Biochemistry and function of nidovirus replicase proteins

Lehmann, K.C.

CHAPTER 4

Arterivirus RNA-dependent RNA polymerase: vital enzymatic activity remains elusive

Kathleen C. Lehmann
Alexander E. Gorbalenya
Eric J. Snijder
and Clara C. Posthuma

ABSTRACT

Polynucleotide polymerases are the central enzymes involved in nucleic acid-based functions of all organisms and viruses. Reflecting this importance, a detailed understanding of their activities is crucial for deciphering biologically important processes like genome replication, transcription, and repair. All plus-stranded RNA viruses encode a conserved RNA-dependent RNA polymerase (RdRp), which was extensively characterized only in viruses of few families. In the order *Nidovirales,* which includes viruses with (very) large genomes, the RdRp is expressed in association with other replicative enzymes as part of the polyprotein encoded in open reading frame 1b (ORF1b). Based on sequence conservation, it was mapped to the C-terminal domain of nonstructural protein (nsp) 9 in arteriviruses and nsp12 in coronaviruses, the two families of mammalian nidoviruses. Potent primer-dependent RdRp activity was demonstrated for the severe acute respiratory syndrome coronavirus enzyme. In contrast, the only study focusing on nsp9 of the arterivirus equine arteritis virus (EAV) reported *de novo* polymerase activity on certain homopolymeric RNA templates in biochemical assays. However, this activity was not maintained when Mn^{2+} ions, which are known to relieve the sequence dependency of polymerases, were omitted or when biologically more relevant templates representing viral sequences were supplied. Due to these observations, we sought to revisit the biochemical properties of this polymerase. We describe here the results of a carefully controlled study involving several preparations of purified recombinant EAV nsp9 that included the wild-type and a set of active site mutants, which were tested for *de novo* and primer-depended polymerase and terminal transferase activities. However, we were unable to reproduce the published EAV nsp9 activity as the RdRp domain of nsp9 was found not to be associated with any of the activities observed in these assays. Also we noticed a striking resemblance between the product profiles of one of the tested preparations of nsp9 and that of T7 phage RNA polymerase. Our results hence emphasize the need to employ diverse controls when utilizing highly sensitive biochemical assays.

INTRODUCTION

Polymerases, which catalyze the templated synthesis of polynucleotides in the 5′-3′ direction, are enzymes encoded by all organisms and RNA viruses, as well as some DNA viruses. Reflecting the principal differentiation into DNA- and RNA-based processes and functions, those enzymes can be grouped into four classes each possessing a distinct combination of specificities for their substrate (NTPs or dNTPs) and template (RNA or DNA) under physiological conditions. Despite these fundamental differences regarding the requirements for their substrates, many polymerases of the four classes, including all characterized RNA-dependent RNA polymerases (RdRps), employ the same catalytic mechanism and a similar three-dimensional fold resembling the shape of a right hand with finger, thumb, and palm domains (1;2). At the sequence level, these polymerases share two sequence motifs, motifs A and C, found in the most conserved palm domain (3). Few conserved residues, primarily aspartates, located in these motifs are implicated in (d)NTP binding and/or catalysis (4;5), and consequently their replacement should abolish or at least severely decrease nucleic acid synthesis (2).

Based on their requirements for initiation of nucleic acid synthesis, two types of polymerases are recognized: primer-dependent and *de novo*-initiating enzymes (2;6). The latter, to our knowledge exclusively RNA polymerases (DNA- or RNA-dependent), are capable of positioning two NTPs, typically two purines, in a manner that allows the formation of a starting dinucleotide. In contrast, primer-dependent polymerases are unable to accommodate the required stable association between the first (d)NTP and the template. As a result the formation of the first dinucleotide is an energetically extremely unfavorable event in these proteins. To overcome this problem, short RNA primers must be produced and placed on the template. For this purpose, organisms and viruses have evolved different initiation mechanisms that are all assisted by additional proteins or domains. They may involve the synthesis of short RNA fragments (by e.g., eukaryotic DNA primase (7)), the formation of covalent RNA-protein complexes (e.g., picornavirus VPg-RNA complexes (8)), or the utilization of tRNAs (by lentivirus tRNA-binding domains (9)) or 5′ fragments of cellular mRNAs (generated by influenza virus, bunyavirus, and arenavirus endoribonuclease and cap-binding domains (10-12)).

For genome replication many viruses rely on a polymerase that is encoded within their genome. In viruses of the order *Nidovirales* (comprising the families *Arteriviridae*, *Coronaviridae*, *Mesoniviridae*, and *Roniviridae*), which are characterized by their large to exceptionally large single-stranded RNA genomes (13;14), a canonical RdRp possessing common motifs of other polymerases with right-hand structure is expressed from ORF1b as part of the pp1ab replicase polyprotein (15-17). After proteolytic cleavage, a protein

subunit (nonstructural protein (nsp) 9 in *Arteriviridae*, nsp12 in *Coronaviridae*) harboring conserved motifs of an RdRp in its C-terminal two-thirds is released (14;15;18;19). Eventually, this cleavage product becomes a key subunit of the membrane-associated multi-subunit replication-transcription complex (RTC) that mediates the synthesis of diverse viral RNAs (20-22). This complex has been characterized *in situ* and through reconstitution of its activities *in vitro*. In one of these studies coronaviruses, prototyped by severe acute respiratory syndrome coronavirus (SARS-CoV), were proposed to express a second, non-canonical RNA polymerase subunit: the ORF1a-encoded nsp8 (23). In agreement with early studies describing nsp8 as an obligatory *de novo* polymerase capable of synthesizing products of less than six nucleotides (23) and nsp12 as strictly primer-dependent (24), it was speculated that the two proteins may work sequentially on the same template, with nsp8 providing the primers required by the nsp12 "main RdRp". Subsequently, also recombinant feline coronavirus nsp8 and human coronavirus 229E nsp7-10 (an nsp8-containing precursor) were reported to be able to synthesize RNA oligonucleotides with a length of up to six nucleotides. Upon addition of the cognate nsp7, the activity of feline coronavirus nsp8 was further enhanced, generating RNA products of up to 67 nucleotides (25).

However, recent studies question this clear division of labor. First, it was shown that recombinant nsp8 expressed without any artificial terminal residues also possesses primer-dependent activity. Furthermore, in complex with its co-factor nsp7, this activity was estimated to be only 2.5-fold lower in terms of NTP incorporation per active site than that of nsp12 (26). Additionally, one study also reported *de novo* activity for nsp12 (27). Finally, in the most recent study, SARS-CoV nsp12 showed non-processive primer extension activity in an *in vitro* assay, which was substantially enhanced by the addition of nsp7 and nsp8. The same combination of three proteins was also required for *de novo* initiation of RNA synthesis. A complex of just nsp7 and nsp8, on the other hand, did not show any activity in this study. Hence it was concluded that the nsp7-nsp8 complex serves as an activator and processivity factor, rather than primase, for the nsp12 RdRp (16). The background of the reported differences and apparent contradictions with respect to the properties of SARS-CoV nsp8 (in complex with nsp7) and nsp12 remain unknown, but technical differences are likely to play a role, especially concerning the expression constructs, protein purification, and templates used.

Besides SARS-CoV nsp12, RdRp activity was characterized for only one other nidovirus "main RdRp", the arterivirus equine arteritis virus (EAV) nsp9 (28). In that study, *de novo* RdRp activity was reported on poly-uridine (pU) and poly-cytidine (pC) single-stranded RNAs while no primer extension or terminal transferase activity, that is, the untemplated elongation of RNA strands, was detected. Thus, it was concluded that EAV nsp9 activity

is restricted to *de novo* initiation. However, the applicability of the observed activity to virus replication remained uncertain since activity on templates containing appropriate virus-specific sequences could not be detected, and the *in vitro* activity required the presence of Mn^{2+}, which is known to relieve template requirements for other polymerases (29). One possible explanation for the lack of initiation on virus-specific templates could be that additional co-factors, e.g. higher-order RNA structures or proteins, are needed for genuine *de novo* initiation *in vivo*. Therefore, the aim of this study was to characterize the RNA polymerase activity of EAV nsp9 in more detail. We report the results of a carefully controlled study involving several preparations of purified recombinant EAV nsp9 that included the wild-type protein and a set of active-site mutants, which were tested for *de novo* and primer-dependent polymerase and terminal transferase activities. However, we were unable to reproduce the published EAV nsp9 activity as the RdRp domain of nsp9 was found not to be associated with any of the activities observed in these assays. Also we noticed a striking resemblance between the product profiles of one of the tested preparations of nsp9 and that of T7 phage RNA polymerase. Our results hence emphasize the need to employ diverse controls when utilizing highly sensitive biochemical assays.

RESULTS AND DISCUSSION

Expression and purification of EAV nsp9 using two vectors

Previously, the purification and *de novo* polymerase activity of recombinant EAV nsp9 were described (28). In that study the viral protein (subsequently designated as nsp9/pDEST) was cloned into a pDEST vector including a C-terminal hexahistidine tag and expressed in *E. coli* BL21 (DE3). As typical for bacterially expressed proteins, an unknown fraction of nsp9 may contain an N-terminal formylmethionine due to saturation of the endogenous protein processing pathway by nsp9 overexpression. Such an N-terminal extension would modify the authentic N-terminus of nsp9, which is expected to be a glycine residue following the proteolytic release of nsp9 from the pp1ab polyprotein by nsp4-mediated cleavage of the Glu1677↓Gly1678 site (30). Previously, it was reported for SARS-CoV nsp8 and nsp12 that artificial tags at the N-terminus may influence RdRp activity and stability, respectively (24;26).

To circumvent this potential problem, we decided to express EAV nsp9 as part of a ubiquitin fusion protein by using a so-called pASK vector (31), the resulting protein is hereafter referred to as nsp9/pASK. In combination with co-expression of the ubiquitin-specific protease UBP1, which will remove the N-terminal ubiquitin fusion partner *in bacterio*,

this enabled us to obtain the natural glycine N-terminus of nsp9 when expressed in the *E. coli* BL21 derived strain C2523/pCG1. An additional advantage of the pASK vector was that its backbone allowed us to drive expression via the endogenous pool of *E. coli* RNA polymerase after induction with anhydrotetracycline. In contrast, nsp9/pDEST was expressed from a T7 promoter after over-expression of the T7 phage RNA polymerase. Although this expression system is well characterized and has proven suitable for a wide range of proteins, the potential presence of this phage RNA polymerase in the ultimate nsp9 preparations could be of concern. Indeed, since the demonstrated activity of recombinant EAV nsp9 was shown to be low (28), even trace quantities of this potent phage polymerase might cause a significant background activity complicating the interpretation of the obtained results.

Both variants of recombinant nsp9 were expressed in their respective *E. coli* strains under identical growth conditions. They were subsequently batch purified in a single step using metal ion chromatography with Co^{2+} targeting the C-terminal hexahistidine tag of both polypeptides. As Figure 1A shows, both proteins could be obtained with similar purity, but nsp9/pASK was expressed in higher quantities than nsp9/pDEST. Attempts to further purify both proteins by gel filtration did not result in a significant improvement as judged by silver staining of SDS-PAGE gels (not shown).

T7 RNA polymerase contamination may account for *de novo* activity observed with EAV nsp9 preparations

nsp9/pDEST and nsp9/pASK preparations were tested side-by-side in a *de novo* polymerase assay in the presence of radioactive ATP using similar reaction conditions as described before (28). The only noteworthy difference from the published protocol was the length of the pU template, which was 30 nucleotides in our experiments compared to an undefined mixture containing RNAs of up to 300 nucleotides in the study of Beerens *et al.*. To our surprise neither of the preparations showed any activity on this template even when the ATP concentration was 15-fold increased to 1.5 mM with the goal to favor polymerase initiation (not shown). Next we tested the RdRp activity using a template whose 3'-terminal dinucleotide matched the CC dinucleotide that is present immediately upstream of the poly(A) tail at the 3' end of the EAV genome. Indeed, as previously shown for homopolymeric pC templates, nsp9/pDEST exhibited some activity with this RNA template, while nsp9/pASK remained essentially inactive (Figure 1B, middle and left panel, respectively, lanes R_1). As noted earlier, the former and latter preparations differed in two respects: the presence of an artificial N-terminal residue in nsp9/pDEST and the induction of T7 RNA polymerase production to achieve expression of nsp9/pDEST. Only this expression of an additional polymerase can reasonably be linked to the (gain of) activity in the nsp9/pDEST preparation.

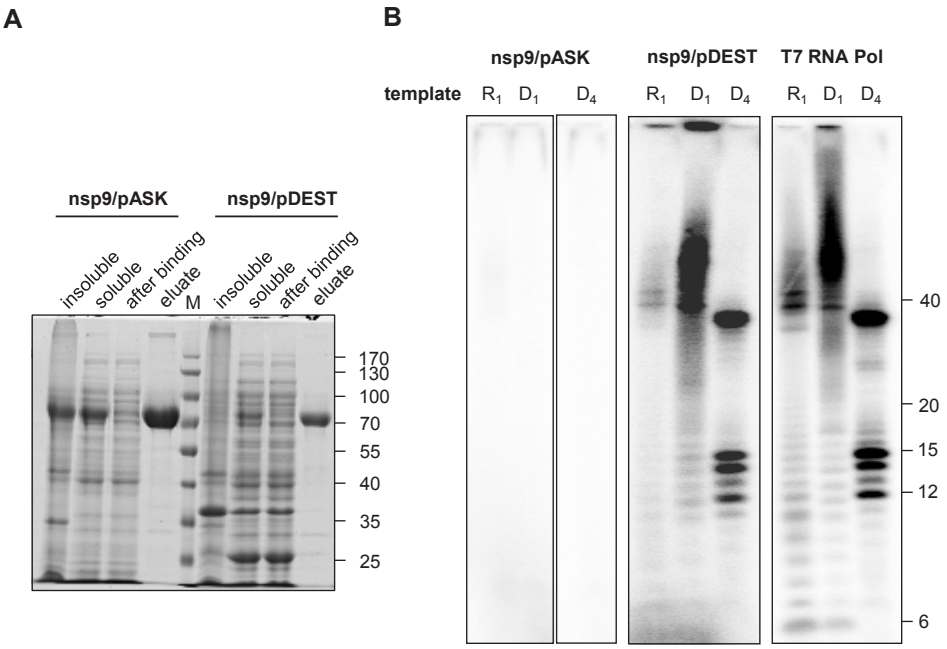

Figure 1. Expression, purification, and *de novo* polymerase activity of two recombinant EAV nsp9-His preparations. (**A**) Coomassie brilliant blue-stained SDS-PAGE gel of samples taken during metal ion chromatography using Co^{2+}. Insoluble and soluble: respective fractions after cell lysis; after binding: unbound protein after removal of Co^{2+} resin; eluate: elution fraction after purification. The molecular weight of nsp9-His is 78 kDa. Size markers are depicted on the right in kDa. (**B**) *de novo* polymerase assay using nsp9 expressed from pASK (final protein concentration 2 µM) or pDest (final protein concentration 0.6 µM) vectors, or using commercial T7 RNA polymerase (0.05 U per sample). R and D indicate the use of RNA and DNA templates, respectively. Identical numbers indicate templates with equivalent sequences. Template sequences are listed in Table 2. Product lengths (nt) are indicated on the right. Note that products longer than template length, 30 nt for R_1 and D_1, result from terminal transferase activity acting on either the template or the newly synthesized strand. D_4 template length 45 nt.

To test this hypothesis, we extended our analysis to include also a highly diluted sample (0.01 U/µl final concentration) of a commercially available T7 RNA polymerase. Since this enzyme is DNA-dependent, we included two single-stranded DNA templates: a DNA variant (D_1) of the RNA template used and a DNA template containing the negative-stranded T7 promoter sequence (D_4). The rationale for the second template was to provide a specific recognition signal in the template for the enzyme and thus increase the chances to observe its activity. Although it has been reported that this DNA-dependent enzyme strictly requires its cognate promoter in a double-stranded form (32), we reasoned that providing DNA with the same polarity as the template that is transcribed under physiological conditions, may at least support some residual activity. Surprisingly, T7 RNA polymerase was active on all of these templates under the employed conditions (Figure 1B, right panel) with the expected preference for DNA templates.

Interestingly, the product pattern from the T7 promoter-containing template was markedly different from the one expected. As already mentioned, *de novo* initiation on any given template can be forced by increasing the concentration of the required NTPs. Likewise, decreasing the concentration of one of the NTPs will force a polymerase to pause and eventually dissociate from the template (or incorporate a non-matching nucleotide) once it encounters the complementary base. In this manner synthesis by enzymes with low processivity can be shifted from the production of evenly distributed but low-intensity products towards a few predominant, high-intensity bands. Thus, limiting the concentration of one nucleotide, in this case CTP, may increase the probability of detecting polymerase activity if the signal-to-noise ratio is a concern. However, this was not evident with nsp9/pDEST (Figure 1B, right panel). While the lack of these prominent bands in lanes R_1 and D_1 may be explained by misincorporation of nucleotides, favored by the high ratio between correct and incorrect NTPs as well as the presence of error-inducing Mn^{2+}, the preference for synthesizing the products of a length of 12, 14, 15, and approximately 38 nucleotides seen in lane D_4 is difficult to reconcile with the template's sequence. Instead it would be expected that, if at all, synthesis would terminate at positions preceding a G residue in the template (nucleotides 5 (in which case the product would not be visible), 8, 10, 12, etc.) as incorporation of CTP is unfavorable under the conditions applied. A possible explanation for the observed product pattern could be internal initiation on this template lacking a strong promoter sequence. Thus, it is tempting to speculate that this particular template interacts with T7 RNA polymerase in a distinctive manner that may not be shared by other polymerases.

With that said, it remains to be noted that the nsp9/pDEST preparation showed the same overall pattern, including the preference for DNA templates, as the commercial T7 RNA polymerase. In line with this notion, an nsp9/pASK preparation gained *de novo* activity once it was expressed in BL21 (DE3) under addition of IPTG (not shown). Hence, this circumstantial evidence suggests that contaminating T7 RNA polymerase, rather than EAV nsp9 itself, is responsible for the *de novo* polymerase activity observed here.

Whether or not this contamination was also present in the nsp9 preparations described in Beerens *et al.* (28), and later on also by te Velthuis *et al.* (33), cannot be established with certainty as the experiments presented here and those published previously deviated in some aspects. Particularly the previously described purification protocol could not be reproduced in our experiments due to technical difficulties with the described purification buffer, which in our hands induced protein precipitation during purification. Furthermore, as mentioned above, we could not observe an additional purifying effect of a second chromatography step. Nevertheless, we investigated whether inclusion of a gel filtration step with a low flow rate (0.3 ml/min) would remove the suspected trace

contamination with T7 RNA polymerase (molecular weight 99 kDa) from a preparation of nsp9/pDEST (molecular weight 78 kDa). We found that this was not the case (not shown).

In conclusion, our results revealed that the radioactive polymerase assay used in this and previous studies is sensitive enough to detect trace activities of contaminating T7 RNA polymerase and also enables this polymerase to act on templates lacking the established T7 promoter requirements. Still, the fact that we did not detect any RdRp activity for nsp9/pDEST, and therefore also not for its suspected contaminant, on a pU template may be used to argue for the detection of genuine nsp9 activity in the previous studies. In this context it is noteworthy that the two coronavirus RdRps were addressed in six independent studies (16;23-27), none of which succeeded in exactly reproducing results of any other. This may indicate that nidovirus RdRps are highly delicate proteins responding to minute changes during purification or in their reaction environment.

EAV nsp9/pASK preparations possess primer-dependent polymerase and terminal transferase activity

Besides *de novo* activity, we decided to test whether EAV nsp9 may possess primer-dependent polymerase activity like its larger coronavirus homolog nsp12 (16;24). To detect this activity, we used a similar assay as the one described above but this time providing partially double-stranded templates. We found that both nsp9 preparations were enzymatically active on these templates and showed the highest extension activity if the template and primer were RNAs (Figure 2A, left and middle panel). This differential reaction towards the type of substrate showed that the measured activity was a direct response to the added nucleic acids, hence not to a co-purified *E. coli*-derived RNA or DNA template. Furthermore, as the presence of a DNA template significantly decreased processivity, it also demonstrated that the responsible polymerase was RNA dependent. Interestingly, while the use of a DNA primer in combination with an RNA template precluded any extension (no products in the size range between 20 and 39 nucleotides), a product corresponding to a length of 40 nucleotides was detected. This suggested that the polymerase possesses terminal transferase activity but only on RNA substrates. To investigate this further, we also compared the elongation of single-stranded RNA and DNA substrates in an assay otherwise identical to the one used for measuring primer-dependent polymerase activity (Figure 2B). As expected, both nsp9 preparations showed a clear selectivity in favor of RNA, again emphasizing their dependence on this substrate type. In this context it is also noteworthy that neither the primer extension nor the terminal transferase assay included Mn^{2+} ions, which can favor activity on sub-optimal templates (29). Together with the demonstrated DNA specificity of T7 RNA polymerase (Figures 2A and B, right panels) this supports the reliability of these assays with respect to the reproduction of physiologically relevant substrate preferences.

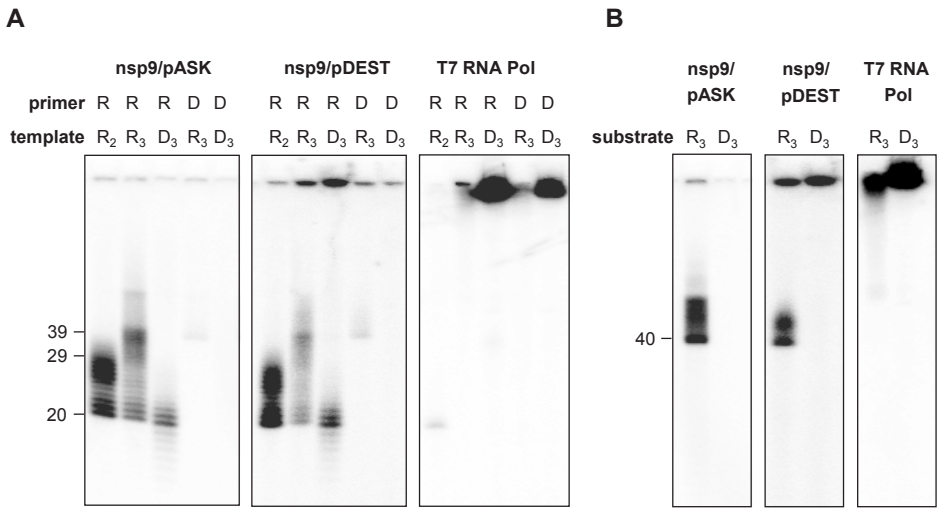

Figure 2. Polymerase assays using recombinant EAV nsp9-His expressed from pASK (final protein concentration 1 μM) or pDEST (final protein concentration 0.3 μM) vectors or using commercial T7 RNA polymerase (0.025 U per sample). R and D indicate RNA and DNA strands, respectively. Identical numbers indicate nucleic acids with equivalent sequences. Nucleic acid sequences are listed in Table 2. Product lengths (nt) are indicated on the left. (**A**) Results from primer extension assay (primer length 19 nt). Note that products longer than template length (29 nt for R_2, 39 nt for R_3 and D_3) must have resulted from terminal transferase activity acting on either the template or the newly synthesized strand. (**B**) Results from terminal transferase assay. The signal at the very top of the gel likely represents products of >200nt that cannot be resolved in the high-percentage acrylamide gel used here. Note that products resulting from end-labeling with ATP may be further extended by a back-priming mechanism.

Finally, to conclude the characterization of the polymerase, its nucleotide preference was examined. To this end, a primed RNA template (Figure 3A) was first elongated in the presence of a low concentration of radioactive ATP, resulting in frequent abortion of transcription after incorporation of the first nucleotide. Subsequently, either dATP or ATP was supplied in a concentration that should allow restarting and completion of the reaction (Figure 3B). As expected, addition of ATP enabled the synthesis of almost fully extended products while dATP did not support any extension beyond one or two nucleotides (Figure 3C). In agreement with the lack of DNA primer extension and the known inability of the prototype viral RdRp of poliovirus to further extend deoxynucleotide chains (4;34), we thus conclude that the observed activity originated from an RNA-dependent RNA polymerase.

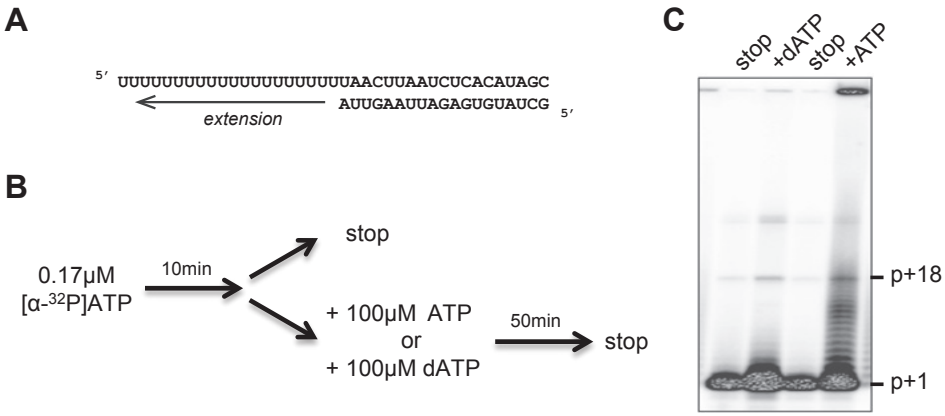

Figure 3. Stop-and-go primer extension assay using EAV nsp9-His expressed from a pASK vector and primer/template R/R_3 in the presence of ATP or dATP. (**A**) Sequence of primer/template. (**B**) Schematic representation of the experimental design. (**C**) Polymerase products after interrupted and resumed synthesis. The sizes of primer extension products are indicated on the right.

Reverse genetics of conserved aspartates of nsp9

To establish whether the observed activity was associated with the RdRp domain of EAV nsp9, we substituted several of the key residues of the (predicted) active site of the enzyme with alanine. Previously, EAV nsp9 residues belonging to conserved polymerase motifs had been identified (17;28). To further support this identification, we constructed EAV full-length cDNA clones encoding alanine substitution mutants of each of the four conserved aspartates of motifs A and C, which coordinate the essential metal ions or interact with the NTP's 2′ and 3′ hydroxyl groups in the better characterized polymerases. After *in vitro* transcription, full-length RNAs representing these mutants were transfected into BHK-21 cells, which were monitored for viral progeny production and protein expression using plaque assays and immunofluorescence microscopy, respectively. Polymerase activity is primarily based on a two-metal-ion mechanism involving several residues. In contrast to other catalysis mechanisms, which may feature a single or few absolutely required residues, individual amino acids rather work in concert during metal catalysis to provide a framework for metal ions and substrates to bind. Consequently, the substitution of single residues may merely reduce binding affinities and may thus, depending on their individual contribution, be either lethal or non-lethal for the enzyme's function and thus for the virus. In agreement with the expected essential role and preliminary unpublished observations for equivalent SARS-CoV nsp12 mutants (not shown), each of the aspartate-to-alanine substitutions tested had a severe impact on viral replication. Whereas all double mutations tested were lethal, viruses carrying single mutations apparently retained a low level of RNA synthesis, ultimately leading to rever-

sion to wild-type virus later in the experiment (by 48 h p.t.; Table 1). In all cases a single nucleotide point mutation was sufficient to restore the codon for the wild-type residue. Nevertheless, this finding is somewhat unexpected given the universal conservation of all four aspartates in positive-stranded RNA viruses. To our knowledge replication, even though severely decreased and undetectable until reversion had occurred, of a single mutant of the enzyme's active site has not been reported for any other RNA virus thus far.

Table 1. Summary of reverse genetics data of EAV nsp9 mutants.

	IFA				plaque phenotype (48 h p.t.)	titer (48 h p.t.) (PFU/ml)	nsp9 sequence of P1*
	16 h p.t.	24 h p.t.	48 h p.t.	72 h p.t.			
wt	++	+++	cells dead	cells dead	large	$2{\cdot}10^7$	n.d.
D445A	-	-	+	+++	large	$6{\cdot}10^5$	wt
D450A	-	-	++	+++	large	$2{\cdot}10^5$	wt
D445/450A	-	-	-	-	-	-	n.d.
D560A	-	-	++	+++	large	$6{\cdot}10^7$	wt
D559/560A	-	-	-	-	-	-	n.d.

IFAs were done with antibodies directed against nsp3 and N proteins; -, negative; +, few, separated positive; ++, clustered positive; +++, all positive; p.t., post transfection; n.d., not done; *P1 was generated by infection of fresh BHK-21 cells with supernatant harvested at 72 h p.t.

Observed primer extension and terminal transferase activities are not correlated with EAV nsp9

Following the results described above, we transferred the same mutations into the nsp9/pASK expression construct to obtain negative controls for the biochemical RdRp assays described in the previous paragraphs. However, none of the proteins with double, triple, and quadruple aspartate-to-alanine substitutions tested showed a decreased primer extension activity compared to two independently purified batches of wild-type recombinant nsp9/pASK (Figure 4). Likewise, D445A and D560A mutant proteins maintained terminal transferase activity (not shown). Thus, the observed activities either derived from a second active site within nsp9, which was not targeted by mutagenesis, or may have originated from a different (contaminating) protein altogether. Both these explanations are quite extraordinary since none of the described RdRps is known to have a second active site and no RdRp activity from *E. coli* has been reported to the best of our knowledge.

To discriminate between these possibilities, we asked whether it was possible to separate nsp9-containing fractions from biochemically active ones during purification

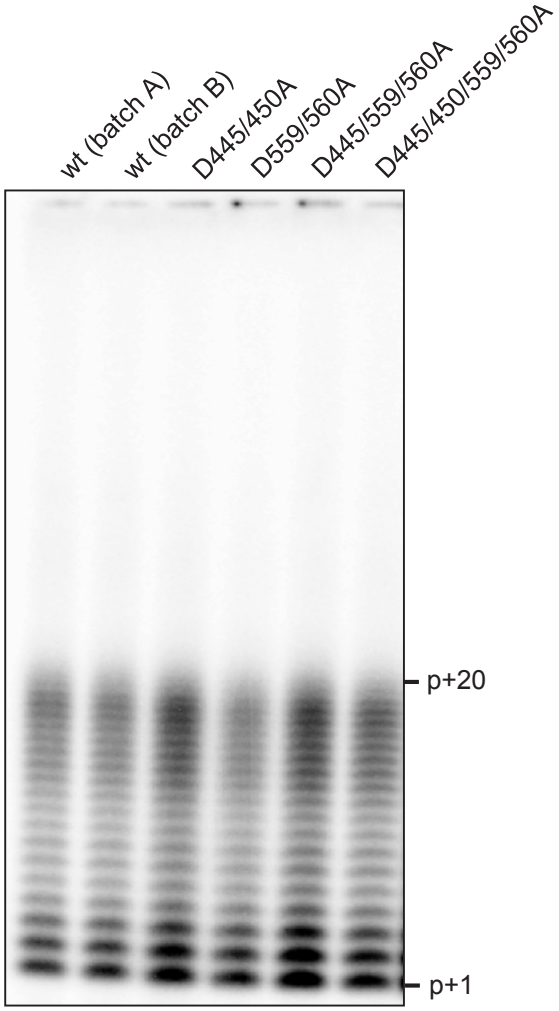

Figure 4. Primer extension assay on primer/template R/R_3 using wild-type EAV nsp9-His expressed from a pASK vector and mutants in which essential aspartate residues of the RdRp domain were replaced with alanine (D445 and D450 of motif A, D559 and D560 of motif C). The sizes of primer extension products are indicated on the right.

of the quadruple mutant of nsp9/pASK. To this end, the wash steps of the previously established purification protocol were modified in either of two ways; first, a decreasing salt gradient was introduced to weaken (disrupt) hydrophobic interactions between a contaminant and nsp9, and second, an increasing imidazole gradient was employed in order to eliminate any contaminant from the Co^{2+}-resin. As shown in Figure 5, the NaCl elution fraction and wash steps 2 and 3 of the imidazole gradient contained almost identical amounts of nsp9-His, as judged by SDS-PAGE, while two of these three fractions were inactive in the polymerase assay. This partial correlation between the presence of

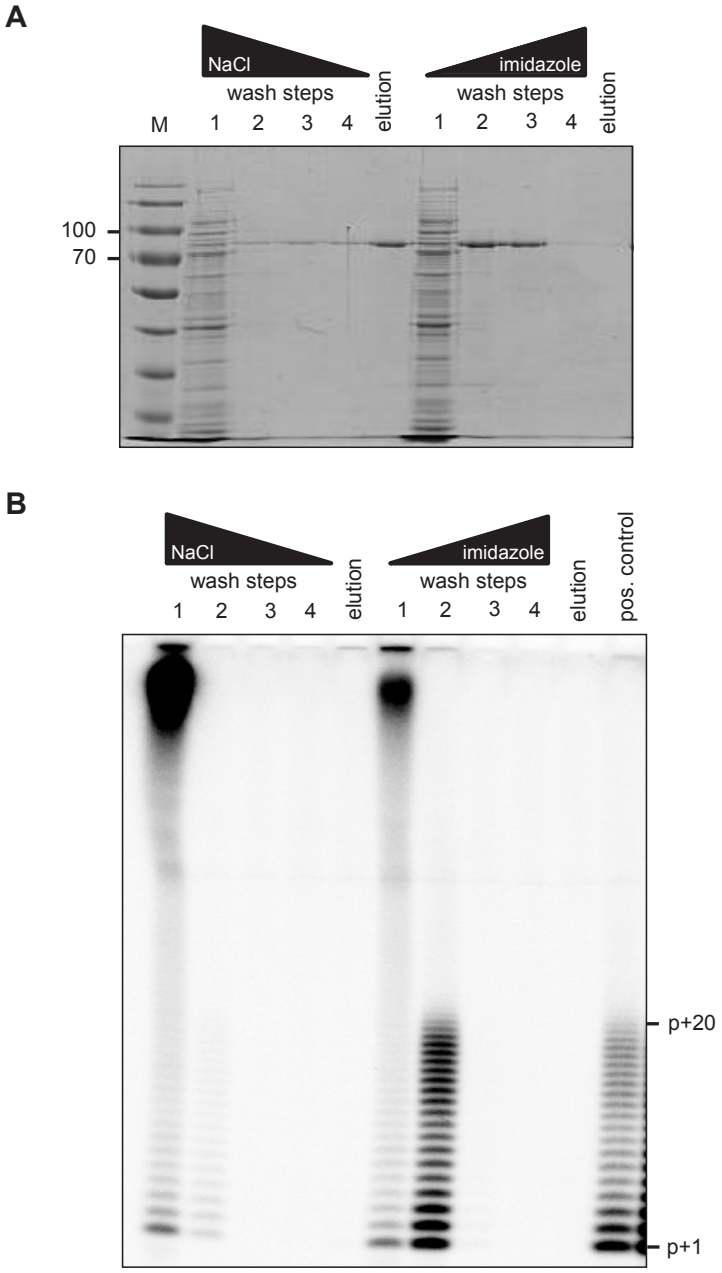

Figure 5. Correlation between EAV nsp9-containing fractions and primer extension activity. (**A**) Coomassie brilliant blue-stained SDS-PAGE gel of samples taken during the purification of nsp9/pASK by Co^{2+} affinity chromatography using wash buffers with either a decreasing NaCl concentration or an increasing imidazole concentration. Size markers are depicted on the left in kDa. (**B**) The samples shown in A were examined for primer extension activity on primer/template R/R_3. The sizes of primer extension products are indicated on the right.

recombinant nsp9-His and primer extension activity could be due to either the presence of two forms of nsp9, enzymatically active and defective, or the presence of a second enzyme responsible for the activity. Resolving the remaining uncertainty is challenging since bacteria are, to our knowledge, not known to encode RNA-dependent RNA polymerases, and the nature and origin of the possible heterogeneity of nsp9 remained elusive.

Concluding remarks

In a final effort to activate the polymerase activity of recombinant EAV nsp9, we included several potential protein co-factors in our primer extension assays. For these experiments we chose the poorly characterized arterivirus subunits encoded immediately upstream of the ORF1a/1b ribosomal frameshift site, at the genomic position equivalent to those of the proven coronavirus nsp12-RdRp co-factors nsp7 and nsp8 (16). Although the proteins of the distantly related corona- and arterivirus families share little similarity, they might have diverged beyond recognition while retaining similar functions. These poorly characterized arterivirus proteins include nsp6, a 22-amino acid peptide in EAV, nsp7α (123 amino acids in EAV), nsp7ß (102 amino acids in EAV), and nsp8 (50 amino acids in EAV), the subunit located immediately upstream of the frameshift site that corresponds to the N-terminus of nsp9. Additionally, as these four subunits are known to be contained in, in part, long lasting cleavage intermediates (nsp6-7α, nsp6-7, nsp6-7-8, nsp7, nsp7-8) (35;36) also those were tested. Finally, the EAV helicase nsp10 was included since its SARS-CoV homolog (nsp13) was shown to interact with its cognate RdRp nsp12 (37;38). Unfortunately, neither of these subunits had any positive impact on the polymerase activity of recombinant EAV nsp9/pASK (not shown) or showed any evidence of interaction with nsp9 in native gel and cross-linking experiments (not shown). However, we should note that we did not probe this possibility exhaustively using different experimental conditions to facilitate complex formation or maybe even co-expression of multiple partners. Hence, there is certainly room to explore the co-factor hypothesis in more detail.

To conclude, in this study we could neither confirm the previously reported *de novo* polymerase activity nor detect any other RNA polymerase activity originating from purified recombinant EAV nsp9-His, indicating that the characterization of the arterivirus RdRp presents a formidable challenge. While the reason(s) underlying the differences to earlier studies remains to be elucidated, the outcome of the present study emphasizes the need for selecting proper controls especially when utilizing highly sensitive biochemical assays for characterizing enzymes with low activity. Furthermore, it demonstrates that *in vitro* assays may reveal activities that are not biologically relevant under physiological

conditions and/or in the presence of interaction partners that may alter substrate preferences by modifying an enzyme's conformation. Being aware of this pitfall probably is one of the most fundamental prerequisites for the deduction of biological roles from biochemical assays.

MATERIAL AND METHODS

Protein expression and purification

C-terminally His-tagged fusion proteins of wild-type and mutant EAV nsp9 were expressed under the control of a tetracycline promoter from a pASK vector in the *E. coli* BL21 derivative C2523/pCG1 as described (26). As a reference, a previously used pDEST construct of nsp9-His_6 was expressed in *E. coli* BL21 (DE3) cells after IPTG induction under otherwise identical conditions. Proteins were purified by metal affinity chromatography using Co^{2+} (Talon beads) as described (26) using a buffer containing 20 mM HEPES, pH 7.5, 500 mM NaCl, 10% glycerol (v/v), 10 mM imidazole, and 5 mM β-mercaptoethanol unless it is explicitly stated otherwise. Where indicated, a second purification step using a Superdex 200 10/300 GL gel filtration column with 20 mM HEPES, pH 7.5, 300 mM NaCl, 1 mM DTT was performed at 4°C using a flow rate of 0.3 ml/min.

Polymerase assays

Three different types of polymerase assays were performed: *de novo*, primer extension, and terminal transferase assays. For *de novo* assays samples contained 10 mM Tris, pH 8.0, 5 mM KCl, 25 mM NaCl (including 20 mM from the protein storage buffer), 6 mM $MgCl_2$, 1.5 mM $MnCl_2$, 1.5 mM DTT, 12.5% glycerol (including 10% from the protein storage buffer), 0.005% Triton X-100, 1.5 U RiboLock RNase inhibitor (Thermo Scientific), 0.5 µM single-stranded nucleic acid template, 1.5 mM ATP, if required 0.7 mM GTP and 0.7 mM UTP, 0.17 µM [α-^{32}P]CTP (Perkin Elmer, 3000 Ci/mmol), and 2 µM nsp9/pASK or 0.6 µM nsp9/pDEST or 0.05 U T7 RNA polymerase from a commercial source (Life Technologies). Primer extension and terminal transferase assays were performed in 20 mM Tris, pH 8.0, 10 mM KCl, 20 mM NaCl (including 10 mM from the protein storage buffer), 6 mM $MgCl_2$, 1 mM DTT, 10% glycerol (including 5% from the protein storage buffer), 0.01% Triton X-100, 0.5 U RiboLock RNase inhibitor, 1 µM partially double-stranded (primer extension) or single-stranded (terminal transferase) nucleic acid, 50 µM ATP, 0.17 µM [α-^{32}P]ATP (Perkin Elmer, 3000 Ci/mmol), and 1 µM nsp9/pASK or 0.3 µM nsp9/pDEST or 0.025 U T7 RNA polymerase (Life Technologies). Sequences of used nucleic acids are listed in Table 2.

Table 2. Sequences of nucleic acids used for polymerase assays.

Primers	
R	GCUAUGUGAGAUUAAGUUA
D	GCTATGTGAGATTAAGTTA
templates/substrates	
R_1	UUUUUUUUUUGCCUCGCUGCCGUCGCCACC
R_2^*	UUUUUUUUUUUAACUUAAUCUCACAUAGC
R_3^*	UUUUUUUUUUUUUUUUUUUUUUAACUUAAUCUCACAUAGC
D_1	TTTTTTTTTTGCCTCGCTGCCGTCGCCACC
D_3^*	TTTTTTTTTTTTTTTTTTTTTAACTTAATCTCACATAGC
D_4	GCTATGTGAGATTAAGTTATCTGAGCCCTATAGTGAGTCGTATTA

*Sequences complementary to both primers are underlined

Nucleic acids were annealed with complementary primers by heating to 95°C for 2 min, then keeping them at 52°C for 30 min, and finally letting them cool to room temperature in 30 min.

In all three assays, samples were incubated for 1 h at 30°C before the reaction was stopped by addition of an equal volume of formamide gel loading buffer (95% formamide, 18 mM EDTA, 0.025% SDS, xylene cyanol, bromophenol blue) and 2 min denaturing at 95°C. Products were separated by gel electrophoresis in 20% polyacrylamide gels (19:1) containing 7 M urea. Gels were run in 0.5x TBE and subsequently exposed to phosphorimager screens overnight. Screens were scanned on a Typhoon variable mode scanner (GE Healthcare), and band intensities were analyzed with ImageQuant TL software (GE Healthcare).

Reverse genetics of EAV

Alanine-encoding mutations of codons specifying conserved nsp9 residues were generated using the QuikChange protocol and were introduced into full-length cDNA clone pEAV211 (39) using appropriate shuttle vectors and restriction enzymes. The presence of the mutations was confirmed by sequencing. pEAV211 plasmid DNA was *in vitro* transcribed and full-length RNA was transfected into BHK-21 cells as described previously (40). Transfected cells were monitored until 72 h post transfection (p.t.) by immunofluorescence microscopy using antibodies directed against the nsp3 and N proteins as described (41). To monitor the production of viral progeny, supernatants were harvested at 48 h p.t. and plaque assays were performed as described (40). To verify the presence of the introduced mutations or reversions in viable mutants, fresh BHK-21 cells were

infected with supernatants harvested at 72 h p.t., RNA was isolated with TriPure after 18 h, and the nsp9-coding region was amplified by RT-PCR and sequenced.

ACKNOWLEDGEMENTS

This work was supported by the European Union Seventh Framework program through the EUVIRNA project (European Training Network on (+) RNA virus replication and Antiviral Drug Development, grant agreement no. 264286). AEG acknowledges support from the Leiden University Fund and MoBiLe Program. The authors thank Jessika Zevenhoven-Dobbe and Aartjan te Velthuis for technical assistance and helpful discussions.

REFERENCE LIST

1. Iyer LM, Koonin EV, Aravind L. Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. BMC.Struct.Biol. 2003;31.

2. Ng KK, Arnold JJ, Cameron CE. Structure-function relationships among RNA-dependent RNA polymerases. Curr.Top.Microbiol.Immunol. 2008;320:137-156.

3. Delarue M, Poch O, Tordo N, *et al.* An attempt to unify the structure of polymerases. Protein Eng 1990; 3(6):461-467.

4. Arnold JJ, Cameron CE. Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mg^{2+}. Biochemistry 2004;43(18):5126-5137.

5. Arnold JJ, Vignuzzi M, Stone JK, *et al.* Remote site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. J.Biol.Chem. 2005;280(27):25706-25716.

6. van Dijk AA, Makeyev EV, Bamford DH. Initiation of viral RNA-dependent RNA polymerization. J.Gen. Virol. 2004;85(Pt 5):1077-1093.

7. Pellegrini L. The Pol alpha-primase complex. Subcell.Biochem. 2012;62:157-169.

8. Steil BP, Barton DJ. Cis-active RNA elements (CREs) and picornavirus RNA replication. Virus Res. 2009; 139(2):240-252.

9. Kleiman L. tRNA(Lys3): the primer tRNA for reverse transcription in HIV-1. IUBMB.Life 2002;53(2):107-114.

10. Boivin S, Cusack S, Ruigrok RW, *et al.* Influenza A virus polymerase: structural insights into replication and host adaptation mechanisms. J.Biol.Chem. 2010;285(37):28411-28417.

11. Morin B, Coutard B, Lelke M, *et al.* The N-terminal domain of the arenavirus L protein is an RNA endonuclease essential in mRNA transcription. PLoS.Pathog. 2010;6(9):e1001038.

12. Reguera J, Weber F, Cusack S. Bunyaviridae RNA polymerases (L-protein) have an N-terminal, influenza-like endonuclease domain, essential for viral cap-dependent transcription. PLoS.Pathog. 2010;6(9): e1001101.

13. de Groot RJ, Baker SC, Baric R, *et al.* Family *Coronaviridae*. In King AMQ, Adams MJ, Carstens EB *et al.* editors, Virus taxonomy. Ninth report of the international committee on taxonomy of viruses, Amsterdam, Elsevier Academic Press, 2012;806-828.

14. Lauber C, Ziebuhr J, Junglen S, *et al. Mesoniviridae*: a proposed new family in the order *Nidovirales* formed by a single species of mosquito-borne viruses. Arch.Virol. 2012;157(8):1623-1628.

15. Gorbalenya AE, Koonin EV, Donchenko AP, *et al.* Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. Nucleic Acids Res. 1989;17(12):4847-4861.

16. Subissi L, Posthuma CC, Collet A, *et al.* One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc.Natl.Acad.Sci.U.S.A 2014;111(37):e3900

17. Den Boon JA, Snijder EJ, Chirnside ED, *et al.* Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. J.Virol. 1991;65(6):2910-2920.

18. van Dinten LC, Wassenaar AL, Gorbalenya AE, *et al.* Processing of the equine arteritis virus replicase ORF1b protein: identification of cleavage products containing the putative viral polymerase and helicase domains. J.Virol. 1996;70(10):6625-6633.

19. Ziebuhr J, Siddell SG. Processing of the human coronavirus 229E replicase polyproteins by the virus-encoded 3C-like proteinase: identification of proteolytic products and cleavage sites common to pp1a and pp1ab. J.Virol. 1999;73(1):177-185.

20. Sawicki SG, Sawicki DL, Siddell SG. A contemporary view of coronavirus transcription. J.Virol. 2007; 81(1):20-29.

21. van Hemert MJ, van den Worm SH, Knoops K, *et al.* SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity *in vitro*. PLoS.Pathog. 2008;4(5): e1000054.

22. van Hemert MJ, de Wilde AH, Gorbalenya AE, *et al.* The *in vitro* RNA synthesizing activity of the isolated arterivirus replication/transcription complex is dependent on a host factor. J.Biol.Chem. 2008;283(24): 16525-16536.

23. Imbert I, Guillemot JC, Bourhis JM, *et al.* A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. EMBO J. 2006;25(20):4933-4942.

24. te Velthuis AJ, Arnold JJ, Cameron CE, *et al.* The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. Nucleic Acids Res. 2010;38(1):203-214.

25. Xiao Y, Ma Q, Restle T, *et al.* Nonstructural proteins 7 and 8 of feline coronavirus form a 2:1 heterotrimer that exhibits primer-independent RNA polymerase activity. J.Virol. 2012;86(8):4444-4454.

26. te Velthuis AJ, van den Worm SH, Snijder EJ. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. Nucleic Acids Res. 2012;40(4):1737-1747.

27. Ahn DG, Choi JK, Taylor DR, *et al.* Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates. Arch.Virol. 2012; 157(11):2095-2104.

28. Beerens N, Selisko B, Ricagno S, *et al.* De novo initiation of RNA synthesis by the arterivirus RNA-dependent RNA polymerase. J.Virol. 2007;81(16):8384-8395.

29. Arnold JJ, Ghosh SK, Cameron CE. Poliovirus RNA-dependent RNA polymerase (3D(pol)). Divalent cation modulation of primer, template, and nucleotide selection. J.Biol.Chem. 1999;274(52):37060-37069.

30. Snijder EJ, Wassenaar AL, van Dinten LC, *et al.* The arterivirus nsp4 protease is the prototype of a novel group of chymotrypsin-like enzymes, the 3C-like serine proteases. J.Biol.Chem. 1996;271(9):4864-4871.

31. Gohara DW, Ha CS, Kumar S, *et al.* Production of "authentic" poliovirus RNA-dependent RNA polymerase (3D(pol)) by ubiquitin-protease-mediated cleavage in *Escherichia coli*. Protein Expr.Purif. 1999; 17(1):128-138.

32. Rong M, He B, McAllister WT, *et al.* Promoter specificity determinants of T7 RNA polymerase. Proc.Natl. Acad.Sci.U.S.A 1998;95(2):515-519.

33. te Velthuis AJ, van den Worm SH, Sims AC, *et al.* Zn(2+) inhibits coronavirus and arterivirus RNA polymerase activity *in vitro* and zinc ionophores block the replication of these viruses in cell culture. PLoS. Pathog. 2010;6(11):e1001176.

34. Gohara DW, Crotty S, Arnold JJ, *et al.* Poliovirus RNA-dependent RNA polymerase (3Dpol): structural, biochemical, and biological analysis of conserved structural motifs A and B. J.Biol.Chem. 2000;275(33): 25523-25532.

35. van Aken D, Zevenhoven-Dobbe J, Gorbalenya AE, *et al.* Proteolytic maturation of replicase polyprotein pp1a by the nsp4 main proteinase is essential for equine arteritis virus replication and includes internal cleavage of nsp7. J.Gen.Virol. 2006;87(Pt 12):3473-3482.

36. Wassenaar AL, Spaan WJ, Gorbalenya AE, *et al.* Alternative proteolytic processing of the arterivirus replicase ORF1a polyprotein: evidence that NSP2 acts as a cofactor for the NSP4 serine protease. J.Virol. 1997;71(12):9313-9322.

37. Adedeji AO, Marchand B, te Velthuis AJ, *et al.* Mechanism of nucleic acid unwinding by SARS-CoV helicase. PLoS.One. 2012;7(5):e36521.

38. von Brunn A, Teepe C, Simpson JC, *et al.* Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFeome. PLoS.One. 2007;2(5):e459.

39. van den Born E, Gultyaev AP, Snijder EJ. Secondary structure and function of the 5'-proximal region of the equine arteritis virus RNA genome. RNA. 2004;10(3):424-437.

40. Nedialkova DD, Gorbalenya AE, Snijder EJ. Arterivirus Nsp1 modulates the accumulation of minus-strand templates to control the relative abundance of viral mRNAs. PLoS.Pathog. 2010;6(2):e1000772.

41. van der Meer Y, van TH, Locker JK, *et al.* ORF1a-encoded replicase subunits are involved in the membrane association of the arterivirus replication complex. J.Virol. 1998;72(8):6689-6698.