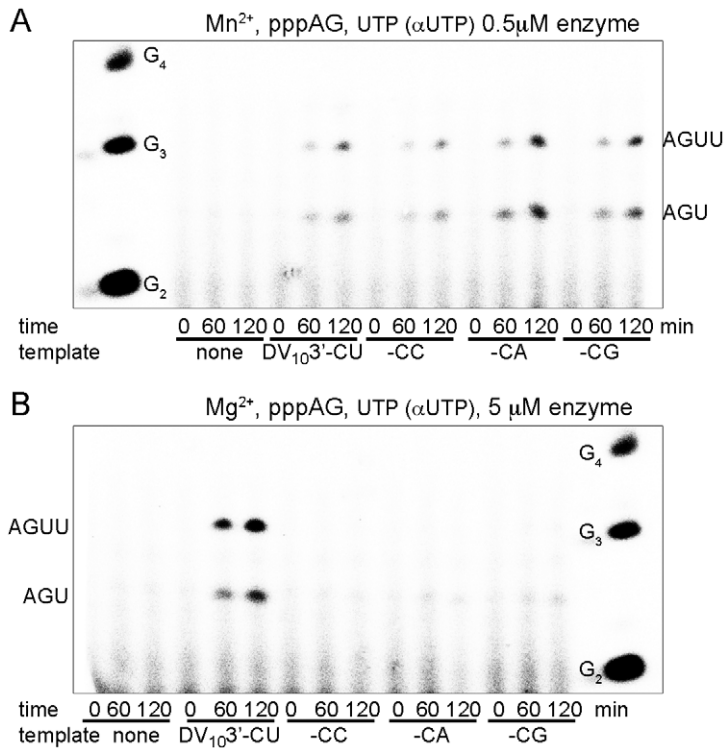
enable NS5Pol<sub>DV</sub> to preferentially generate pppAG even in the presence of incorrect templates, which could represent an indirect way of 3'-end repair.

We then considered the elongation of the correct pppAG primer over templates with incorrect 3'-ends. We thus tested the elongation of a chemically synthesized pppAG primer (see Materials and Methods) either without template or in the presence of the four  $DV_{10}3'$ - variants (Figure 4). The most prominent result is that NS5Pol<sub>DV</sub> is able to productively elongate pppAG on the correct template in the presence of  $Mn^{2+}$  (Figure 4A) and  $Mg^{2+}$  ions (Figure 4B). We also observe that NS5Pol<sub>DV</sub> in the presence of  $Mn^{2+}$  is able to productively elongate pppAG on incorrect templates (Figure 4A), thus demonstrating that the enzyme is able to indirectly correct the error in the template and conserve the 5'-end of the DV genome. Note that as expected there is no primer elongation detectable in the absence of a template.

#### The predicted T794-A799 priming-loop of NS5Pol<sub>DV</sub> provides the built-in ATP-specific priming site

NS5Pol<sub>DV</sub> harbors an ATP-specific priming site, which is essential for the formation, accumulation, and elongation of the correct primer pppAG. Which elements of NS5Pol<sub>DV</sub> form this site? The crystal structure of NS5Pol<sub>DV</sub> (Figure 5A) allowed the prediction of a priming loop comprising residues 792 to 804 [16], which is expected to provide the priming site during *de novo* RNA





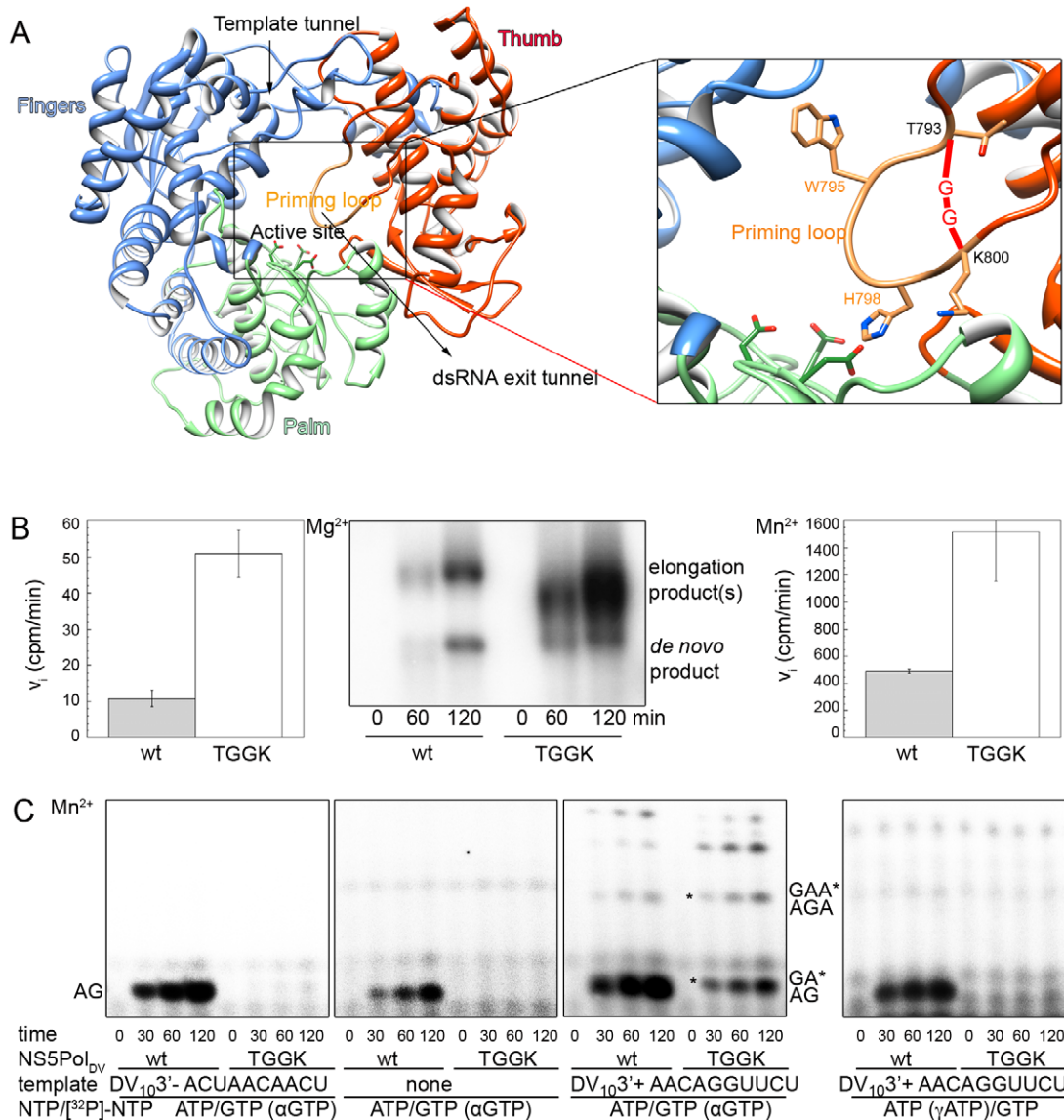
**Figure 4. pppAG-elongation on the correct antigenome 3'-end and on variants with an incorrect last nucleotide.** pppAG-elongation by NS5Pol<sub>DV</sub> on DV<sub>10</sub>3'- templates (ACUAACAA-CN) varying the last nucleotide (correct -CU versus -CC, -CA and -CG). Control reactions were included without template. **(A)** pppAG elongation in the presence of  $Mn^{2+}$ . pppAG (100  $\mu$ M) and UTP (100  $\mu$ M, containing  $\alpha$ UTP) were used as substrates. Reaction mixtures were prepared as given in Materials and Methods plus 2 mM  $MnCl_2$ , 500 nM NS5Pol<sub>DV</sub>, and 1  $\mu$ M template. Reactions were started by addition of  $MnCl_2$  and UTP. Samples were taken at given time points and analyzed by PAGE and autoradiography. OligoG marker is shown on the left, the identity of product bands is given on the right. **(B)** pppAG-elongation in the presence of  $Mg^{2+}$ . pppAG (100  $\mu$ M) and UTP (100  $\mu$ M, containing  $\alpha$ UTP) were used as substrates. Reaction mixtures were prepared as given in Materials and Methods plus 5 mM  $MgCl_2$ , 5  $\mu$ M NS5Pol<sub>DV</sub>, and 1  $\mu$ M template. Reactions were started by addition of  $MgCl_2$  and UTP. Samples were taken at given time points and analyzed by PAGE and autoradiography. OligoG marker is shown on the right; the identity of product bands is given on the left. doi:10.1371/journal.ppat.1002912.g004

synthesis initiation. We generated a deletion mutant (NS5Pol<sub>DV</sub> TGGK) by replacing residues T794-A799 between T793 and K800 by two glycines (see close-up in Figure 5A). The overall correct folding of the purified, recombinant mutant protein was verified by a fluorescent thermal shift assay giving identical temperatures of denaturation (melting temperature  $T_m$ ) for both proteins (wild type (wt) NS5Pol<sub>DV</sub>  $T_m$  49.0°C  $\pm$  0.5°C, NS5Pol<sub>DV</sub> TGGK  $T_m$  48.4°C  $\pm$  0.05°C).

The TGGK mutant is expected to have an open active site, which impedes correct ATP-specific *de novo* initiation over the 3'-end of a single-stranded RNA template but may favor the accommodation of double-stranded RNA. Its RNA synthesis initiation and elongation activity was first tested using a "minigenomic" RNA template consisting of 224 nucleotides of the 5'-end of the DV genome fused to 492 nucleotides of the 3'-end [14]. It has been shown before using this template and analyzing the products on a denaturing agarose-formaldehyde gel [29] that two types of product are formed (see wt reaction kinetics in the center panel of Figure 5B). Firstly, the *de novo* RNA synthesis product is generated corresponding to the size of the template. Secondly, an elongation product is generated by back-primed RNA synthesis. There, the 3'-end (...AACAGGUUCU-3') forms a short hairpin annealing the last di-nucleotide to nucleotides -6 and -7 (underlined in the sequence) and is then elongated [29]. The length of the product is thus ~twice the size of the template. Reactions were carried out using either  $Mg^{2+}$  or  $Mn^{2+}$  as catalytic

ions. The left and right panels of Figure 5B show that in both cases the mutant TGGK shows an increased overall activity on this template compared to wt activity. The center panel shows that this is mainly caused by increased back-priming. Interestingly, instead of one product species of twice the template size NS5Pol<sub>DV</sub> TGGK produces a range of elongated products of different lengths. This might be due to the accommodation of long hairpins, which then create longer products than the template but shorter than the elongation product of wt NS5Pol<sub>DV</sub>.

*De novo* RNA synthesis initiation by wt NS5Pol<sub>DV</sub> and the TGGK mutant were then tested on DV<sub>10</sub>3'-, in the absence of a template and on DV<sub>10</sub>3'+ using  $Mn^{2+}$  as the catalytic ion, ATP and GTP containing  $\alpha$ GTP. Figure 5C (panel 1) shows that in contrast to wt NS5Pol<sub>DV</sub>, NS5Pol<sub>DV</sub> TGGK is not able to catalyze *de novo* initiation on DV<sub>10</sub>3'-. Secondly, NS5Pol<sub>DV</sub> TGGK does not catalyze pppAG formation without template (panel 2). In contrast, it is able to catalyze *de novo* initiation on DV<sub>10</sub>3'+ presenting ca. 32% of wt activity (panel 3). In order to understand this apparent contradiction, we used  $\gamma$ ATP instead of  $\alpha$ GTP as radioactive NTP. It became clear that NS5Pol<sub>DV</sub> TGGK was unable to generate the pppAG primer product (panel 4). We conclude that the product observed with  $\alpha$ GTP corresponds to pppGA formed by internal *de novo* initiation being only possible on DV<sub>10</sub>3'+. When using  $Mg^{2+}$  as catalytic ion again we did not observe formation of the *de novo* RNA synthesis initiation product pppAG on either template (for DV<sub>10</sub>3'- see below Figure 6B).

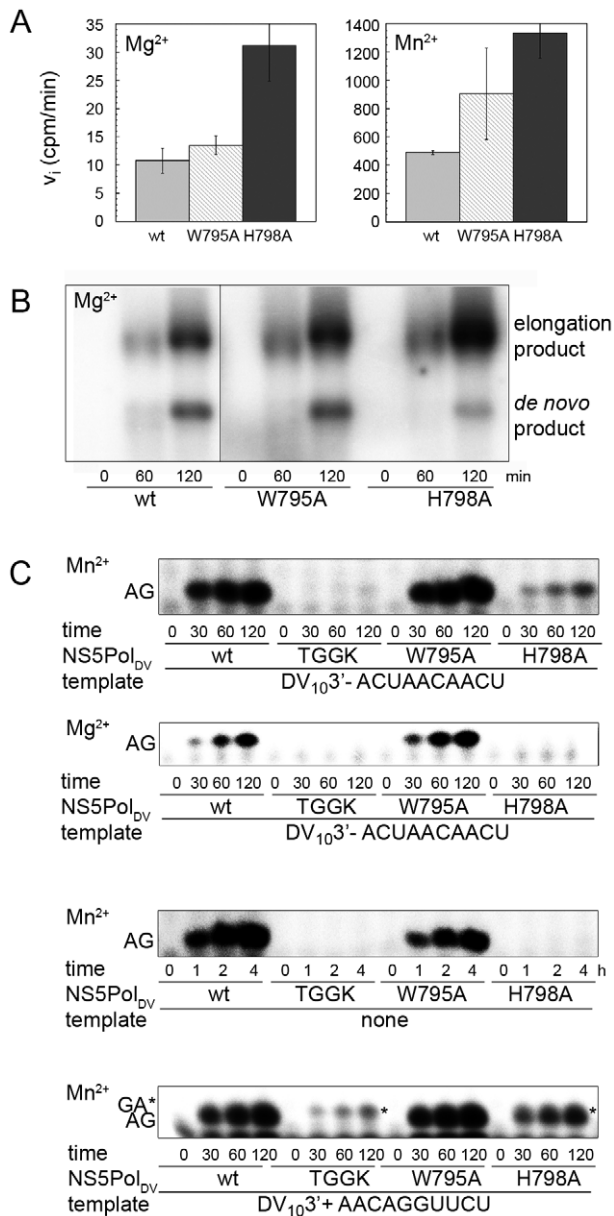


**Figure 5. Role of the predicted priming loop T794-A799 in correct *de novo* initiation.** (A) 3D-structural model of NS5Pol<sub>DV</sub> used in this study (DV serotype 2 strain New Guinea C) derived from the structure of serotype 3 NS5Pol<sub>DV</sub> (PDB code 2J7W [16]). NS5Pol<sub>DV</sub> adopts the typical closed right-hand structure of RdRps containing the palm (light green), fingers (light blue) and thumb (red) subdomains. Between fingers and thumb subdomains the template tunnel runs down to the active site harbored mainly by the palm subdomain. The side chains of the three conserved catalytic residues D533 (motif A), D663 and D664 (motif C) in the active site are shown in sticks (C-atoms light green, O-atoms red). The priming loop emerges from the thumb subdomain and closes the dsRNA exit tunnel and the active site. The close-up shows aromatic residues W795 and H798 (in sticks) within the putative priming loop T794 to A799. In the mutant TGGK the priming loop was replaced by two glycines situated between T793 and K800 (in sticks). (B) Activity of wt NS5Pol<sub>DV</sub> and its deletion mutant TGGK NS5Pol<sub>DV</sub> was determined on a specific minigenomic template. Reaction mixtures were prepared as given in Materials and Methods. Initial velocities in cpm/min determined by filter-binding assays in the presence of [<sup>3</sup>H]-UTP and liquid scintillation counting, are compared in the presence of Mg<sup>2+</sup> (left panel) and Mn<sup>2+</sup> (right panel). The center panel shows agarose-formaldehyde gel analysis of reaction kinetics in the presence of [α-<sup>32</sup>P]-UTP and Mg<sup>2+</sup> ions. Product bands are labeled on the right side of the gel. (C) *De novo* initiation of wt NS5Pol<sub>DV</sub> and its deletion mutant TGGK was followed in the presence of Mn<sup>2+</sup> using either 1 μM DV<sub>10</sub>3<sup>-</sup>, in the absence of a template, or 1 μM DV<sub>10</sub>3<sup>+</sup> (from left to right as indicated). Reaction mixtures also contained 2 mM MnCl<sub>2</sub>, 500 nM enzyme, 500 μM of NTPs, which were not labeled, and 100 μM labeled NTP (containing αGTP or γATP as indicated). Reactions were started by addition of MnCl<sub>2</sub> and samples were taken at given time points. Identities of labeled product bands are given on the right and left side of the reaction kinetics. pppGA and pppGAA internal *de novo* initiation products on DV<sub>10</sub>3<sup>+</sup> are labeled by an asterisk. doi:10.1371/journal.ppat.1002912.g005

We conclude that NS5Pol<sub>DV</sub> TGGK is unable to pre-form the ATP-specific priming site necessary for *de novo* RNA synthesis initiation at the very 3'-end. The predicted priming loop plays

indeed an essential role in providing the correct priming site. We explain the increased activity of NS5Pol<sub>DV</sub> TGGK on minigenomic RNA templates by its increased propensity to catalyze





**Figure 6. Role of NS5Pol<sub>DV</sub> residues His798 and Trp795 as initiation platform.** (A) Activity of wt NS5Pol<sub>DV</sub> and its mutants W795A and H798A was determined on a specific minigenomic template. Reaction mixtures in the presence of [<sup>3</sup>H]-UTP were prepared as given in Materials and Methods. Initial velocities in cpm/min determined using filter-binding assays and liquid scintillation counting are compared in the presence of Mg<sup>2+</sup> (left panel) and Mn<sup>2+</sup> (right panel). (B) Reaction kinetics of wt NS5Pol<sub>DV</sub> and its mutants W795A and H798A on the minigenomic template were analyzed on an agarose-formaldehyde gel. Reaction mixtures in the presence of [<sup>32</sup>P]-UTP and Mg<sup>2+</sup> ions were prepared as given in Materials and Methods. Product bands are labeled on the right side of the gel. (C) *De novo* initiation of wt NS5Pol<sub>DV</sub>, deletion mutant TGGK and mutants W795A and H798A was followed using either DV<sub>10</sub>3' (1 μM in the presence of Mn<sup>2+</sup> and 10 μM in the presence of Mg<sup>2+</sup>), in the absence of a template or 1 μM DV<sub>10</sub>3'+ (from top to bottom as indicated). Reaction mixtures contained 2 mM MnCl<sub>2</sub> or 5 mM MgCl<sub>2</sub> as indicated, 500 nM enzyme, 500 μM of ATP, and 100 μM GTP (containing αGTP). Reactions were started by addition of catalytic ions and samples were taken at given time points. Identities of labeled product bands are given on the left side of the gels. pppGA internal *de novo* initiation product on DV<sub>10</sub>3'+ is labeled by an asterisk.

doi:10.1371/journal.ppat.1002912.g006

back-priming due its more accessible catalytic site, *i.e.* to harbor the minigenome in different hairpin conformations allowing 3' elongation.

### Residue H798 and not W795 is important for ATP-specific *de novo* initiation

Two aromatic residues, W795 and H798, within the priming loop were proposed to play a particular role in providing an initiation platform to which the base of the priming ATP could establish a stacking interaction [16]. Residue W795 was given special attention because it was found near the triphosphate moiety of a 3'-dGTP bound to NS5Pol<sub>DV</sub> [16]. In addition, this tryptophan was better placed than the histidine for stacking a priming ATP in two models of *de novo* RNA synthesis initiation complexes of NS5Pol<sub>DV</sub> and NS5Pol<sub>WNV</sub> [16,22]. We generated two mutants of NS5Pol<sub>DV</sub>, W795A and H798A. Overall correct folding of the purified recombinant mutants was equally verified by a fluorescent thermal shift assay giving T<sub>m</sub> values corresponding to the wt protein (wt NS5Pol<sub>DV</sub> T<sub>m</sub> 49.0°C ± 0.5°C, W795A mutant T<sub>m</sub> 48.6°C ± 0.6°C, H798A mutant T<sub>m</sub> 48.1°C ± 0.04°C).

The RNA initiation and elongation activities of wt NS5Pol<sub>DV</sub> and the W795A and H798A mutants were tested using the minigenomic RNA template and either Mg<sup>2+</sup> or Mn<sup>2+</sup> as catalytic ions (Figure 6A). In both cases the H798A mutant shows an increased activity on this template whereas W795A shows a similar overall activity compared to wt NS5Pol<sub>DV</sub>. Figure 6B shows the analysis of the reaction products on a denaturing agarose-formaldehyde gel. The W795A mutant behaves indeed like wt NS5Pol<sub>DV</sub>; the percentage of the *de novo* RNA synthesis initiation product of template size is unchanged. In contrast the H798A mutant generates considerably less *de novo* RNA synthesis product whereas the yield of RNA elongation products is higher.

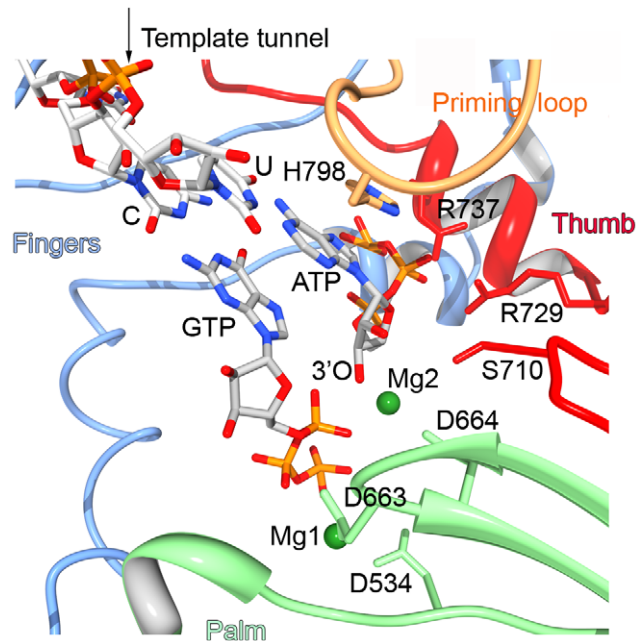
We then compared the capacities of wt and all mutant NS5Pol<sub>DV</sub> proteins to catalyze *de novo* RNA synthesis initiation on DV<sub>10</sub>3', without template and on DV<sub>10</sub>3'+ using Mn<sup>2+</sup> as catalytic ion (Figure 6C panels 1, 3 and 4). Indeed, the H798A mutant is considerably less capable of correct *de novo* RNA synthesis initiation than wt NS5Pol<sub>DV</sub> whereas W795A behaves as wt NS5Pol<sub>DV</sub>. Note that the product formed by NS5Pol<sub>DV</sub> TGGK on DV<sub>10</sub>3'+ (panel 4) corresponds to pppGA generated by internal RNA synthesis initiation (see also Figure 5C); and therefore part of the product formed by the H798A mutant may correspond to pppGA. When Mg<sup>2+</sup> is used on both templates, the same results are obtained (Figure 6C panel 2 for template DV<sub>10</sub>3'). We thus conclude that residue H798 is essential for the formation of the correct ATP-specific priming site and may act as a priming platform.

### Discussion

In this study, we present evidence that the dengue virus NS5 polymerase domain (NS5Pol<sub>DV</sub>) alone is responsible for maintenance of A and U as first and last nucleotides of the DV genome, respectively. NS5Pol<sub>DV</sub> was used instead of full-length NS5 in the frame of this study in order to avoid any interference of the RNA-binding, NTP-binding, or enzymatic activities of the N-terminal domain of NS5. We report that NS5Pol<sub>DV</sub> is endowed with several structural and mechanistic features converging to the specific *de novo* synthesis and elongation of the correct ATP-initiated primer even on templates that lack the correct corresponding U at the 3'-end. The first and last nucleotides of the genome are strictly conserved in the genus *Flavivirus* thus the results presented here may apply to the entire genus.

We demonstrate the generation of a dinucleotide primer pppAG on both genomic and antigenomic RNA templates. We have previously observed the production of such dinucleotide primer on homopolymeric templates [14]. In the following step pppAG(A/U) trinucleotides are formed before processive RNA elongation occurs. During the latter, NS5Pol<sub>DV</sub> continues RNA synthesis to the very end of the template. We do not know if di- and tri-nucleotide primers as detected in the reaction, originate from a slow but processive RNA synthesis reaction, or are actually released from the complex and re-used by the polymerase acting in a distributive RNA synthesis mode. We also show that the pppAG primer is effectively elongated in the presence of Mg<sup>2+</sup> or Mn<sup>2+</sup> and the correct template. Thus, after initial phosphodiester bond synthesis, the pppAG primer is aligned at the correct position in order to be elongated. The efficient use of the short primer pppAG reported here is in apparent contrast to the inefficient use of 5'-OH-AG dinucleotide previously reported [13,30]. The 5'-triphosphate moiety of the chemically synthesized pppAG primer is most probably an important binding determinant allowing efficient elongation (see discussion of the proposed *de novo* initiation complex Figure 7).

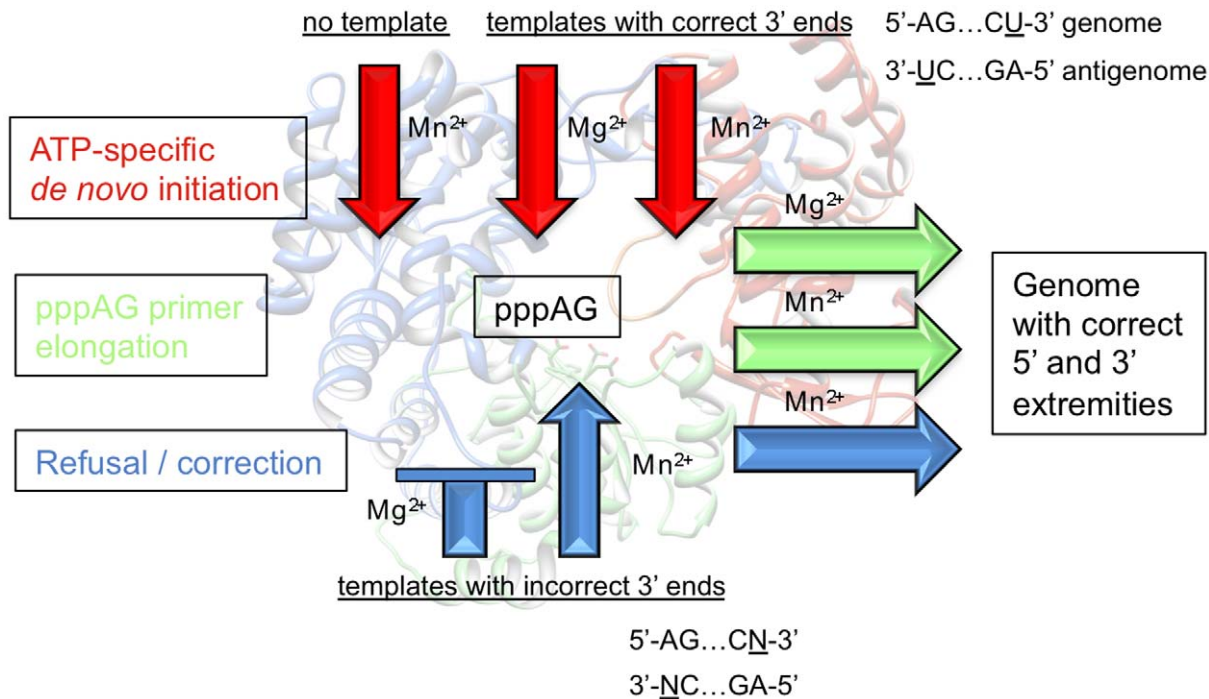
We then demonstrate that in its *de novo* RNA synthesis initiation state NS5Pol<sub>DV</sub> contains a built-in ATP-specific priming site. Major structural elements of NS5Pol<sub>DV</sub> contributing to this site reside within residues T794 to A799. Their deletion forces NS5Pol<sub>DV</sub> to initiate *de novo* RNA synthesis internal to the template using GTP as the first nucleotide (Figure 5C panel 1) and



**Figure 7. *De novo* initiation model of NS5Pol<sub>DV</sub> in complex with RNA template UUCU (3'-end of DV genome), ATP, GTP and catalytic Mg<sup>2+</sup> ions.** The model was generated as explained in Materials and Methods. Fingers, thumb and palm subdomains are colored as in Figure 5A. Only the 3'-end CU of the RNA template is shown, it base pairs the initiating ATP and the second nucleotide GTP. The triphosphate of ATP contacts residues S710, R729 and R737 of the thumb subdomain motif E. The 3'-O atom of the ribose of ATP lies near the  $\alpha$ -phosphate of GTP, which is coordinated to Mg1 bound to D534 of motif A and D663 of motif C. Mg2 is coordinated to the leaving pyrophosphate of GTP. The proposed priming platform H798 stacks to adenine base of the priming ATP.  
doi:10.1371/journal.ppat.1002912.g007

to perform primer-dependent RNA synthesis (Figure 5B). In analogy to the structure of HCV NS5B in complex with a nucleotide in its priming site [31] and because of the amino acid conservation observed within a larger group of *de novo* RdRps [25], we expect that NS5Pol<sub>DV</sub> residues R472 (RdRp catalytic motif F3, see [14]) as well as S710 and R729 (motif E) are involved in triphosphate binding. This might explain why *de novo* RNA synthesis initiation by the loop-deleted mutant is still possible, albeit internal to the template. We conclude that indeed the T794-A799 loop plays a major role both in correct *de novo* initiation and in shaping the priming site. Within the priming loop, residue H798 is essential for primer synthesis (Figure 6). We propose that H798 provides the initiation platform against which the priming nucleotide ATP is stacked. Using the structure of the *de novo* initiation complex of the RdRp of bacteriophage  $\phi 6$  [23] as a starting point, we generated a model of the initiation complex of DV serotype 2 RdRp in complex with the 3'-end of the genome UUCU and both ATP and GTP as first and second nucleotide, respectively (Figure 7). In this model, the triphosphate moiety of ATP indeed interacts with residues S710, R729 and R737 of the thumb subdomain of NS5Pol<sub>DV</sub>. The aromatic ring of H798 stacks the adenine nucleobase of ATP in a similar position to a  $\phi 6$  RdRp tyrosine residue against which the guanine nucleobase of its priming GTP is stacked. In several protein complex structures histidine has been shown to bind an adenine nucleobase by stacking interactions [32]. Nevertheless, histidine does not seem to provide any specificity towards adenine versus guanine [33]. Our model does not propose any obvious specific interaction with the adenine base. This might be due to the fact that the structure of NS5Pol<sub>DV</sub> has been captured in a pre-initiation state. In this state, motif F, which provides the upper part of the NTP entry tunnel in the active initiation and elongation conformation of viral RdRps, is not yet correctly positioned [34]. The fine characterization of the ATP-specific built-in priming site of NS5Pol<sub>DV</sub> awaits the crystal structure of a *de novo* RNA synthesis initiation complex.

We provide a mechanistic basis for the conservation of nucleotides A and U as the first and last nucleotides of the DV genome, respectively. Figure 8 summarizes the different levels of control that ensure ATP-specific *de novo* RNA synthesis initiation. Firstly, it generates and elongates the *bona fide* pppAG primer (red arrows and green arrows on the right). Even in the absence of any template and in the presence of Mn<sup>2+</sup> (Figure 8 left red arrow) NS5Pol<sub>DV</sub> is able to exclusively synthesize the pppAG primer (Figure 2B and C, Figure 3A and C). Note that we have also observed pppAG synthesis by full-length NS5 in the absence of a template (not shown). Since a sufficiently high Mn<sup>2+</sup> concentration is present in the cell (0.1  $\mu$ M to 40  $\mu$ M Mn<sup>2+</sup> in blood, brain, and other tissues [35]), NS5 in the replication complex might already be loaded with pppAG and thus be ready to elongate pppAG on the viral template. The same pppAG primer is preferentially synthesized in the presence of the correct template irrespective of the metal ion present at the polymerase active site (Figure 8 right red arrows, Figure 2A and B, Figure 3). In the presence of Mg<sup>2+</sup>, NS5Pol<sub>DV</sub> supports neither formation nor elongation of pppAG on incorrect templates (Figure 8 blue blocked arrow, Figure 4B). In the presence of Mn<sup>2+</sup>, NS5Pol<sub>DV</sub> is able to synthesize cognate dinucleotides on incorrect templates (Figure 2C), but in the presence of all nucleotides and all templates (a probably biased and more unfavorable set-up compared to the situation in the replication complex *in vivo*), pppAG is still a major product (Figure 3C). Remarkably, the pppAG/Mn<sup>2+</sup>-loaded polymerase is able to mismatch and extend pppAG in order to restore the correct 5'-end (Figure 8 blue arrows, Figure 4). The selective extension reaction thus refrains synthesis of incorrect RNAs that



**Figure 8. Dengue virus RdRp conserves the correct 5'- and 3'-ends of the genome.** DV RdRp conducts strict ATP-specific *de novo* initiation in the absence of a template and in the presence of the correct template using the indicated catalytic ions  $Mn^{2+}$  or  $Mg^{2+}$ . The pppAG primer is then elongated. When DV RdRp encounters templates with incorrect 3'-end nucleotides it refuses *de novo* initiation (when  $Mg^{2+}$  is present) or corrects the error by preferentially generating and elongating pppAG (using  $Mn^{2+}$  as catalytic ion). The structure of the DV2 RdRp domain is shown in the background.

doi:10.1371/journal.ppat.1002912.g008

could occur in the presence of incorrect templates. All these reactions converge to the formation of pppAG and the conservation of A as the starting nucleotide at the 5'-end of viral genomic and antigenomic RNAs. Note that the mechanistic basis of the conservation of the second nucleotide G is beyond the scope of this study. Preliminary results generated in our laboratory indicate that both template and polymerase are important to ensure the specific incorporation of a G as the second nucleotide (not shown).

Several ways of viral RNA genome maintenance and repair concerning terminal damage have been discussed [4], among others the generation of "non-templated" primers and the use of abortive transcripts as primers. Here we demonstrate that NS5Pol<sub>DV</sub> uses these two mechanisms. Non-templated primers are generated only in the presence of  $Mn^{2+}$ . Abortive transcripts are used as primers in the presence of either  $Mg^{2+}$  or  $Mn^{2+}$ . A third mechanism observed here is the discrimination against an incorrect template in the presence of  $Mg^{2+}$ . In addition, in the case that a 3'-end might be shortened, the correction upon *de novo* initiation should be preceded by the addition of (a) nucleotide(s) by the terminal transferase activity of NS5. This activity has also been listed as another way of repairing terminal damage of viral RNA genomes [4]. For NS5Pol<sub>DV</sub> we have observed this activity before [14] and now again in the presence of  $Mn^{2+}$  (Figure 1B).

The DV polymerase endows several of the proposed mechanisms to maintain the correct 5' and 3'-ends of the DV genome and antigenome. The ability of DV and WNV to restore a U at the very 3'-end of genomes with 3'-end deletions has been demonstrated [2,36]. This observation is in accordance with the existence of an ATP-specific priming site in NS5Pol<sub>DV</sub>. Tilgner *et al.* [2,36] reported the complete reversion of WNV replicon CA and CG 3'-ends to CU whereas CC was only partially reverted. Since we have

not seen preferential *de novo* RNA synthesis initiation starting with GG in comparison to UG or CG (all three are possible in presence of  $Mn^{2+}$ , Figure 2), this might be due to an intrinsic difference between DV and WNV RdRp or caused by different propensities of the erroneous templates to allow pppAG elongation. Indeed CA and CG 3'-ends allow pppAG elongation more readily than the CC 3'-end (Figure 4, two independent reactions were performed). Thus the CC 3'-end might therefore take longer to revert. Furthermore, Teramoto *et al.* [2,36] observed the correction of the 5'-end from pppGAG to pppAG. Our work provides a mechanistic explanation for their observation.

The observation of non-templated pppAG formation in the presence of  $Mn^{2+}$  by a viral RdRp has not been reported before using recombinant RdRp assays. However, previous reports convey the occurrence of non-templated dinucleotide formation. RSV, a member of the ns-RNA virus family *Paramyxoviridae* restores the correct 5'-pppA although minireplicons did not encode the correct 3'-U [10]. The authors propose that RSV RdRp contains a built-in ATP-specific priming site and cite the observation that the RdRp of the related ns-RNA vesicular stomatitis virus (VSV, *Rhabdoviridae*) contains a specific ATP-binding site [37] as an argument in favor of their proposition. When VSV RdRp assays were carried out using recombinant RdRp in the presence of  $Mg^{2+}$ , non-templated 5'-initiation was not observed [6]. There is either the possibility that RSV and VSV belong to two different ns-RNA viral families and thus developed different strategies or, in analogy to our results that their RdRps use  $Mn^{2+}$  to correctly initiate RNA synthesis on erroneous templates as observed for NS5Pol<sub>DV</sub> here. It is generally believed that  $Mg^{2+}$  is the activating cofactor of polymerases *in vivo* because viral RdRp properties observed with  $Mg^{2+}$  *in vitro* are more

consistent with properties observed biologically. A second reason for giving the preference to  $Mg^{2+}$  is its cellular abundance in comparison to  $Mn^{2+}$  (*i.e.*, 0.5 mM free  $Mg^{2+}$  versus 0.7  $\mu$ M free  $Mn^{2+}$  in rat hepatocytes [38,39] and 0.1  $\mu$ M–40  $\mu$ M  $Mn^{2+}$  in blood, brain and other tissues [35] versus 0.2 to 0.7 mM  $Mg^{2+}$  in human blood [40]). Nevertheless, some events especially involved in correct and efficient *de novo* RNA synthesis initiation may require the specific use of  $Mn^{2+}$  by viral RdRps under physiological conditions (our study and [10,36,41,42]).

The pppAG primer synthesis by the DV RdRp can be considered as the first line of control of the conservation of *Flavivirus* genome and antigenome ends. However, there might be other mechanisms to tighten the selection. The first one could be the base pairing of the genome ends maintaining specific RNA secondary structures, which are necessary to recruit the replication machinery. Computer simulations of such structures [43] indicate that the last U of the 3'-end of the genome may be unpaired or paired (structure I or II, respectively in [43]). Thus, requested base pairing may exert selective pressure to keep a U at the end of the *Flavivirus* genome. Another selection level concerns only the 5'-end of the genome and is due to the counterselection of incorrect 5'-ends through the NS5 RNA-cap methyltransferase. Indeed, several crystal structures of the cap-dependent bi-functional methyltransferase domain of NS5 show that specific binding of the 5'-cap involves specific recognition of the first transcribed 5'-adenosine through its N1 position and residue Asn18 [44,45]. Therefore, for the genomic strand, methylation at the cap *N*-guanine and the subsequent 2'-*O* position of the first transcribed adenosine should be efficiently achieved only when ATP is the starting 5'-nucleotide. Finally, cap addition seems to involve 5'-ATP selectivity as well [20]. Collectively, we propose that the RdRp of flaviviruses is the first actor responsible for the conservation of the correct ends of their genome, and that other mechanisms such as genome cyclization and the specificity of guanylyltransferase and methyltransferase activities add to the selective pressure. These mechanisms of maintenance might also apply to other RNA virus genera with conserved genome ends and viral RdRps initiating RNA synthesis *de novo*.

## Materials and Methods

### RNA templates

DV<sub>10</sub>3'+ (5'-AACAGGUUCU-3') and DV<sub>10</sub>3'- (5'-ACUAA-CAACU-3') were synthesized at the RIBOXX GmbH Dresden and by Dharmacon. The templates are devoid of stable secondary structure when submitted to the Mfold server [46] ( $\Delta G = 3.60$  kcal/mol for DV<sub>10</sub>3'+ and no folding for DV<sub>10</sub>3'-).

### Large-scale chemical synthesis and purification of pppAG

**Chemical synthesis of AG on solid support.** Chemical synthesis of the diribonucleotide AG was performed on an ABI 394 synthesizer (Applied Biosystems) from commercially available (Link Technologies) long chain alkylamine controlled-pore glass (LCAA-CPG) solid support with a pore size of 1000 Å derivatized through the succinyl linker with 5'-*O*-dimethoxytrityl-2'-*O*-acetyl-*N*<sup>2</sup>-dimethylformamide guanosine. The dinucleotide AG was assembled at a 8- $\mu$ mol scale in Twist oligonucleotide synthesis columns (8×1- $\mu$ mol scale) (Glen research) using 5'-*O*-DMTr-2'-*O*-pivaloyloxymethyl-*N*<sup>6</sup>-phenoxyacetyladenosine]-3'-*O*-(*O*-cyanoethyl-*N,N*-diisopropyl-phosphoramidite (Chemgenes) and following a previously described procedure [47]. After assembly completion, the CPG beads in the eight columns were dried under a stream of argon. The beads were pooled, divided in two, and around 4  $\mu$ mol AG transferred into two Twist oligonucleotide

synthesis columns (size for 10- $\mu$ mol scale). The 5'-functionalization of AG with triphosphate moiety was performed in parallel with both columns following previously described conditions [48].

**5'-triphosphorylation of solid-supported AG.** A solution (8 ml) of 1 M diphenyl phosphite (1.6 ml) in dry pyridine (6.4 ml) was manually passed with a glass syringe through the columns containing AG still attached to the solid support and left to stand for 30 minutes at room temperature. After several washings, the oxidation solution containing imidazole (375 mg, 5 mmol) in *N,O*-bis-trimethylsilylacetamide (1 ml, 4.1 mmol), CH<sub>3</sub>CN (1.875 ml), CCl<sub>4</sub> (1.875 ml) and triethylamine (0.25 ml) was added under argon and left to react for 5 h at 30°C. After washing and drying the support, the TBAPP solution (0.23 M, 2 ml) was applied to the column and left to react for 18 h at 30°C. The solution was removed and the support was washed with dry CH<sub>3</sub>CN (4×8 ml). Finally, the column was dried by 1-min argon flush.

**Deprotection and release of solid-supported pppAG.** A 0.1 M solution of 1,8-diazadicyclo-[5,4,0]undec-7-ene (DBU) (1.2 ml) in anhydrous CH<sub>3</sub>CN (6.8 ml) was applied to each column for 3 min. Subsequently, a 30% aqueous ammonia solution was applied to each column in three batches (6 ml, 4 ml, 2 ml) for 30 min each. The three ammonia fractions were collected in screw-capped glass vials and were left to react at 30°C for 1.5 h. The fully deprotected pppAG was transferred to 50 ml round-bottomed flasks and isopropylamine (15% of total volume) was added to the solutions. Then the mixtures were evaporated under reduced pressure at 30°C until the volumes were reduced to 0.5 ml. The residues were redissolved in water (1.5 ml), transferred to 2 ml Eppendorf-vials and then lyophilized.

**Analysis and purification of pppAG by reverse-phase HPLC.** Analytical and semi-preparative HPLC was performed on a Dionex DX 600 HPLC system equipped with reverse-phase columns (Nucleodur C<sub>18</sub>, 100 Å, 3  $\mu$ m, 4.6×70 mm for analysis and Nucleodur C<sub>18</sub>, 100 Å, 7  $\mu$ m, 8×125 mm for purification, Macherey Nagel). The following solvent system was used: 5% CH<sub>3</sub>CN in 50 mM TEAB buffer, pH 8 (buffer A) and 80% CH<sub>3</sub>CN in 50 mM TEAB buffer, pH 8 (buffer B). Flow rates were 1 ml.min<sup>-1</sup> and 2 ml.min<sup>-1</sup> for analysis and semi-preparative purposes, respectively. Elution was performed with a linear gradient of 0% to 10% buffer B in buffer A in 20 min. The fractions containing the pure pppAG were pooled in a 100 ml round-bottomed flask and were concentrated to a volume of 0.5 ml under reduced pressure at 30°C. The residue was coevaporated ten times with 2 ml of water. The residue was redissolved in 1.5 ml water, transferred to 2 ml Eppendorf-vials and lyophilized. MALDI-TOF mass spectra were recorded on a Voyager-DE spectrometer (Perseptive Biosystems, USA) using a 10:1 (m/m) mixture of 2,4,6-trihydroxyacetophenone/ammonium citrate as a saturated solution in acetonitrile/water (1:1, v/v) for the matrix. Analytical samples were mixed with the matrix in a 1:5 (v/v) ratio, crystallized on a 100-well stainless steel plate and analyzed. UV quantitation of pppAG was performed on a Varian Cary 300 Bio UV/Visible spectrometer by measuring absorbance at 260 nm. Two  $\mu$ mol of pure pppAG were obtained corresponding to 25% total yield. Lyophilized aliquots of 100 nmol have been stored at -20°C for several months without any sign of degradation.

### Protein expression and purification

The gene coding for N-terminal His<sub>6</sub>-tagged NS5Pol<sub>DV</sub> (serotype 2, New Guinea C) as defined in [14] cloned in a pQE30 plasmid was expressed in *E.coli* (Tuner (Novagen) or NEB Express (New England Biolabs)) cells carrying helper plasmid pRare2LacI (Novagen). Expression was carried out in Luria broth

overnight at 17°C after induction with 50  $\mu$ M IPTG, addition of 2% EtOH and a cold shock (2 h at 4°C). Sonication was done in 50 mM sodium phosphate lysis buffer, pH 7.5, 500 mM NaCl, 20% glycerol, 0.8% Igepal (10 ml of this lysis buffer for around 2 g cell pellet from 11 culture) in the presence of DNase I (22  $\mu$ g/ml), 0.2 mM benzamide, protease inhibitor cocktail (SIGMA), 5 mM  $\beta$ -mercaptoethanol and 1 mg/ml lysozyme after 30 min incubation at 4°C. After centrifugation the soluble fraction was incubated in batch with 2 ml TALON metal-affinity resin slurry (Clontech) for 40 min at 4°C. Protein bound to the beads was washed once with 10 volumes of sonication buffer containing 1 M NaCl and 10 mM imidazole and once with the former buffer without Igepal. Protein fractions were then eluted with sonication buffer containing 250 mM imidazole, no Igepal and 250 mM glycine. After dialysis into 10 mM Tris buffer, pH 7.5 containing 300 mM NaCl, 20% glycerol, 250 mM glycine and 1 mM DTT the protein was diluted with the same volume of this buffer without NaCl and loaded onto a HiTrap heparin column (GE Healthcare). Pure NS5Pol<sub>DV</sub> was then eluted in a single peak applying a gradient from 150 mM to 1 M NaCl. Alternatively, gel filtration was used as a second purification step using a Superdex 75 HR 16/60 column (GE Healthcare) and the dialysis buffer. NS5Pol<sub>DV</sub> was stored at -20°C at a concentration of 40 to 60  $\mu$ M after a final extensive dialysis into 10 mM Tris buffer, pH 7.5 containing 300 mM NaCl, 40% glycerol and 1 mM DTT. Purity was higher than 98% as judged by SDS-PAGE.

### Mutant NS5Pol<sub>DV</sub> genes and proteins

Mutant TGGK, W795A and H798A NS5Pol<sub>DV</sub> expression plasmids were generated using the kit QuikChange (Stratagene). Protein expression and purification was done as for the wt protein. Analysis by gel filtration showed a single peak eluting at the same volume as wt NS5Pol<sub>DV</sub>.

### Determination of $T_m$ values

Melting temperature ( $T_m$ ) values of wt and mutant NS5Pol<sub>DV</sub> were determined using a thermofluor-based assay [49]. In 96-well thin-wall PCR plates 3.5  $\mu$ l of a fluorescent dye (Sypro Orange, Molecular Probes, 714-fold diluted in H<sub>2</sub>O) was added to 21.5  $\mu$ l protein solutions at a concentration of 0.5 or 1 mg/ml (6.7 or 13.4  $\mu$ M) in storage buffer. Thermal denaturation of the proteins was followed by measuring fluorescence emission at 575 nm (excitation 490 nm).  $T_m$  values were calculated using GraphPad Prism software and the Boltzmann equation as in [49].

### *In vitro* RdRp assays on DV<sub>103'</sub>+ and DV<sub>103'</sub>-

Reactions were done in 50 mM HEPES buffer, pH 8.0 containing 10 mM KCl, 10 mM DTT and template, NS5Pol<sub>DV</sub>, non-labeled NTPs, and catalytic ions at final concentration as given in the figure legends. Radiolabeled [ $\gamma$ -<sup>32</sup>P]-ATP, [ $\alpha$ -<sup>32</sup>P]-GTP, or [ $\alpha$ -<sup>32</sup>P]-UTP was used at 0.4  $\mu$ Ci per  $\mu$ l reaction volume (3000 Ci/mmol, Perkin-Elmer). Reactions were started by addition of a mixture of HEPES buffer, KCl, catalytic ions and UTP and CTP when used (given in Figures). After given time points samples were taken and reactions stopped by adding an equal volume of formamide/EDTA gel-loading buffer. Reaction products were separated using sequencing gels of 20% acrylamide-bisacrylamide (19:1), 7 M Urea with TTE buffer (89 mM Tris pH 8.0, 28 mM taurine (2-aminoethanesulfonic acid), 0.5 mM EDTA). RNA product bands were visualized using photo-stimulated plates and the Fluorescent Image Analyzer FLA3000 (Fuji) and quantified using Image Gauge (Fuji). The oligoG marker was produced as explained in [14].

### *In vitro* RdRp assays on minigenomic template

The minigenomic template was produced by *in vitro* transcription and tests carried out as described in [14].

Reactions analyzed by filter-binding and liquid scintillation counting contained 50 mM HEPES buffer, pH 8.0, 10 mM KCl, 10 mM DTT, 100 nM RNA template, 200 nM NS5Pol<sub>DV</sub>, 500  $\mu$ M NTP except for UTP (4  $\mu$ M), [<sup>3</sup>H]-UTP at 0.2  $\mu$ Ci/ $\mu$ l and either 5 mM MgCl<sub>2</sub> or 2 mM MnCl<sub>2</sub>. Reactions were started by the addition of a mixture of HEPES, KCl, catalytic ions, CTP, and UTP. After 30, 60, 90, and 120 min 10- $\mu$ l samples were taken and diluted into 50  $\mu$ l of 100 mM EDTA, pH 8.0 to quench the reaction. Samples were then transferred onto a DEAE filter mat. Non-incorporated [<sup>3</sup>H]-UTP was removed by washing with 300 mM ammonium formate and the radioactively labeled product quantified in counts per minute (cpm) using liquid scintillation counting. Product formation was then plotted against time and initial velocities calculated in cpm/min.

Reactions analyzed on formaldehyde-agarose gels contained 50 mM HEPES buffer, pH 8.0, 10 mM KCl, 10 mM DTT, 100 nM RNA template, 200 nM NS5Pol<sub>DV</sub>, 500  $\mu$ M NTP except for UTP (4  $\mu$ M), [ $\alpha$ -<sup>32</sup>P]-UTP at 0.4  $\mu$ Ci/ $\mu$ l, and 5 mM MgCl<sub>2</sub>. Reactions were started by a mixture of HEPES, KCl, MgCl<sub>2</sub>, CTP and UTP and stopped after 60 and 120 min by adding an equal volume of sample buffer (40 mM MOPS pH 7.0, 83.3% formamide, 2 M formaldehyde, 10 mM sodium acetate, 85 mM EDTA). Samples were denatured for 10 min at 70°C and 1/10 of loading buffer (50% glycerol, 10 mM EDTA, xylene cyanol and bromophenol) added. Samples were then analyzed on a 1.2% agarose-formaldehyde gel in 20 mM MOPS buffer pH 7.0, 5 mM sodium acetate, 1 mM EDTA. Gels were dried and RNA product bands visualized using photo-stimulated plates and the Fluorescent Image Analyzer FLA3000 (Fuji) and quantified using Image Gauge (Fuji).

### Modeling of the NS5Pol<sub>DV</sub> initiation complex

A homology model of NS5Pol<sub>DV</sub> serotype 2 strain New Guinea C was generated using the Swiss-model server [50] and the X-ray structure of NS5Pol<sub>DV</sub> serotype 3 (PDB code 2J7W [16]). NS5Pol<sub>DV</sub> and the RdRp of bacteriophage  $\phi$ 6 in complex with a template RNA strand and initiating NTPs (PDB code 1HI0) were then superimposed using the three catalytic aspartate residues of both proteins. The structural model of the initiation complex of NS5Pol<sub>DV</sub> serotype 2 was then generated by changing the RNA template to UUCU (3'-end of the DV genome) and the initiating NTP to ATP, and by manually adapting the conformation of the priming loop using the UCSF Chimera software [51]. Subsequently using the same program the computed free energy of the model was minimized.

### Acknowledgments

We acknowledge contributions of Marilyne Blémond and Armando Vega Osorno in protein production. We are grateful to M.-C. Bergogne for assistance in 5'-triphosphate and capped RNA bibliography search. Finally, we would like to thank Etienne Decroly and François Ferron for reading the manuscript and contributing ideas and critics, and Mats Dreyer for correcting the English.

### Author Contributions

Conceived and designed the experiments: B. Selisko, S. Potisopon, S. Priet, F. Debart, B. Canard. Performed the experiments: B. Selisko, S. Potisopon, R. Agred, S. Priet, I. Varlet, Y. Thillier, C. Sallamand. Analyzed the data: B. Selisko, S. Potisopon, R. Agred, S. Priet, F. Debart, B. Canard. Contributed reagents/materials/analysis tools: Y. Thillier, C. Sallamand, F. Debart, J. J. Vasseur. Wrote the paper: B. Selisko, F. Debart, B. Canard.



## References

- Decroly E, Ferron F, Lescar J, Canard B (2012) Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol* 10:51–65.
- Tilgner M, Shi PY (2004) Structure and function of the 3'-terminal six nucleotides of the west Nile virus genome in viral replication. *J Virol* 78:8159–8171.
- Khromykh AA, Kondratieva N, Sgro JY, Palmenberg A, Westaway EG (2003) Significance in replication of the terminal nucleotides of the flavivirus genome. *J Virol* 77:10623–10629.
- Barr JN, Fearn R (2010) How RNA viruses maintain their genome integrity. *J Gen Virol* 91:1373–1387.
- Farsetta DL, Chandran K, Nibert ML (2000) Transcriptional activities of reovirus RNA polymerase in re-coated cores. Initiation and elongation are regulated by separate mechanisms. *J Biol Chem* 275:39693–39701.
- Morin B, Rahmeh AA, Whelan SP (2012) Mechanism of RNA synthesis initiation by the vesicular stomatitis virus polymerase. *EMBO J* 31:1320–1329.
- Zhong W, Ferrari E, Lesburg CA, Maag D, Ghosh SK, et al. (2000) Template/primer requirements and single nucleotide incorporation by hepatitis C virus nonstructural protein 5B polymerase. *J Virol* 74:9134–9143.
- Nagy PD, Carpenter CD, Simon AE (1997) A novel 3'-end repair mechanism in an RNA virus. *Proc Natl Acad Sci U S A* 94:1113–1118.
- Guan H, Simon AE (2000) Polymerization of nontemplate bases before transcription initiation at the 3'-ends of templates by an RNA-dependent RNA polymerase: an activity involved in 3'-end repair of viral RNAs. *Proc Natl Acad Sci U S A* 97:12451–12456.
- Noton SL, Cowton VM, Zack CR, McGivern DR, Fearn R (2010) Evidence that the polymerase of respiratory syncytial virus initiates RNA replication in a nontemplated fashion. *Proc Natl Acad Sci U S A* 107:10226–10231.
- Lindenbach BD, Thielt H-J, Rice C (2007) *Flaviviridae*: the viruses and their replication. In: Knipe DM, Howley PM, editors. *Fields Virology*. 5th ed. Lippincott Williams & Wilkins. pp. 1101–1152.
- Gubler D, Kuno G, Markoff L (2007) *Flaviviruses*. In: Knipe DM HP, editor. *Fields Virology*. 5th ed. Lippincott Williams & Wilkins. pp. 1155–1252.
- Nomaguchi M, Ackermann M, Yon C, You S, Padmanabhan R (2003) *De novo* synthesis of negative-strand RNA by Dengue virus RNA-dependent RNA polymerase in vitro: nucleotide, primer, and template parameters. *J Virol* 77:8831–8842.
- Selisko B, Dutartre H, Guillemot JC, Debarnot C, Benarroch D, et al. (2006) Comparative mechanistic studies of *de novo* RNA synthesis by flavivirus RNA-dependent RNA polymerases. *Virology* 351:145–158.
- Filomatori CV, Lodeiro MF, Alvarez DE, Samsa MM, Pietrasanta L, et al. (2006) A 5' RNA element promotes dengue virus RNA synthesis on a circular genome. *Genes Dev* 20:2238–2249.
- Yap TL, Xu T, Chen YL, Malet H, Eglhoff MP, et al. (2007) Crystal structure of the dengue virus RNA-dependent RNA polymerase catalytic domain at 1.85-angstrom resolution. *J Virol* 81:4753–4765.
- Latour DR, Jekle A, Javanbakht H, Henningsen R, Gee P, et al. (2010) Biochemical characterization of the inhibition of the dengue virus RNA polymerase by beta-d-2'-ethynyl-7-deaza-adenosine triphosphate. *Antiviral Res* 87:213–222.
- Eglhoff MP, Benarroch D, Selisko B, Romette JL, Canard B (2002) An RNA cap (nucleoside-2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. *EMBO J* 21:2757–2768.
- Ray D, Shah A, Tilgner M, Guo Y, Zhao Y, et al. (2006) West Nile virus 5'-cap structure is formed by sequential guanine N-7 and ribose 2'-O methylations by nonstructural protein 5. *J Virol* 80:8362–8370.
- Issur M, Geiss BJ, Bougie I, Picard-Jean F, Despines S, et al. (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.
- Bollati M, Alvarez K, Assenberg R, Baronti C, Canard B, et al. (2010) Structure and functionality in flavivirus NS-proteins: Perspectives for drug design. *Antiviral Res* 87: 125–148
- Malet H, Eglhoff MP, Selisko B, Butcher RE, Wright PJ, et al. (2007) Crystal structure of the RNA polymerase domain of the West Nile virus non-structural protein 5. *J Biol Chem* 282:10678–10689.
- Butcher SJ, Grimes JM, Makeyev EV, Bamford DH, Stuart DI (2001) A mechanism for initiating RNA-dependent RNA polymerization. *Nature* 410:235–240.
- Tao Y, Farsetta DL, Nibert ML, Harrison SC (2002) RNA synthesis in a cage-structural studies of reovirus polymerase lambda3. *Cell* 111:733–745.
- Bressanelli S, Tomei L, Rey FA, De Francesco R (2002) Structural analysis of the hepatitis C virus RNA polymerase in complex with ribonucleotides. *J Virol* 76:3482–3492.
- Choi KH, Groarke JM, Young DC, Kuhn RJ, Smith JL, et al. (2004) The structure of the RNA-dependent RNA polymerase from bovine viral diarrhoea virus establishes the role of GTP in *de novo* initiation. *Proc Natl Acad Sci U S A* 101:4425–4430.
- Ranjith-Kumar CT, Gutshall L, Kim MJ, Sarisky RT, Kao CC (2002) Requirements for *de novo* initiation of RNA synthesis by recombinant flavivirus RNA-dependent RNA polymerases. *J Virol* 76:12526–12536.
- Ranjith-Kumar CT, Gajewski J, Gutshall L, Maley D, Sarisky RT, et al. (2001) Terminal nucleotidyl transferase activity of recombinant Flaviviridae RNA-dependent RNA polymerases: implication for viral RNA synthesis. *J Virol* 75:8615–8623.
- Ackermann M, Padmanabhan R (2001) *De novo* synthesis of RNA by the dengue virus RNA-dependent RNA polymerase exhibits temperature dependence at the initiation but not elongation phase. *J Biol Chem* 276:39926–39937.
- Iglesias NG, Filomatori CV, Gamarnik AV (2011) The F1 motif of dengue virus polymerase NS5 is involved in promoter-dependent RNA synthesis. *J Virol* 85:5745–5756.
- Harris D, Ahmed-El-Sayed N, Simister PC, Miller S, Triconnet M, et al. (2010) Further insights into the roles of GTP and the C terminus of the hepatitis C virus polymerase in the initiation of RNA synthesis. *J Biol Chem* 285:32906–32918.
- Cauët E, Rooman M, Wintjens R, Lievin J, Biot C (2005) Histidine-Aromatic Interactions in Proteins and Protein-Ligand Complexes: Quantum Chemical Study of X-ray and Model Structures. *J Chem Theory Comput* 1:472–483.
- Firoz A, Malik A, Joplin KH, Ahmad Z, Jha V, et al. (2011) Residue propensities, discrimination and binding site prediction of adenine and guanine phosphates. *BMC Biochem* 12:20.
- Malet H, Masse N, Selisko B, Romette JL, Alvarez K, et al. (2008) The flavivirus polymerase as a target for drug discovery. *Antiviral Res* 80:23–35.
- Versieck J (1985) Trace elements in human body fluids and tissues. *Crit Rev Clin Lab Sci* 22:97–184.
- Teramoto T, Kohno Y, Mattoo P, Markoff L, Falgout B, et al. (2008) Genome 3'-end repair in dengue virus type 2. *RNA* 14:2645–2656.
- Massey DM, Lenard J (1987) Inactivation of the RNA polymerase of vesicular stomatitis virus by N-ethylmaleimide and protection by nucleoside triphosphates. Evidence for a second ATP binding site on L protein. *J Biol Chem* 262:8734–8737.
- Ash DE, Schramm VL (1982) Determination of free and bound manganese(II) in hepatocytes from fed and fasted rats. *J Biol Chem* 257:9261–9264.
- Gaussin V, Gailly P, Gillis JM, Hue L (1997) Fructose-induced increase in intracellular free Mg<sup>2+</sup> ion concentration in rat hepatocytes: relation with the enzymes of glycogen metabolism. *Biochem J* 326:823–827.
- Wang S, McDonnell EH, Sedor FA, Toffaletti JG (2002) pH effects on measurements of ionized calcium and ionized magnesium in blood. *Arch Pathol Lab Med* 126:947–950.
- Wright S, Poranen MM, Bamford DH, Stuart DI, Grimes JM (2012) Noncatalytic ions direct the RNA-dependent RNA polymerase of bacterial double-stranded RNA virus varphi6 from *de novo* initiation to elongation. *J Virol* 86:2837–2849.
- Ranjith-Kumar CT, Kim YC, Gutshall L, Silverman C, Khandekar S, et al. (2002) Mechanism of *de novo* initiation by the hepatitis C virus RNA-dependent RNA polymerase: role of divalent metals. *J Virol* 76:12513–12525.
- Alvarez DE, Lodeiro MF, Luduena SJ, Pietrasanta LI, Gamarnik AV (2005) Long-range RNA-RNA interactions circularize the dengue virus genome. *J Virol* 79:6631–6643.
- Eglhoff MP, Decroly E, Malet H, Selisko B, Benarroch D, et al. (2007) Structural and functional analysis of methylation and 5'-RNA sequence requirements of short capped RNAs by the methyltransferase domain of dengue virus NS5. *J Mol Biol* 372:723–736.
- Yap IJ, Luo D, Chung KY, Lim SP, Bodenreider C, et al. (2010) Crystal structure of the dengue virus methyltransferase bound to a 5'-capped octameric RNA. *PLoS One* 5:e12836.
- Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31:3406–3415.
- Lavergne T, Bertrand JR, Vasseur JJ, Debart F (2008) A base-labile group for 2'-OH protection of ribonucleosides: A major challenge for RNA synthesis. *Chemistry* 14:9135–9138.
- Zlatev I, Lavergne T, Debart F, Vasseur JJ, Manoharan M, et al. (2010) Efficient solid-phase chemical synthesis of 5'-triphosphates of DNA, RNA, and their analogues. *Org Lett* 12:2190–2193.
- Ericsson UB, Hallberg BM, Detitta GT, Dekker N, Nordlund P (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem* 357:289–298.
- Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18:2714–2723.
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.