Chapter 8

Dual functionality for the SARS-CoV nsp(7+8) RNA polymerase at the genomic 3′ end

Aartjan J.W. te Velthuis[1], Isabelle Imbert[2], Bruno Canard[2], Alexander E. Gorbalenya[1], Sjoerd H.E. van den Worm[1] and Eric J. Snijder[1]

[1] Molecular Virology laboratory, Department of Medical Microbiology, Center of Infectious Diseases, Leiden University Medical Center, PO Box 9600, 2300RC Leiden, The Netherlands.
[2] Architecture et Fonction des Macromolecules Biologiques, UMR 6098 Centre National de la Recherche Scientifique and Universites d'Aix-Marseille I et II, Case 925, 13288 Marseille Cedex 9, France.

Present address: Vaccine and Gene Therapy Institute, Oregon Health and Science University, Beaverton, Oregon, USA

Abstract

Despite identification of two virus-encoded RNA-dependent RNA polymerases (RdRps) genome, many critical steps in coronavirus RNA synthesis remain to be elucidated at the molecular level. Here, using the SARS-coronavirus model, we describe intriguing biochemical observations suggesting that the initiation of minus-strand RNA synthesis and the polyadenylation of both genome-length and subgenomic mRNAs may depend on a unique platform that involves an interplay of the nsp(7+8) RNA polymerase with a conserved 3′-terminal signature of the viral genome. Specifically, our *in vitro* studies implicate the 3′-terminal GAC^{OH}, which can be polyadenylated or used as template for *de novo* initiation of minus strand synthesis. We hypothesise that this specificity is partly achieved through the template-independent formation of a 5′-pppGpU^{OH} product that is complementary to this sequence. Lastly we explore here the intrinsic regulation between the two SARS-CoV encoded RdRps and the implications thereof for regulation of their activities.

Introduction

The initiation of RNA synthesis, *i.e.* the formation of a dinucleotide product through condensation of the first nucleotide bond, is a crucially important function of the viral RNA-dependent RNA polymerases (RdRp) during genome replication. Accordingly, initiation is typically coupled to the recognition of specific cis-acting "promoter elements" in the 3′ untranslated region (UTR) of the viral genome in order to secure (i) specificity for the viral promoter in the infected cell, and (ii) negative self-regulation when these RNA sequences were not faithfully copied in the previous cycle [188,197,406,407,408]. So far, a number of initiation mechanisms have been documented that guide RNA viral RdRps to accurate duplication of these important RNA sequences, including i) internal prime-realign initiation [190], ii) *de novo* initiation directed by the 'priming loop' of the RdRp and RNA structures [139], and iii) protein-primed initiation [206,207,208]. In addition, non-templated initiation of RNA synthesis has been proposed as a fourth mechanism and is thought to be used by the RdRps of non-segmented negative strand RNA (-RNA) viruses and the dsRNA rotaviruses. Conclusive biochemical support for the latter mechanism, however, is presently still lacking [409,410].

In addition to various replication elements, the genomic 3′ UTR contains a polyadenylyl (polyA) tail in many positive stranded (+RNA) viruses. Although it is believed that this structure is chiefly involved in regulating the translational efficacy and stability of RNA molecules in the cytoplasm of the cell, it may also contribute to viral replication [183,251,252,411]. In RNA viruses, mRNA polyadenylation is typically based on mechanisms such as i) repeated 'stuttering' on a short polyU tract in the viral minus strand [253,412,413], ii) direct copying of a 20 nt long polyU sequence encoded at the 5′ end of the minus strand RNA [251,408], or iii) non-templated, terminal extension of the viral mRNA [247,248,250]. Additionally, whereas cellular polyA tails, are mainly synthesised in a template-independent manner by nuclear polyA polymerases (PAPs) following the 3′ processing of pre-mRNAs by the nuclear cleavage and polyadenylation specificity factor (CPSF) [414,415] most viral +RNA genomes appear to be polyadenylated at the site of viral replication in the cytoplasm.

In contrast to the above examples, it is completely unknown which initiation or polyadenylation mechanisms are used during the complex RNA synthesis of coronaviruses (CoVs). This +RNA virus group infects a wide range of vertebrates and is renowned for its exceptionally large polycistronic genome of approximately 30 kilobases (Fig. 1A) [65]. Moreover, CoV genome replication must operate in balance with the synthesis of four to nine sg mRNAs [51,265] - crucial molecules that are 5′-capped and 3′-polyadenylated like the genome. They are thought to be produced from complementary subgenome-length -RNA templates that derive from discontinuous negative strand synthesis [51,265,416,417,418,419,420] and it is likely that the synthesis of sg RNAs utilises the same initiation and extension signals as those used for anti-genome synthesis and

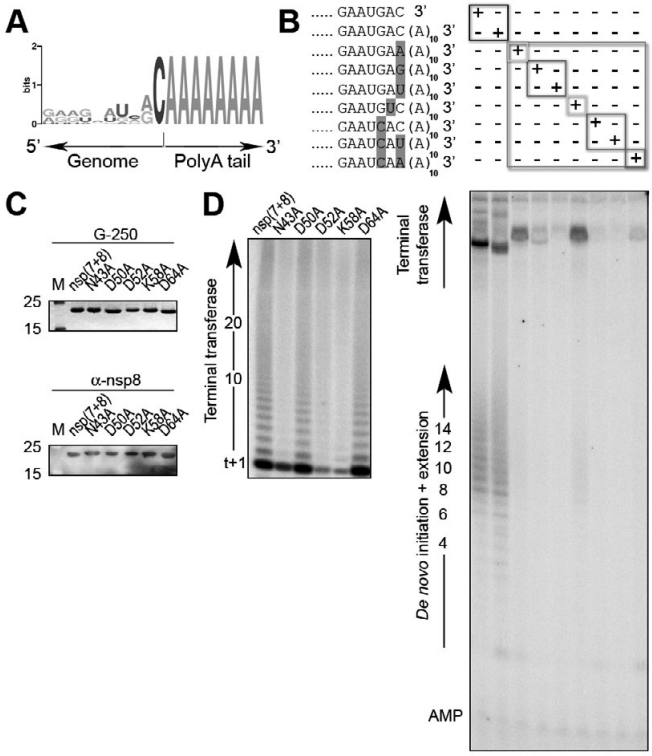

Figure 1: SARS-CoVnsp(7+8) has two distinct activities on the SARS-CoV genomic 3′ end. (**A**) Conservation of the genomic 3′ residues of the CoV genome. For sequence alignment see Fig. S1. (**B**) To test whether nsp(7+8) could use the SARS-CoV genomic 3′ end as template, purified nsp(7+8) was incubated with [α-^{32}P]ATP as readout, 10 μM ATP, 500 μM GTP, 100 μM of CTP and UTP, and a 111 nt long RNA template that was based on the SARS-CoV genomic 3′ end. As controls, side-directed mutagenesis was used to introduce mutations in the 3′ end template. Mutations are clustered based on their effect on nsp(7+8) activity. The AMP contaminant present in the radioactive label is indicated as internal loading control. (**C**) To test whether the terminal transferase activity on the 3′ genomic end was specific, mutations were introduced in conserved nsp8 residues. Mutant proteins were purified using the same protocol used for the wild type enzyme and analysed by SDS-PAGE. Upper panel shows Coomassie G-250 staining. Lower panel shows Western-blot analysis with a monoclonal nsp8 antibody. (**D**) Terminal transferase activity of the nsp8 mutants in the presence of nsp7 relative to the nsp(7+8) RdRp. Mutation of residues D52A and K58A resulted in a ~80% loss of terminal transferase activity. The activity of mutant N43A was reduced by ~30%, while mutants D50A and D64A showed wild type-like activities.

genome polyadenylation. Coronavirus genomes are further characterised by the strict conservation of a 3′ terminal cytosine just upstream of the polyA tail (Fig. 1A, Fig. S1). The mechanism and purpose of its conservation are unclear, however, just like it remains to be elucidated how the previously defined RNA sequences and structures in the 3′ end of the CoV genome minister control over the initiation of viral minus strand synthesis [334,421,422].

Complicating matters further, current evidence suggests that, uniquely among RNA viruses, CoV RNA synthesis involves not one, but two viral proteins with RdRp activity [154,155,156]. Both these proteins belong to the set of 16 CoV nonstructural proteins (nsps) that are produced through autoproteolytic cleavage of the pp1a and pp1ab precursor polyproteins [44,58] and co-localise with the membrane-associated viral replication machinery in infected cell [68,294]. Presently, these enzymes have only been studied in some detail for the Severe Acute Respiratory Syndrome-associated coronavirus (SARS-CoV) [154,155,156], which emerged in 2003 and caused worldwide concern due to the ~10% mortality rate associated with its infection of humans [65,324]. Of the two CoV RdRps, the first was found to employ a primer-dependent initiation mechanism and to be residing in the conserved canonical RdRp domain of the 106-kDa nsp12 [154,325]. The second on the other hand, the 22-kDa nsp8, was reported to be capable of *de novo* RNA synthesis and primer extension, both alone and in complex with the 10-kDa nsp7, *i.e.* nsp(7+8) [155,156].

In spite of identification of the two CoV RdRps, many fundamental questions regarding CoV RNA synthesis remain unanswered. For instance, it is unclear what viral template is used by nsp8 for *de novo* initiation and whether this includes specific RNA signals in the genomic 3′ UTR. In addition, it is unknown whether nsp8's *de novo* activity may be able to support the primer-dependent activity of nsp12. To investigate the RdRp properties of nsp8 in more detail and seek answers to these questions, we here studied the activity of the SARS-CoV nsp(7+8) polymerase and found it capable of both non-templated dinucleotide synthesis and subsequent extension of these products into ~10-nt long RNAs using a template representing the 3′ end of the viral genome. Interestingly, 3′-terminal extension of this template with adenylyl moieties was also observed, while mutation of the 5′-GAC^{OH} signature specific for the 3′ end of the SARS-CoV genome or amino acid substitutions in SARS-CoV nsp8 could abolish or separate the initiation of -RNA strand synthesis and +RNA polyadenylation *in vitro*. Overall, we propose that these results define the CoV nsp(7+8) complex as a unique RNA polymerase that recognises genomic 3′-terminal sequences and plays a critical role in both initiation of minus strand synthesis, the first catalytic step of SARS-CoV replication, as well as the termination of plus-strand synthesis.

Results

SARS-CoV nsp(7+8) has two activities on genome 3′ end-derived templates

SARS-CoV nsp7 and nsp8 were reported to interact and form the previously crystallised hollow ring structure that is composed of an intricate nsp8 octamer supported by eight copies of nsp7 [153,289]. In addition, the nsp(7+8) complex was shown to have RNA polymerase activity that is catalysed by residues in the nsp8 subunit [155,156]. Previous

biochemical analysis of recombinant SARS-CoV His-nsp8 revealed the enzyme's preference for *de novo* initiation on RNA templates containing a 5′-(G/U)CC^{OH} signature [155]. Interestingly, the coronavirus genome contains a fully conserved terminal cytosine as well (Fig. 1A and S1). Together with the compensating mutations identified in nsp8 after mutation of the 3′ end of the mouse hepatitis virus genome, this suggests that nsp8 may interact with the genomic 3′ end during -RNA synthesis [334].

To explore this hypothesis we incubated purified SARS-CoV nsp(7+8) with 111-nt long transcripts representing the SARS-CoV genomic 3′ end. In addition, point mutations were introduced into this RNA sequence to investigate template-dependence and the role of the terminal cytosine in template recognition. Compellingly, nsp(7+8) was readily able to synthesise RNA products of >8 nt in size on templates containing the wild-type genomic sequence, both in the presence or absence of a polyA$_{10}$ tail (Fig. 1B). Mutation of the 3′-terminal cytosine to any other nucleoside, however, abolished [α-^{32}P]AMP incorporation into short RNA products (Fig. 1B). Strikingly though, some residual [α-^{32}P]AMP incorporation was still observable for products beyond template length when the 3′ cytosine was mutated to adenosine (Fig. 1B). We assumed that this activity was identical to the terminal transferase activity of nsp(7+8) on primed templates [156] and further gel analysis of the products from assays using the wild-type, polyA-tailed template did indeed reveal a classic tailing pattern (Fig. 1D). In addition, the observed products appeared to be longer on mutant templates than on the wild type 3′ end, suggesting that the formation of the short RNAs competes with the terminal transferase activity on the 3′ genomic end. Similar results were obtained when we mutated the 3′ terminal 5′-GAC^{OH} to 5′-GUC^{OH} (Fig. 1B), whereas a loss of both activities was observed for all other substitutions in the 3′ terminal 5′-GAC^{OH} sequence (Fig. 1B).

Mutational analysis of SARS-CoV nsp(7+8) identifies positions involved in terminal transferase activity

To verify that the observed terminal transferase-like activity was nsp(7+8) derived, mutations were engineered in nsp8 via site-directed mutagenesis. Mutant proteins were purified in parallel with the wild-type protein and analysed by SDS-PAGE, Western blot and gel filtration (Fig. 1C). When we subsequently investigated the terminal transferase activity of these enzymes, we found that substitution of aspartate-52 to alanine (D52A) and lysine-58 to alanine (K58A) significantly impaired the observed activity (Fig. 1D). Mutation of aspartate-50 and the non-conserved aspartate-64 to alanine (D50A and D64A) on the other hand did not affect the activity at all, whereas alanine substitution of the conserved of asparagine-43 resulted in a ~30% loss of activity. In part, these observations are in line with the importance of D52 for nsp(7+8) primer-extension activity and the role of K58 in RNA binding [156].

SARS-CoV nsp(7+8) terminal transferase activity prefers ATP, but requires GTP for template specificity

As observed in Fig. 1B and described elsewhere [156], activity reminiscent of terminal transferase activity was observed for SARS-CoV nsp(7+8). To confirm the nature of this activity, an oligonucleotide mimicking the last 10 nucleotides of the SARS-CoV genome was first blocked with cordycepin (3′deoxyadenosine triphosphate, 3′dATP) using a commercially available polyA polymerase and subsequently incubated with purified SARS-CoV nsp(7+8) and [α-^{32}P]ATP. As shown in Fig. 2A, ^{32}P-labelling of 3′-terminally blocked template did not occur, whereas tailing was still observed for the unblocked control substrate (Fig. 2A).

We next explored the nucleotide preference of SARS-CoV nsp(7+8) for 3′ extension of the SARS-CoV genome and titrated ~30-, ~300- or ~3000-fold excess of either ATP, GTP, CTP or UTP in a terminal transferase assay containing 0.17 μM [α-^{32}P] ATP, to assess their ability to compete for incorporation. In line with the preference of poliovirus 3D^{pol} [247], unlabelled ATP was the most effective competitor of [α-^{32}P] ATP, suggesting that nsp(7+8) preferentially synthesises polyA tails (Fig. 2B, Fig. S2). At larger excesses, pyrimidines were also efficiently incorporated into the tails, while semi-efficient GMP incorporation was observed at a 3000-fold excess over [α-^{32}P]AMP as was evident from the 'G-jumps' in the tailing pattern (Fig. 2B, Fig. S2). This thus leads us to suggest that the order of nucleotide preference for nsp(7+8) is ATP>CTP>UTP>GTP.

Following up on the above finding and our observation that mutation of the 3′ sequence of the SARS-CoV genome influenced terminal transferase activity (Fig. 1B), we investigated the relation between the nucleotide substrates present and the template sequence. Interestingly, when nsp(7+8) was incubated with [α-^{32}P]AMP in absence of other nucleotides, nsp(7+8) preferentially tailed our control oligonucleotide consisting of a polyU tail followed by a 3′-terminal 5′-CCAAAA^{OH} sequence and disregarded the templates that were based on the genomic 3′ end (Fig. 2C, left panel). This picture completely changed, however, upon the addition of GTP, which significantly shifted the tailing preference to the substrates matching the viral 3′-terminal sequence (Fig. 2C, compare left and middle panel). This shift in template selection also included an equal preference for the mutant 3′ end oligonucleotides, in the presence of either GTP only or all four NTPs (Fig. 2C, middle and right panel), suggesting that although GTP plays a role in template selection, nsp(7+8) likely uses additional template-enzyme contacts for fine-tuning and to achieve the specificity observed in Fig. 1B.

SARS-CoV nsp(7+8) can synthesise dinucleotides in the absence of an RNA template

To investigate nsp(7+8)'s activity on the 3′ end during *de novo* RNA synthesis and to study in particular the first step leading to dinucleotide formation - which was previ-

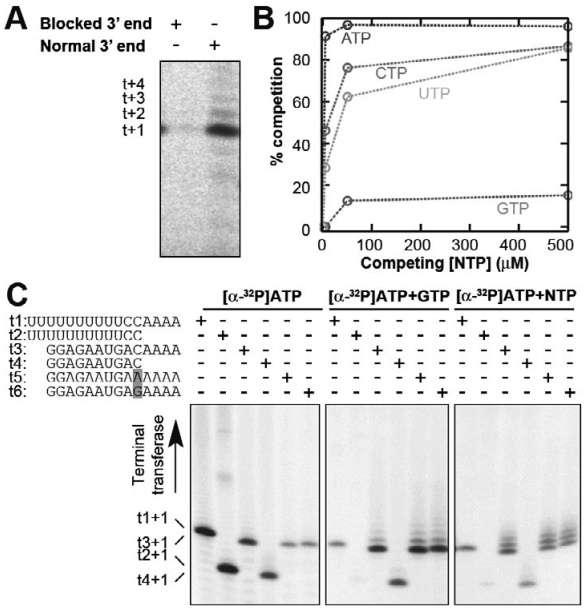

Figure 2. Terminal transferase specificity of SARS-CoVnsp(7+8). (**A**) No terminal transferase activity of nsp(7+8) was observed when the 3′ end of an oligoribonucleotide representing the 3′ end of the SARS-CoV genome (oligo t3, see Fig. 2C) was blocked with cordycepin (3′dA) and polyA polymerase. Some residual labelling was visible due to the incomplete blocking of the oligo. (**B**) Using [α-^{32}P]ATP as readout, we tested the competition for other NTPs with the 0.17 μM labelled nucleotide for their use in the terminal transferase activity of nsp(7+8). Ratios were 1:30, 1:300 and 1:3000. For original gel data see Fig. S2. (**C**) To further explore the requirements for the nsp(7+8) terminal transferase activity, we reduced the RNA template length and removed all nucleotides except the [α-^{32}P]ATP readout. Interestingly, this resulted in a general change in terminal transferase specificity (left panel) and a preference for the control templates t1 and t2. The effect could be reversed via the addition of 500 μM GTP alone (middle panel), or through the addition of 500 μM GTP and 100 μM CTP and UTP (right panel). No unlabelled ATP was added to prevent competition for the radioactive signal. All reactions were incubated for 30 min at 30 °C and analysed on 20% 7M Urea PAGE gels.

ously postulated to be pppGpA synthesis [155] - we incubated nsp(7+8) with either wild-type or mutant 3′ template, [α-^{32}P] ATP and guanosine. This latter component was added to replace GTP in the reaction and to exclude the incorporation of GMP at the +2 position. Strikingly, dinucleotide products were observed in all reactions and their formation appeared to be unabated by mutation of the penultimate 3′ C or substitution of the upstream A with U (Fig. 3A). The nps(7+8)-dependent terminal transferase activity was affected, however, in line with the results presented in Fig. 1B. These results therefore suggested that nsp(7+8) was able to identify the template correctly under these conditions and that dinucleotide synthesis had taken place in a template-independent manner. Indeed, GpA formation was also observed in the absence of an RNA template (Fig. 3A).

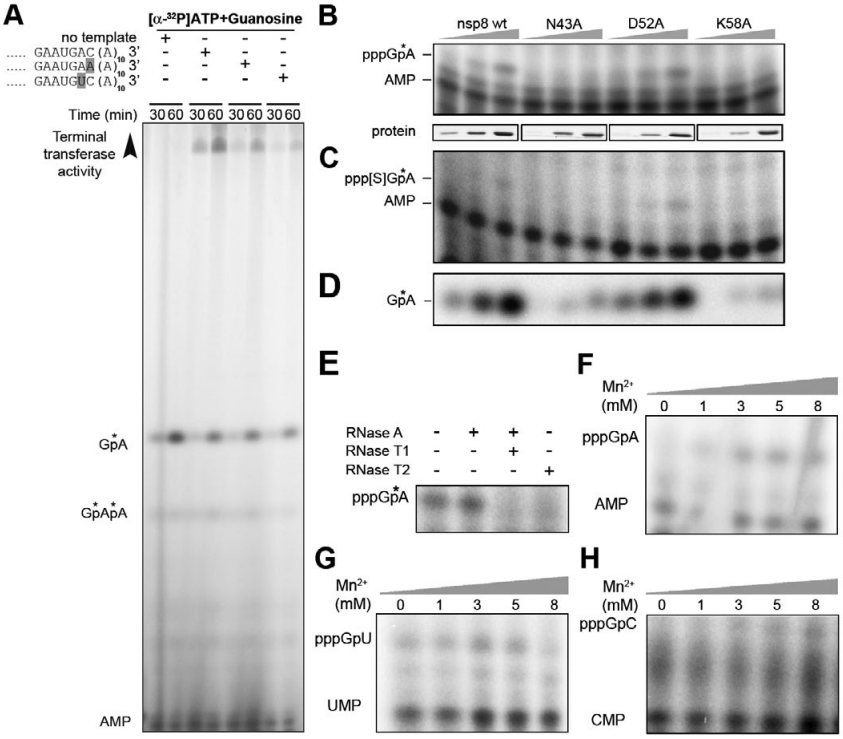

Figure 3. Initiation of RNA synthesis by SARS-CoVnsp(7+8) requires GTP and is non-templated. (**A**) To test whether GTP was used for primer synthesis nsp(7+8), we incubated purified nsp(7+8) with 111 nt long SARS-CoV 3′ end templates, 500 μM guanosine and [α-^{32}P]ATP. The guanosine was used to better separate the formed dinucleotide from the unincorporated label, distinguish it from putative pppApA products and prevent nsp(7+8) from forming pppGpG products. Interestingly, a GpA dinucleotide was also formed in absence of template. Asterisk indicates position of the [α-^{32}P]-group. The AMP contaminant present in the radioactive label is indicated as internal loading control. (**B**) Control reactions showing that dinculeotides can be efficiently formed using GTP, (**C**) alpha-S-GTP or (**D**) guanosine at the +1 position in absence of template. In addition, these figures show that the activity is nsp8-specific using nsp8 mutant N43A, D52A and K58A. Reactions contained 0.1, 0.5 or 1 μM of wild-type nsp8, mutant D52A or mutant K58A. Mutation of conserved residues N43 and K58 to alanine results in a complete loss (<1%) of *de novo* dinucleotide synthesis in the presence of GTP. Similar to the terminal transferase assay, the activity of D52A was not significantly altered in comparison to the wild-type protein. Lower panels show the SDS-PAGE analysis of the diluted protein stocks of wild-type or mutant nsp8 used in the assay. (**E**) Synthesis of the pppGpA dinucleotide product by nsp8 in the presence or absence of RNases A, T1 and T2. No template was added to these reactions. (**F**) Comparative analysis of [α-^{32}P]ATP, (**G**) [α-^{32}P]UTP, and (**H**) [α-^{32}P]CTP incorporation at the +2 site under different Mn^{2+} concentrations. The ANP contaminants present in the radioactive labels are indicated as internal loading controls.

To exclude that this effect was the result of a contamination or the absence of the triphosphate on the guanosine, we performed dinucleotide assays in the presence of GTP, the non-hydrolysable GTP analogue GTP-alpha-S (guanosine 5′-[alpha-thiotriphosphate];

ppp[S]G) and guanosine, and compared the *de novo* activity of wild-type nsp(7+8) and nsp8 mutants N43A, D52A and K58A. As shown in Fig. 3B and 3C, nsp(7+8) was able to synthesise equal amounts of pppGpA and ppp[S]GpA dinucleotide products, which not only confirmed the important role of the guanosine at the +1 position, but also suggested that is was not significantly competing with [α-^{32}P]ATP for incorporation at the +2 site. In addition, this result also implied that hydrolysis of the GTP's triphosphate was not required to drive the reaction, which is in contrast to its role as energy donor for the covalent uridylylation of the poliovirus VPg protein primer [206]. Of all nsp8 mutations tested, K58A had the strongest effect on dinucleotide synthesis in all control reactions, thus confirming the specificity of the assay (Fig. 3B-D). Mutation D52A was not able to impair formation of pppGpA, in stark contrast to its role in the terminal transferase assay (Fig. 1).

To exclude that the observed reactions were templated by inadvertently co-purified RNAs, we supplemented the template-free dinucleotide reactions with RNase A (Fig. 3E). Known for its ability to hydrolyse RNA 3′ of C and U residues, this RNase could thus specifically degrade any RNAs containing sequences that may have templated pppGpA synthesis. However, the presence of RNase A did clearly not affect the synthesis of a pppGpA product at all, whereas all product formation was lost in negative control reactions that included RNAse T1 or T2 to cleave ssRNA products 3′ of G or N, respectively (Fig. 3F). These results thus suggest that dinucleotide synthesis by nsp(7+8) is insensitive to both the absence of template and the presence of nucleases capable of degrading all RNA templates complementary to the product. Moreover, together with our other observations, they provide a strong indication that the observed pppGpA dinucleotide product was synthesised *de novo* without the use of an RNA template.

SARS-CoV nsp(7+8) preferentially synthesises a pppGpU dinucleotide

Up to this point, all dinucleotide activity experiments presented in Fig. 3 had been performed in the presence of Mn^{2+}, given the reported dependence of His-nsp8's *de novo* activity on this cation [155]. In light of the effect of Mn^{2+} on the SARS-CoV nsp12 and poliovirus 3D^{pol} fidelity [154,279], we tested whether the presence of Mn^{2+} had the ability to alter the nucleotide specificity for nucleotide incorporation at the +2 position. To this end, we titrated Mn^{2+} in dinucleotide synthesis assays containing either [α-^{32}P]AMP, [α-^{32}P]UMP or [α-^{32}P]CMP as readout. Interestingly, the incorporation of both AMP and CMP required millimolar concentrations of Mn^{2+} in addition to Mg^{2+} (Fig. 3G and 3H), whereas the incorporation of UMP only required Mg^{2+} (Fig. 3I). This thus leads us to suggest that the preferred nsp(7+8) dinucleotide product is pppGpU, which is strikingly complementary to the 3′-terminal sequence immediately upstream of the polyA tail in the genome of SARS-CoV and all closely related betacoronaviruses (Fig. S1). In addition, this result explains the absence of a significant dinucleotide product in the reactions

presented in Fig. 1B, as these were visualised through [α-^{32}P]AMP incorporation rather than [α-^{32}P]UMP.

A two-way regulation of SARS-CoV nsp12 and nsp8 activity

With the previously proposed primase hypothesis in mind [155], we next sought to investigate whether the products produced by nsp(7+8) under the conditions in Fig. 1 could be extended by nsp12. Unfortunately, however, no change in the position of the products or product accumulation was observed, both when nsp12 was added after a preincubation with nsp8 or nsp(7+8) and when nsp12 was present from the start. Given that three studies have independently shown that SARS-CoV nsp12 and nsp8 are able to interact [85,289,423], we hypothesised that one of the enzymes was able to modulate the activity of the other and prevent the reaction from progressing. To verify the existence of such an effect, we designed an experiment to model the hypothesised nascent strand hand-off between nsp(7+8) and nsp12 (Fig. 4). Specifically, we employed a primed-template that consisted of a duplex derived from an internal sequence of the 3′ UTR of the genome [325] and a single-stranded template sequence with low secondary structure (Fig. 4D) on which both enzymes were active (Fig. S3) [154,325].

To minimise the number of potential outcomes of this experiment and simplify interpretation we decided to fix the nsp12 concentration and vary the amount of nsp8 or nsp(7+8) added to the reaction. As presented in Fig. 4A and 4B, putative outcomes were likely to be qualitatively similar (multiple were either inhibiting or stimulating), but given that nsp12 is a stronger polymerase and has a lower RNA binding affinity than nsp8 and nsp(7+8) [156], we expected the differences in signal to be significant enough to postulate whether nsp8 inhibited nsp12, nsp12 inhibited nsp8, or whether the inhibition was reciprocal (Fig. 4A and 4B). In addition, we hoped to achieve further discrimination between the outcomes by performing a titration experiment under conditions where nsp8 had gained or lost activity, since this could either negate part of the effect or have no effect at, depending on whether the influence of the two enzymes was mediated through the template or via a direct protein-protein interaction (compare Fig. 4B and C).

Interestingly, when a fixed concentration of 0.1 μM wild-type nsp12 was titrated with either wild-type nsp8 or the nsp8 template-binding mutant K58A in the presence of an excess of RNA template and NTPs, we observed a steep drop in the overall polymerase activity (Fig. 4E and F). The inhibitory effect of K58A was significantly stronger than that of wild-type nsp8, however, and essentially resulted in a reduction to the lower detection limits of the assay. A smaller effect on the initial nsp12 activity was observed when nsp(7+8) was added to the reaction (Fig. 4E panel 3, Fig. S3). To confirm that the highest concentrations used did not result in competition for template binding, we analysed the reactions by native PAGE. As shown in Fig. 4G, mutant K58A was unable to bind substrate and inhibit nsp12 template binding, whereas the presence of wild-type nsp8

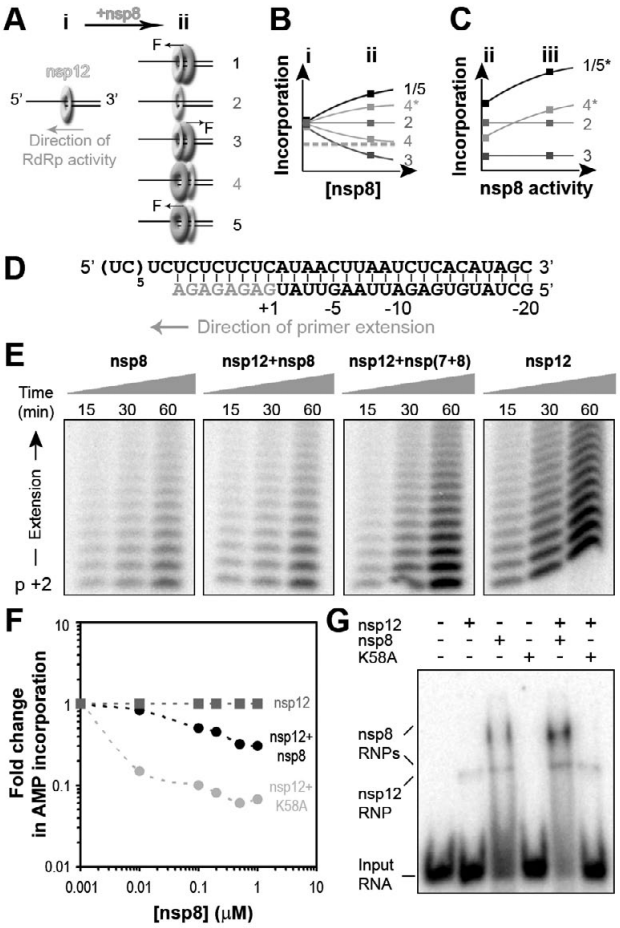

Figure 4: SARS-CoV nsp8 inhibits nsp12-dependent primer-extension activity. (**A**) Addition of nsp8 to a model primer-extension reaction by the main SARS-CoV RdRp nsp12 may result in four potential scenarios: 1) nsp8 has a stimulatory effect on nsp12 (indicated by F in the direction of RNA synthesis) through binding to nsp12, while we assume that the observed activity is dependent on nsp12's interaction with the primer; 2) nsp8 has no effect and is outcompeted by nsp12 for access to the nascent strand; 3) nsp8 has an inhibitory effect on nsp12 (indicated by F in the direction antipodal to RNA synthesis) by binding to nsp12, while catalysis is performed by nsp12; 4) nsp8 is active rather than nsp12, prevent nsp12 from accessing the 3'OH of the nascent strand and slows down the reaction as a results of its lower activity; 5) nsp8 is active rather than nsp12 and is stimulated by nsp12 through an allosteric effect, indicated with F in the direction of RNA synthesis. The dark grey dashed line indicates the basic nsp8 activity level. (**B**) The addition of nsp8 to nsp12 results in an overall activity that is one of five hypothetical functions of the nsp8 concentration. The effect indicated by 4* is presented to portray the difference with Fig. 4C. (**C**) In theory, the combined activity of nsp12 and nsp8 can also be studied as a function of the nsp8 activity, while keeping the nsp8 concentration constant. (**D**) RNA template used to study the primer extension activity of nsp12 and nsp8. (**E**) Side-by-side comparison of the incorporation activities of nsp8 alone, nsp12+nsp8, nsp12+nsp(7+8), and nsp12 alone in time. (**F**) Average change in AMP incorporation. In the presence of nsp8, we observe a decrease in the total activity (black circles).

Figure 4 continued
A much stronger decrease in activity was observed when we combined nsp12-RdRp with nsp8 template binding mutant K58A (grey circles). (**G**) Electromobility shift assays showing that the presence of nsp8 K58A does not negatively affect nsp12 template binding. In fact, when both nsp8 and nsp12 were present the total amount of bound template was higher. This was particularly noticeable at the height of the upper nsp8 RNP complex.

resulted in an overall increase in template binding that was not only additive due to the fact that both enzymes bound RNA, but also evident as a ~2 fold enhanced complex formation with the largest wild-type nsp8 complex (Fig. 4G). Given that the smaller migrating nsp8 complex runs at a similar position as the nsp12 enzyme (106 kDa), it is likely that the smaller complex represents the nsp8 tetramer, while the larger complex would represent the nsp8 octamer (8x 22 kDa) bound to RNA [153,156].

Increased nsp(7+8) activity overcomes nsp12 inhibition

To further untwine the interplay between both enzymes and differentiate between the possible interactions of nsp8 and nsp12, we aimed to stimulate the activity of nsp(7+8) relative to nsp12 (Fig. 4C). Specifically, we hoped to find whether the effect of nsp8 was interaction- or activity-dependent. Presently, however, no chemical has been identified that can specifically inhibit or stimulate one SARS-CoV RdRp relative to the other. We therefore attempted to raise the overall activity level of the assay through the addition of Mn^{2+} and a reduction of the Mg^{2+} concentration, thereby stimulating the activity of both nsp12 and nsp8 (Fig 5A). To negate the effect of these conditions on incorporation fidelity, we simplified the template to U_{20}. Both enzymes are active on this template as shown in Fig. S3 and elsewhere [154,156].

As shown in Fig. 5, the more active nsp(7+8) now stimulated the overall incorporation level of the assay, while the inactive mutant K58A showed an unaltered inhibitory effect in line with its inability to benefit from the stimulating reaction conditions (Fig. 5C and D). Lastly and as a further control measure, we also titrated the nsp12 active site mutant D618A in reactions containing either wild type nsp8 or nsp8 mutant K58A. As shown in Fig. 5E, nsp12 was able to fully inhibit nsp8 activity and thus essentially behaved as a permanent 'brake' on the nsp8 activity (Fig. 5E). Overall, these results thus lead us to conclude that on the one hand, nsp8 must be able to exert an influence on nsp12 via a direct interaction given the effect of the template-binding mutant, but that the interplay is also regulated by the interplay on the template on the other hand.

Discussion

The complex replication and transcription process that coronaviruses initiate upon infection involves up to sixteen viral nsps and at least one host factor [51,79,290]. In spite of increased effort to untwine their interactions since the outbreak of SARS in 2003, the

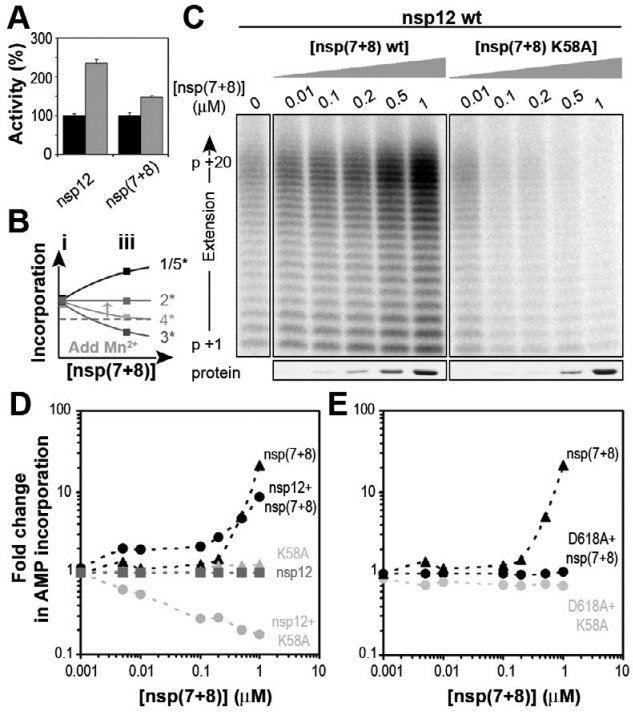

Figure 5: Stimulated SARS-CoV nsp8 can stimulate nsp12 in a concentration-dependent manner, whereas inactive nsp12 blocks nsp8 activity. (**A**) To stimulate the activity of SARS-CoV nsp8 we reduced the Mg^{2+} concentration from 6 mM Mg^{2+} (black bars) to 4 mM and added 1mM Mn^{2+} (green bars). This also resulted in a larger increase in activity of nsp12, albeit at a previously observed loss of fidelity. (**B**) The increase in nsp12-RdRp activity due to Mn^{2+} raised the basal nsp12 activity level in the assay (green arrow). Furthermore, using the higher nsp8 activity we were now able to explore its corresponding effects as proposed in Fig. 4B and 4C by varying the nsp8 concentration. (**C**) Fixed amounts of 0.1 μM nsp12-RdRp and 1 μM of U_{20} template were titrated with either wild-type nsp8 or nsp8 template-binding mutant K58A. Both nsp8 and nsp(7+8) were active on this template (see Fig. S3). Clearly, wild type nsp8 now stimulates the total RdRp activity, whereas the K58A mutant, which cannot benefit from the stimulating reaction conditions, still inhibits activity. (**D**) Quantification of the nucleotide incorporation activity relative to the activity of nsp12 alone (red squares). In absence of nsp12, the nsp8 activity increases 21 ± 0.5 fold over the course of the curve (black triangles), whereas in the presence of nsp12, this increase is only 8 ± 1 fold (n = 3) (black circles). Addition of nsp8 mutant K58A reduces the total AMP incorporation 9 ± 1 fold (grey circles). (**E**) To test whether nsp8-dependent nucleotide incorporation was still observed in the presence of inactive nsp12, we titrated a fixed amount of nsp12 D618A with either wild type nsp8 (black circles) or nsp8 mutant K58A (grey circles). In both cases no activity was observed. For reference, the activity of wild type nsp8 in absence of nsp12 is shown as black triangles.

exact mechanism that these enzymes use to catalyse RNA synthesis in the CoV RTC is still poorly understood. Moreover, the replication strategy is likely more complex than proposed for other +RNA viruses, since, uniquely among RNA viruses, CoVs employ two viral RdRps to catalyse CoV RNA synthesis [154,155,156]. The prevailing hypothesis to explain

this unique strategy suggests that these enzymes are functioning closely together and form a primase-replicative polymerase pair [155], putatively similar to the eukaryotic replisome, while one alternative hypothesis suggests that their function is separable and that their activity may be strictly regulated in space by other viral proteins [156].

Following up on the description of the polymerase activity of nsp8 and nsp(7+8) [155,156], we here demonstrate that the nsp(7+8) polymerase is capable of recognising the genomic 3′ end in order to perform two distinct activities: *de novo* RNA synthesis and terminal transferase activity (Fig. 1B and 6A). We further show that this enzyme can preferentially synthesise a pppGpU product (Fig. 3), which is complementary to the 3′ terminal 5′-GAC^{OH} of the SARS-CoV genome and may assist in recognition of the template and regulation of the two activities (Fig. 1 and 2). Moreover and significant for our understanding of CoV RNA synthesis, the recognition of the genomic 3′ end facilitated the synthesis of 8-14 nt short RNA products (Fig. 1B), which is in line with previous observations implicating 3′ terminal sequences and their interaction with nsp8 in the initiation of CoV replication [334,424]. Subsequent mutational analysis of the conserved 3′ terminal 5′-GAC^{OH} signature abolished the synthesis of these RNAs (Fig. 1B), similar to the observed effect of the mutational analysis of the 3′-terminal 5′-CC^{OH} for rotavirus and arterivirus RNA synthesis *in vitro* and *in vivo* [196,407].

Together, these findings suggest that although nsp(7+8) can initiate nucleotide condensation in the absence of an RNA template, it likely prefers an RNA template with a sequence complementary to the pre-formed dinucleotide to catalyse further incorporations (Fig. 6A). We assume therefore that the initiation of -RNA synthesis depends on the correct substrate stoichiometry as well as base pairing and interactions with secondary structures in the template, as also suggested by Züst *et al*. [334]. Future experiments should investigate these steps in more detail. Of particular interest in that regard is the mutational analysis of the nsp8 subunit presented here. As shown in Fig. 1D, this analysis revealed that contrary to D50's involvement in primer-extension [156], this aspartate was not required for terminal transferase activity. Similarly, the for primer-extension crucial D52 was not involved in pppGpA formation (Fig. 1D) [156]. These observations suggest that the various activities noted for nsp(7+8) thus far can be separated *in vitro*, which, for instance, opens up avenues to complementation experiments.

In addition, we also investigated the interplay between nsp8 or nsp(7+8) and nsp12 on a primed template, and observed in Fig. 4 and Fig. 5 that nsp8 has an inhibitory effect on nsp12 activity. Specifically, we found that this effect was dependent on the nsp(7+8) concentration and activity, suggesting that i) nsp8 is allosterically limiting nsp12 activity (Fig. 4A, model 3) or ii) preventing it from accessing the 3′-OH of the primer (Fig. 4A. model 4). In the latter case, the observed activities would become solely nsp8-driven at the higher concentrations, which in turn would postulate a requirement for additional factors to induce nsp(7+8)'s dissociation from the primer. Unfortunately, we can pres-

ently not ascribe a conclusive preference to either model, since the results obtained with nsp8 mutant K58A suggest that part of the influence must be mediated through an interaction, whereas the complete inhibitory effect of inactive nsp12 on active nsp8 suggests that the interplay may be organised through the template. Alternatively, it is a distinct possibility that the regulation between the two enzymes requires both direct (protein-protein) and indirect (substrate) influences, and possibly even additional viral proteins.

In summary, our results explain a number of ill-understood observations and touch upon at least three important questions regarding CoV replication, including the genetic linkage between mutations in the SARS-CoV genomic 3′ end and the nsp8-coding sequence, and the conservation of the 3′ terminal residue (Fig. S1). In light of the above two subjects, we demonstrated here that the nsp(7+8) complex can synthesise 8-14 nt short RNA products on RNA templates representing the 3′-terminal domain of the SARS-CoV genome and that these templates appeared to stimulate the RdRp activity compared to non-biological sequences in the presence of NTPs (Fig. 1B and 2). These results make it tempting to envision a mechanism in which these products are utilized by an elongation complex containing the putative main primer-dependent CoV RdRp, nsp12, as its key enzyme [155] (Fig. 6B, panel i). This brings us to a third question, however, because based on the activity of nps(7+8) on non-structured RNA templates [156] (Fig. S3), we cannot formally exclude the possibility that nsp(7+8) may synthesise substantially longer products *in vivo* (Fig. 6B, panel iii), possibly stimulated by the presence of additional viral protein factors. Moreover, the observed and apparently complex inhibitory effect of nsp8 and nsp(7+8) on the nsp12-dependent primer-extension activity (Fig. 4 and 5) asks whether i) additional factors are required to regulate their interaction and the nascent strand hand-off, possibly even through the release of nsp(7+8) from the template as proposed by Li *et al*. [425] (Fig. 6B, panel ii), or ii) that these two enzymes simply operate in separated complexes that each have their own dedicated RdRp and function in viral +RNA or -RNA synthesis in the infected cell. It will therefore become now particularly crucial to study whether these different polymerases associate *in vivo* and to what extent they influence each other's activity there.

Material and methods

Nucleotides and RNA modification enzymes

RNAases RNase A, T1 and T2 were purchased from Ambion. PolyA polymerase was purchased from USB, cordycepin 5′-triphosphate (3′dATP) from Sigma, and T4 kinase and Shrimp Alkaline Phosphatase (SAP) from Invitrogen. Radiolabelled nucleotides were purchased from Perkin-Elmer, marker dinucleotides GpA, GpU and ApA, and all RNA and DNA oligonucleotides from Eurogentec, nucleotide analog GTP-alpha-S from Jena Bi-

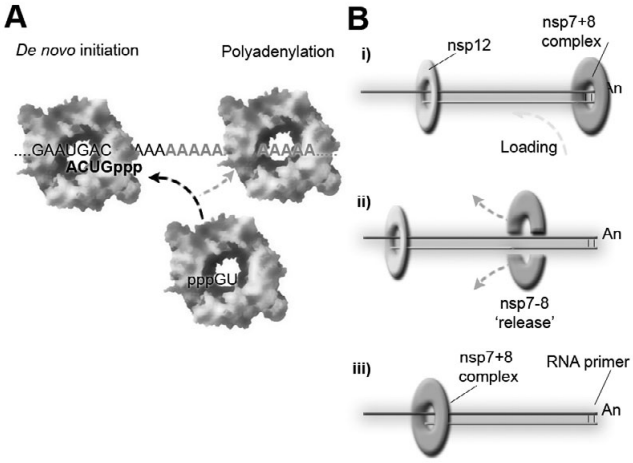

Figure 6: Models for nsp8 activity and the initiation of negative strand RNA synthesis. (**A**)Schematic presentation of the activities on SARS-CoVnsp(7+8). We hypothesise nsp(7+8) preferentially forms a pppGpU dinucleotide in absence of template that can subsequently be extended in a template dependent manner (black arrow). Additionally, nsp(7+8) may use the preformed product to align itself with the genomic end and catalyse the formation of a polyA tail (dark grey arrow). It is unknown whether in absence of *de novo* initiation, the dinucleotide product is release or whether it remains associated with nsp(7+8). (**B**) Panel i) shows a schematic presentation of the 'primase' hypothesis first proposed by Imbert *et al*, in which nsp8 or nsp(7+8) synthesises a primer that can be extended by SARS-CoV nsp12. Based on the data present in Fig. 1, the initiation site for nsp(7+8) may be on the genomic 3′ end. Panel ii) demonstrates an extension of the 'primase' hypothesis, in which SARS-CoVnsp(7+8) must first release its primer before nsp12 can commence extension of the primer. In panel iii) a model is presented in which no primer-hand off between nsp(7+8) and nsp12 is required at all and nsp(7+8) is assumed to fully catalyse minus strand formation.

osience, whereas guanosine and all standard nucleoside triphosphates were purchased from Sigma. For T7 transcription, an Ambion MegaShortScript kit was used according to the manufacturer's instructions.

Mutagenesis, protein purification and polymerase assays

All described nsp8 mutants were engineered via site-directed mutagenesis according to the QuikChange protocol (Stratagene) using the primers listed in table S1. All proteins were expressed and analysed as described previously [156]. SARS-CoV nsp7+8 assays were essentially performed as described elsewhere [156], with modifications. Specifically, nsp8 dinucleotide assays were performed either in the presence or absence of 1 µM synthetic oligoribonucleotides as listed in Fig. 2 or the presence or absence of SARS-CoV genomic 3′ end templates (see section below). Per nsp8 *de novo* reaction, typically 1µM monomeric nsp8 or nsp8 mutant and 1µM monomeric nsp7 was incubated with 4 mM $MgCl_2$, 1 mM $MnCl_2$, 1 mM GTP, 10 µM ATP, 0.17 µM [α-^{32}P]ATP, 1mM DTT, 10 mM

KCl and 20 mM Tris (pH 8.0). ATP and GTP were substituted with other nucleotides or nucleosides as indicated in the main text. Note that we performed the activity assays at pH 8.0 instead of the for nsp8 optimal pH of 9.5 [156] to better preserve RNA integrity. For reactions containing guanosine, we compensated for the addition of NaOH that was used to dissolve guanosine by lowering the pH of the reaction buffer to 6.5 prior to the addition of guanosine and enzyme.

RNA templates

Oligoribonucleotides for polymerase assays were excised as a single band from 7 M urea/15% polyacrylamide gels, eluted overnight in deionized water and desalted using NAP-10 columns (GE healthcare) that were equilibrated with deionized water. To anneal RNA duplexes, oligonucleotide mixtures in annealing buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl and 5 mM EDTA) were heat denatured and slowly cooled to room temperature.

Genomic 3′ end templates of SARS-CoV Frankfurt-1 were amplified via *pfu*-based PCR (Fermentas) using forward primers containing a T7 promoter region (see table S2). After the PCR reaction, the cDNA template was digested with *Dpn*I for 1 hour at 37 °C and the product purified via gel-extraction. Subsequently, T7 transcriptions were performed and the RNA products extracted with Trizol (Invitrogen), and glycogen (Roche) co-precipitation in ethanol.

Sequence alignment

Alignments of the genomic 3′ terminal residues were made using Muscle [327] and visualized with Weblogo 3.0. Sequences used included the alpha coronaviruses human CoV 229E (NC_002645), rat CoV (NC_012936), and bat CoV HKU8 (NC_010438); the beta coronaviruses SARS-CoV (AY291315), mouse hepatitis virus (NC_001849), and human CoV OC43 (NC_005147); and the gamma coronaviruses beluga whale CoV (NC010646), bulbul CoV (FJ376620) and avian infectious bronchitis virus (IBV, AJ311317).

Acknowledgements

The authors thank Dr. Danny Nedialkova, Dr. Clara Posthuma and Lorenzo Subissi for stimulating discussions. This work was supported by the Netherlands Organization for Scientific Research (NWO) through Toptalent grant 021.001.037 and ECHO grant 700.55.002 from the Council for Chemical Sciences (NWO-CW).

CHAPTER 8 - SUPPLEMENTAL INFORMATION

```
                                                  *  polyA
α | RAT-CoV      UGAAUGAAGUUGAUCAUGGCAAAAAAAAA
α | HCoV-229E    CGGUUUCGAUAUGGAUACACAAAAAAAA
α | Bat-CoV      UAGUUUUGAUAGGGAUUCACAAAAAAAA
β | MHV-A59      UGGCCAAUUGGAAGAAUCACAAAAAAAA
β | SARS-CoV     UAGCUUCUUAGGAGAAUGACAAAAAAAA
β | HCoV-OC43    UGGCCAAUUGGAAGAAUCACAAAAAAAA
γ | IBV-B        GGCUAGUAUAGAGUUAGAGCAAAAAAAA
γ | BW-CoV       GACUAUAGGUAAUUGUUAGCAAAAAAAA
γ | Turkey-CoV   GGCTAGTATAGAGTTAGAGCAAAAAAAA
δ | Bulbul-CoV   GCUUAAAUGGGGGAGGGAGCAAAAAAAA
```

Figure S1.Sequence alignment of genomic 3′ terminal sequences of alpha, beta, gamma and delta *coronavirinae*. Sequences used included the alpha coronaviruses Rat Parker CoV (NC_012936), Human CoV 229E (NC_002645), and Bat CoV HKU8 (NC_010438); the beta coronaviruses MHV A59 (NC_001849), SARS-CoV Fr-1 (AY291315), and Human CoV OC43 (NC_005147); the gamma coronaviruses avian infectious bronchitis virus (IBV, AJ311317), Beluga Whale CoV (NC010646), and Turkey CoV (NC_010800); and the delta coronavirus Bulbul CoV (FJ376620). The conserved 3′ cytosine is marked with an asterisk.

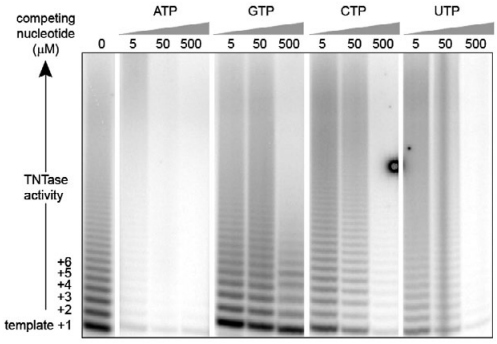

Figure S2:Nucleotide competition during terminal transferase activity. A SARS-CoV 3'end template was incubated with SARS-CoVnsp(7+8) and [α-^{32}P]ATP (lane 1). In parallel, 5, 50 or 500 μM concentrations of either ATP, GTP, CTP and UTP were added to this mixture to study their potential to compete with [α-^{32}P]ATP for incorporation into the template RNA. Note that although the incorporation preference for GTP is lowest, it is effectively incorporated when present in ~3000 fold excess over the [α-^{32}P]ATP as is evident of the 'G-jumps' in the migration pattern on 20% 7M Urea PAGE.

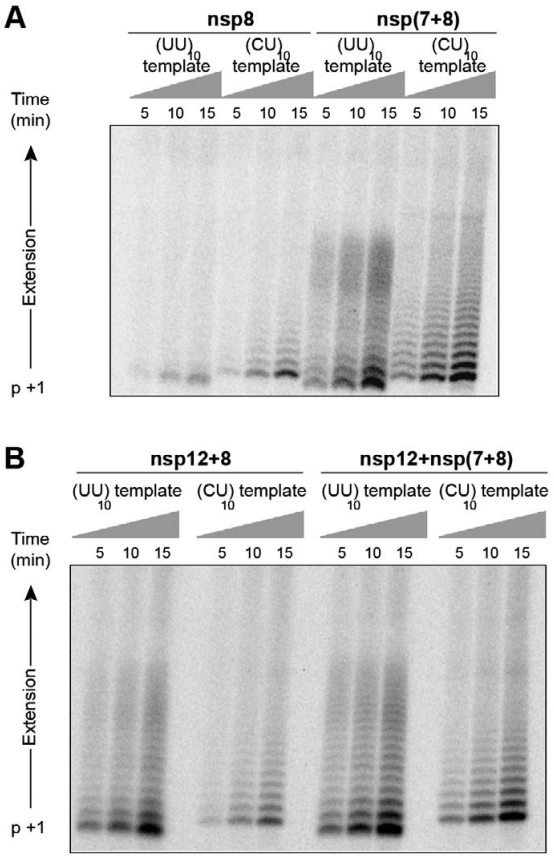

Figure S3: Primer-extension activities on various templates. RdRp reactions containing either the U20 or the UC10 template, SARS-CoV polymerases nsp8, nsp(7+8) or nsp12 were resolved by 20% 7M Urea PAGE. Since extension of the UC10 template requires a GTP for position +1 and the readout of this assay is [α-^{32}P]ATP, the first visible product is p+2.

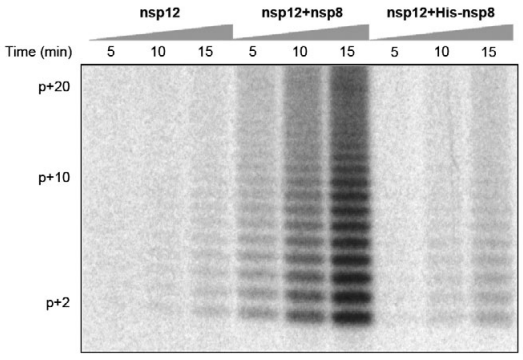

Table S1

nsp8 mutation	PCR primers	Sequence
N43A	SAV570	5′-TTAAAGAAATCTTTGGCTGTGGCTAAATCTGAG-3′
	SAV571	5′-CTCAGATTTAGCCACACCAAAGATTTCTTTAA-3′
D50A	SAV574	5′-GCTAAATCTGAGTTTGCCCGTGATGCTGCCATG-3′
	SAV575	5′-CATGGCAGCATCACGGGCAAACTCAGATTTAGC-3′
D52A	SAV590	5′-TCTGAGTTTGACCGTGCTGCTGCCATGCAACGC-3′
	SAV591	5′-GCGTTGCATGGCAGCAGCACGGTCAAACTCAGT-3′
K58A	SAV402	5′-GCCATGCAACGCGCTTTGGAAAAGATGG-3′
	SAV403	5′-CCATCTTTTCCAAAGCGCGTTGCATGGC-3′
D64A	SAV580	5′-TTGGAAAAGATGGCAGCTCAGGCTATGACCCAA-3′
	SAV581	5′-TTGGGTCATAGCCTGAGCTGCCATCTTTTCCAA-3′

Table S2

3′ terminal sequence	PCR primers	Sequence
---GACpA	SAV453	TTTTTTTTTT**GTC**ATTCTCCTAAGAAGCTATTA
	SAV209	TAATACGACTCACTATAGGGAGAGCTGCCTATATGGAAGAGC
---GAC	SAV212	**GTC**ATTCTCCTAAGAAGCTATTAAAATCACATGGGGATAGC
	SAV209	TAATACGACTCACTATAGGGAGAGCTGCCTATATGGAAGAGC
---GAGpA	SAV489	TTTTTTTTTT**CTC**ATTCTCCTAAGAAGCTATTA
	SAV209	TAATACGACTCACTATAGGGAGAGCTGCCTATATGGAAGAGC
---GAApA	SAV483	TTTTTTTTTT**TTC**ATTCTCCTAAGAAGCTATTA
	SAV209	TAATACGACTCACTATAGGGAGAGCTGCCTATATGGAAGAGC
---GAUpA	SAV490	TTTTTTTTTT**ATC**ATTCTCCTAAGAAGCTATTA
	SAV209	TAATACGACTCACTATAGGGAGAGCTGCCTATATGGAAGAGC
---GUCpA	SAV486	TTTTTTTTTT**GAC**ATTCTCCTAAGAAGCTATTA
	SAV209	TAATACGACTCACTATAGGGAGAGCTGCCTATATGGAAGAGC