

Favipiravir strikes the SARS-CoV-2 at its Achilles heel, the RNA polymerase

A. Shannon¹, B. Selisko¹, NTT Le¹, J. Huchting², F. Touret³, G. Piorkowski³, V. Fattorini¹, F. Ferron¹, E. Decroly¹, C Meier², B. Coutard³, O. Peersen^{4*} & B. Canard^{1*}.

¹ Architecture et Fonction des Macromolécules Biologiques, CNRS and Aix-Marseille Université, UMR 7257, Polytech Case 925, 13009 Marseille, France.

² University of Hamburg, Faculty of Sciences, Department of Chemistry, Organic Chemistry, Martin-Luther-King-Platz 6, D-20146 Hamburg, Germany.

³ Unité des Virus Émergents (UVE: Aix-Marseille Univ - IRD 190 - Inserm 1207 - IHU Méditerranée Infection), Marseille, France.

⁴ Department of Biochemistry & Molecular Biology, Colorado State University, Fort Collins, CO 80523-1870, USA.

*Correspondence to: either Olive.Peersen@colostate.edu or bruno.canard@afmb.univ-mrs.fr

Abstract: The ongoing Corona Virus Disease 2019 (COVID-19) pandemic, caused by severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), has emphasized the urgent need for antiviral therapeutics. The viral RNA-dependent-RNA-polymerase (RdRp) is a promising target with polymerase inhibitors successfully used for the treatment of several viral diseases. Here we show that Favipiravir exerts an antiviral effect as a nucleotide analogue through a combination of chain termination, slowed RNA synthesis and lethal mutagenesis. The SARS-CoV RdRp complex is at least 10-fold more active than any other viral RdRp known. It possesses both unusually high nucleotide incorporation rates and high-error rates allowing facile insertion of Favipiravir into viral RNA, provoking C-to-U and G-to-A transitions in the already low cytosine content SARS-CoV-2 genome. The coronavirus RdRp complex represents an Achilles heel for SARS-CoV, supporting nucleoside analogues as promising candidates for the treatment of COVID-19.

Coronaviruses (CoV) are large genome, positive-strand RNA viruses of the order *Nidovirales* that have recently attracted global attention due to the ongoing COVID-19 pandemic. Despite significant efforts to control its spread, SARS-CoV-2 has caused substantial health and economic burden, emphasizing the immediate need for antiviral treatments. As with all positive strand RNA viruses, an RdRp lies at the core of the viral replication machinery and for CoVs this is the nsp12 protein. The pivotal role of nsp12 in the viral life-cycle, lack of host homologues and high level of sequence and structural conservation makes it an optimal target for therapeutics. However, there has been remarkably little biochemical characterization of nsp12 and a lack of fundamental data to guide the design of antiviral therapeutics and study their mechanism of action (MoA). A promising class of RdRp inhibitors are nucleoside analogues (NAs), small molecule drugs that are metabolized intracellularly into their active ribonucleoside 5'-triphosphate (RTP) forms and incorporated into the nascent viral RNA by error-prone viral RdRps. This can disrupt RNA synthesis directly via chain termination, or can lead to the accumulation of deleterious mutations in the viral genome. For CoVs, the situation is complicated by the post-replicative repair capacity provided by the nsp14 exonuclease (ExoN) that is essential for maintaining the integrity of their large ~30 kb genomes¹⁻³. Nsp14 has been shown to remove certain NAs after insertion by the RdRp into the nascent RNA, thus reducing their antiviral effects⁴⁻⁶. Despite this, several NAs currently being used for the treatment of other viral infections have been identified as potential anti-CoV candidates⁷⁻⁹. Among these is the purine base analogue T-705 (Favipiravir, Avigan, Extended data Fig. 1a,b) that has broad-spectrum activity against a number of RNA viruses and is currently licensed in Japan for use in the treatment of influenza virus¹⁰. Clinical trials are currently ongoing in China, Italy, and the UK for the treatment of COVID-19, although a precise MoA of the drug against CoVs is lacking.

We infected Vero cells with CoV-SARS-2 in the presence or absence of 500 μ M T-705 and performed deep sequencing of viral RNA. A 3-fold ($P < 0.001$) increase in total mutation frequencies is observed in viral populations grown in the presence of the drug as compared to the no-drug samples (Fig. 1). Similar to previous findings with influenza¹¹, Coxsackie B3⁵ and ebola¹² viruses, a 12-fold increase in G-to-A and C-to-U transition mutations is observed, consistent with T-705 acting predominantly as a guanosine analogue. The increase in the diversity of the virus variant population suggests that once incorporated into viral RNA, T-705 is acting as a mutagen capable of escaping the CoV repair machinery. Interestingly, the SARS-CoV-2 genome has an already low cytosine content of ~17.6% and T-705 treatment may therefore place additional pressure on the already unbalanced CoV nucleotide content. Associated with this increase in mutation frequency, T-705 has an antiviral effect on SARS-CoV-2, as illustrated by a reduction in virus-induced cytopathic effect, viral RNA copy number, and infectious particle yield. Altogether these observations show that the mutagenic effect induced by T-705 is, at least in part, responsible for the inhibition of the replication.

To determine the efficacy and MoA of T-705 against SARS-CoV we first characterized nsp12 primer-dependent activity using traditional annealed primer-template (PT) and self-priming hairpin (HP) RNAs that may confer additional stability on the elongation complex

(Extended data Fig. 1c). Consistent with prior findings, nsp12 alone is essentially inactive^{13,14} and RNA synthesis requires the presence nsp7 and 8 cofactors whose stimulatory effect is enhanced by linking them as a nsp7L8 fusion protein^{6,15}. Structures of nsp12 show a four-component complex with a nsp7/nsp8 heterodimer and an additional nsp8 monomer^{16,17} and accordingly we found that addition of supplementary nsp8 to the nsp12:nsp7L8 complex further increases activity (Extended data Fig. 2). The resulting nsp12:7L8:8 complex is highly active on both PT and HP RNAs, with reactions containing 0.2 μ M of each substrate and 1 μ M nsp12 showing comparable initiation rates, with a rapid burst phase resulting in >50% primer consumption followed by remaining primer use over a period of a few minutes (Fig. 2a, Extended data Fig. 3). Interestingly however, the apparent processivity for the two substrates differs substantially. When provided with an annealed PT pair, intermediate products account for ~60% of the total lane intensity across all timepoints suggesting distributive polymerase activity (Fig. 2a, Extended data Fig. 4). In contrast, extension of a HP substrate has few intermediate products, indicating a processive elongation mode. This pattern is consistently observed across RNA substrates of different lengths, showing that the distributive PT mode does not convert into a processive state within a 30-nucleotide long template. Notably, the *Nidovirales* RNA replication/transcription scheme involves precise recombination-like events to generate subgenomic RNAs through a discontinuous mechanism along the ~30,000 nt genome¹⁸. The different processivities and complex stabilities we observe may be connected to these peculiar RNA synthesis events.

We sought to determine the extent to which the RdRp complex could incorporate the NA inhibitors T-705 and T-1105 (a non-fluorinated T-705-related analogue, Extended data Fig. 1a,b) into RNA. T-1105 shows improved potency against influenza virus in certain cell lines and is more stable than T-705 in the RTP form used in enzyme assays^{19–21}. The MoA for these compounds is currently controversial. T-705 has been shown to act through lethal mutagenesis for several viruses, predominantly by competing with guanosine to cause transition mutations^{11,22–25}. However, two separate studies support an antiviral effect mediated by chain termination, with incorporation of either a single or two consecutive T-705 molecules blocking further extension by influenza polymerase^{26,27}.

For the nsp12 complex, omission of ATP and/or GTP from elongation reactions results in rapid incorporation of T-705 and T-1105 at multiple sites with both substrates (Fig. 2b,c). In reactions with 10 μ M T-1105-RTP, multiple incorporation events are seen within 10 seconds, while T-705-RTP is less efficient. Neither compound is incorporated in the place of cytosine or uracil, clearly showing they function as purine analogues (Extended data Fig. 5). For the processive HP complex, efficient incorporation and elongation occurs opposite both uracil and cytosine, resulting in rapid accumulation of full-length products, advocating for lethal mutagenesis as the MoA for these compounds (Fig. 2c). Interestingly, the PT substrate reveals a striking difference in the MoA depending on whether the analogues are incorporated in place of guanine or adenine. Opposite uracil, both are rapidly incorporated but further extension is slow and inefficient, leading to an accumulation of RNA products consistent with a non-obligate chain

115 termination mechanism. The presence of a non-natural 3' nucleotide analogue may destabilize
the complex and promote template dissociation. In contrast, opposite cytosine there is a stall
before each analogue incorporation step, but elongation past the analogue is rapid, even at
consecutive sites. More full-length product is observed in these no-GTP experiments, suggesting
120 both compounds are more efficient as guanosine analogues. A similar observation has been made
for the poliovirus RdRp that efficiently incorporated and bypassed T-1105 opposite cytosine but
was prone to pausing opposite uridine. These pause events were attributed to RdRp backtracking,
which was therefore assigned the primary cause of the inhibitory activity²⁸. We conclude that for
nsp12 the T-705/T-1105 MoA is a combination of chain termination, slowed RNA synthesis and
lethal mutagenesis, dictated by the structural and functional properties of both the polymerase
125 and the RNA.

The reactions performed without ATP also showed detectable amounts of GTP:U
mismatch products, allowing us to make a direct comparison of the T-1105 incorporation rate
with that of a natural GTP:U mismatch (Extended data Fig. 6). Reactions using 50 μ M GTP and
1 μ M T-1105-RTP show ~5-fold more T-1105:U produced relative to GTP:U. Considering the
130 concentration difference, T-1105 incorporation is calculated to be ~250-fold more efficient than
the native GTP mismatch, the most common naturally occurring transition mutation. These high
incorporation rates likely overwhelm the CoV error-correcting mechanisms, consistent with our
infectious virus data, although it remains to be determined if nsp14 can excise T-705/T-1105, as
is the case for ribavirin and 5-fluorouracil^{4,6,29}.

135 We carried out pre-steady state rapid-quench experiments to further understand the
molecular basis of nsp12 elongation rates and fidelity. We initially attempted to form stalled
elongation complexes, as is commonly done for the structurally related picornaviral RdRps³⁰⁻³²
but nsp12 complexes showed half-lives of only ~1 minute across a range of conditions and
cofactor stoichiometries (Extended data Fig. 7). The significant rapid burst phase, however,
140 allowed characterization of elongation rates under single-turnover conditions using millisecond
timescale EDTA quench-flow (Fig. 3 and Extended data Fig. 8). Experiments were performed at
three NTP concentrations based on estimated steady-state K_d values of 3-15 μ M. CTP was
omitted to allow elongation by only 7 and 8 nucleotides on PT and HP substrates respectively,
while avoiding end-effects that can slow elongation rates. At 17 μ M NTP we observe multiple
145 incorporation steps within a mere 10-20 ms and formation of the +7/8 products by 100 ms (Fig.
3). Analysis of the data using a minimal model of identical nucleotide incorporation steps yields
elongation rates of 66 ± 1.3 and 95 ± 8 nt/s for the PT and HP substrates at 50 μ M NTP at 22 °C
(Extended data Table 1). Fitting the rate concentration dependence yields maximal elongation
rates of 90 ± 4 nt/s on annealed primer-templates and 150 ± 30 nt/s on hairpin templates (Fig. 3d).

150 These data reveal that the SARS-CoV nsp12 is the fastest viral RdRp known, with rates
significantly faster than the 5-20 nt/s observed for picornaviral polymerases at room
temperature³²⁻³⁴ and 4-18 nt/s for hepatitis C and dengue virus polymerases at 30 and 37° C^{35,36}.
Based on its structure, nsp12 is expected to use the palm domain based active site closure

mechanism that is unique to viral RdRps³¹ and which exhibits a 3-to-5 fold rate increase between 22 °C and 37 °C^{33,34}, suggesting nsp12 can elongate at a staggering 600-700 nt/s at physiological temperatures.

Such a fast viral RdRp is consistent with the need to rapidly replicate ~30,000 nt long RNA genomes, but raises questions as to how fidelity of nucleotide incorporation is impacted. In our data, we consistently observe nucleotide misincorporations and subsequent elongation on templates where nsp12 should stall due to the lack of CTP (Extended data Fig. 2b, 3, 9). In contrast, RdRps that form highly stable elongation complexes show limited or no read-through products under comparable conditions^{30,35,37}. The efficiency of the nsp12 bypass is UTP concentration dependent (Extended data Fig. 3), indicative of uracil misincorporation opposite a templating guanosine (UTP:G). The UTP:G elongation product is also observed in the PT quench-flow data at reaction times as short as 100 ms, yielding a misincorporation rate of ~0.2 nt/s (Extended data Table 1). This is only 400–500-fold lower than the cognate UTP:A rate measured in the same experiment, more than one order of magnitude less accurate than the generally admitted 10⁻⁴-10⁻⁶ error rate of viral RdRps³⁸.

The molecular basis for fast and low fidelity replication by nsp12 is not yet known, but a comparison of RdRp structures reveals that a key NTP interaction is absent in CoV enzymes (Fig. 4). Viral RdRps use an electrostatic interaction with an arginine residue in motif F to position the NTP during catalysis³¹. For most positive strand RNA virus polymerases, this arginine is stabilized by a salt bridge to a glutamic acid residue, also from motif F. Notably, the CoV nsp12 has an alanine in place of this glutamate (A547) and as a result, the arginine (R555) is not rigidly anchored above the active site (Fig. 4a)^{31,39}. This flexibility could allow catalysis to occur with a relaxed triphosphate positioning, decreasing fidelity by lowering the requirement for strict Watson-Crick base pairing in the active site. Interestingly, the CoV and other large-genome nidoviral polymerases also have an SDD rather than a more typical GDD sequence in the palm domain of motif C (Fig. 4b). This serine residue can hydrogen bond with the 2' hydroxyl of the priming nucleotide, theoretically stabilizing the position of the primer in the active site and compensating for the less rigid interactions with the bound NTP. These unique features have likely played a central role in genome expansion and stability by providing a fast RNA synthesis machinery whose inaccuracy is mitigated by the presence of an RNA repair exonuclease. Our data demonstrate that nucleoside analogues are pertinent candidates for the treatment of COVID-19. Favipiravir, with its already defined safety profile and mode of action, may well find a place as an anti-RdRp component in combination therapies targeting coronaviruses.

References and Notes:

1. Snijder, E. J. *et al.* Unique and Conserved Features of Genome and Proteome of SARS-coronavirus, an Early Split-off From the Coronavirus Group 2 Lineage. *J. Mol. Biol.* **331**, 991–1004 (2003).
2. Bouvet, M. *et al.* RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proc. Natl. Acad. Sci.* **109**, 9372 (2012).
3. Minskaia, E. *et al.* Discovery of an RNA virus 3'→5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5108 (2006).
4. Eckerle, L. D., Lu, X., Sperry, S. M., Choi, L. & Denison, M. R. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J. Virol.* **81**, 12135–12144 (2007).
5. Abdelnabi, R. *et al.* Understanding the Mechanism of the Broad-Spectrum Antiviral Activity of Favipiravir (T-705): Key Role of the F1 Motif of the Viral Polymerase. *J. Virol.* **91**, e00487-17 (2017).
6. Ferron, F. *et al.* Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *Proc. Natl. Acad. Sci.* **115**, E162–E171 (2018).
7. Gordon, C. J., Tchesnokov, E. P., Feng, J. Y., Porter, D. P. & Götte, M. The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J. Biol. Chem.* **295**, 4773–4779 (2020).
8. Gordon, C. J. *et al.* Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. *J. Biol. Chem.* jbc.RA120.013679 (2020) doi:10.1074/jbc.RA120.013679.
9. Pruijssers, A. J. & Denison, M. R. Nucleoside analogues for the treatment of coronavirus infections. *Curr. Opin. Virol.* **35**, 57–62 (2019).
10. Delang, L., Abdelnabi, R. & Neyts, J. Favipiravir as a potential countermeasure against neglected and emerging RNA viruses. *Antiviral Res.* **153**, 85–94 (2018).
11. Goldhill, D. H. *et al.* Determining the Mutation Bias of Favipiravir in Influenza Virus Using Next-Generation Sequencing. *J. Virol.* **93**, e01217-18, /jvi/93/2/JVI.01217-18.atom (2018).
12. Guedj, J. *et al.* Antiviral efficacy of favipiravir against Ebola virus: A translational study in cynomolgus macaques. *PLOS Med.* **15**, e1002535 (2018).

13. te Velthuis, A. J. W., Arnold, J. J., Cameron, C. E., van den Worm, S. H. E. & Snijder, E. J. The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Res.* **38**, 203–214 (2010).
14. Cheng, A. *et al.* Expression, purification, and characterization of SARS coronavirus RNA polymerase. *Virology* **335**, 165–176 (2005).
15. Subissi, L. *et al.* One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proc. Natl. Acad. Sci.* **111**, E3900–E3909 (2014).
16. Kirchdoerfer, R. N. & Ward, A. B. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat. Commun.* **10**, 2342 (2019).
17. Gao, Y. *et al.* Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* eabb7498 (2020) doi:10.1126/science.abb7498.
18. Sola, I., Almazán, F., Zúñiga, S. & Enjuanes, L. Continuous and Discontinuous RNA Synthesis in Coronaviruses. *Annu. Rev. Virol.* **2**, 265–288 (2015).
19. Huchting, J., Winkler, M., Nasser, H. & Meier, C. Synthesis of T-705-Ribonucleoside and T-705-Ribonucleotide and Studies of Chemical Stability. *ChemMedChem* **12**, 652–659 (2017).
20. Huchting, J., Vanderlinden, E., Van Berwaer, R., Meier, C. & Naesens, L. Cell line-dependent activation and antiviral activity of T-1105, the non-fluorinated analogue of T-705 (favipiravir). *Antiviral Res.* **167**, 1–5 (2019).
21. Huchting, J. *et al.* Prodrugs of the Phosphoribosylated Forms of Hydroxypyrazinecarboxamide Pseudobase T-705 and Its De-Fluoro Analogue T-1105 as Potent Influenza Virus Inhibitors. *J. Med. Chem.* **61**, 6193–6210 (2018).
22. Baranovich, T. *et al.* T-705 (Favipiravir) Induces Lethal Mutagenesis in Influenza A H1N1 Viruses *In Vitro*. *J. Virol.* **87**, 3741 (2013).
23. Escribano-Romero, E., Jiménez de Oya, N., Domingo, E. & Saiz, J. C. Extinction of West Nile Virus by Favipiravir through Lethal Mutagenesis. *Antimicrob. Agents Chemother.* **61**, e01400-17 (2017).
24. de Ávila, A. I. *et al.* Lethal Mutagenesis of Hepatitis C Virus Induced by Favipiravir. *PloS One* **11**, e0164691–e0164691 (2016).
25. Arias, A., Thorne, L. & Goodfellow, I. Favipiravir elicits antiviral mutagenesis during virus replication in vivo.

26. Sangawa, H. *et al.* Mechanism of Action of T-705 Ribosyl Triphosphate against Influenza Virus RNA Polymerase. *Antimicrob. Agents Chemother.* **57**, 5202 (2013).
27. Jin, Z., Smith, L. K., Rajwanshi, V. K., Kim, B. & Deval, J. The Ambiguous Base-Pairing and High Substrate Efficiency of T-705 (Favipiravir) Ribofuranosyl 5'-Triphosphate towards Influenza A Virus Polymerase. *PLOS ONE* **8**, 10 (2013).
28. Dulin, D. *et al.* Signatures of Nucleotide Analog Incorporation by an RNA-Dependent RNA Polymerase Revealed Using High-Throughput Magnetic Tweezers. *Cell Rep.* **21**, 1063–1076 (2017).
29. Smith, E. C., Blanc, H., Vignuzzi, M. & Denison, M. R. Coronaviruses Lacking Exoribonuclease Activity Are Susceptible to Lethal Mutagenesis: Evidence for Proofreading and Potential Therapeutics. *PLoS Pathog.* **9**, e1003565 (2013).
30. Arnold, J. J. & Cameron, C. E. Poliovirus RNA-dependent RNA Polymerase (3Dpol): Assembly Of Stable, Elongation-Competent Complexes By Using A Symmetrical Primer-Template Substrate (sym/sub). *J. Biol. Chem.* **275**, 5329–5336 (2000).
31. Gong, P. & Peersen, O. B. Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proc. Natl. Acad. Sci.* **107**, 22505–22510 (2010).
32. Shi, W. *et al.* A nucleobase-binding pocket in a viral RNA-dependent RNA polymerase contributes to elongation complex stability. *Nucleic Acids Res.* **48**, 1392–1405 (2020).
33. Gong, P., Campagnola, G. & Peersen, O. B. A quantitative stopped-flow fluorescence assay for measuring polymerase elongation rates. *Anal. Biochem.* **391**, 45–55 (2009).
34. Seifert, M., van Nies, P., Cameron, C. E., Depken, M. & Dulin, D. Temperature controlled high-throughput magnetic tweezers show striking difference in activation energies of replicating viral RNA-dependent RNA polymerases. 12.
35. Jin, Z., Leveque, V., Ma, H., Johnson, K. A. & Klumpp, K. Assembly, Purification, and Pre-steady-state Kinetic Analysis of Active RNA-dependent RNA Polymerase Elongation Complex. *J. Biol. Chem.* **287**, 10674–10683 (2012).

36. Jin, Z., Deval, J., Johnson, K. A. & Swinney, D. C. Characterization of the Elongation Complex of Dengue Virus RNA Polymerase: ASSEMBLY, KINETICS OF NUCLEOTIDE INCORPORATION, AND FIDELITY. *J. Biol. Chem.* **286**, 2067–2077 (2011).
37. Campagnola, G., McDonald, S., Beaucourt, S., Vignuzzi, M. & Peersen, O. B. Structure-Function Relationships Underlying the Replication Fidelity of Viral RNA-Dependent RNA Polymerases. *J. Virol.* **89**, 275–286 (2015).
38. Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M. & Belshaw, R. Viral Mutation Rates. *J. Virol.* **84**, 9733–9748 (2010).
39. Peersen, O. A Comprehensive Superposition of Viral Polymerase Structures. *Viruses* **11**, 745 (2019).
40. Tvarogová, J. *et al.* Identification and Characterization of a Human Coronavirus 229E Nonstructural Protein 8-Associated RNA 3'-Terminal Adenylyltransferase Activity. *J. Virol.* **93**, e00291-19, /jvi/93/12/JVI.00291-19.atom (2019).
41. Xiao, Y. *et al.* Nonstructural Proteins 7 and 8 of Feline Coronavirus Form a 2:1 Heterotrimer That Exhibits Primer-Independent RNA Polymerase Activity. *J. Virol.* **86**, 4444–4454 (2012).
42. te Velthuis, A. J. W., van den Worm, S. H. E. & Snijder, E. J. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Res.* **40**, 1737–1747 (2012).
43. Imbert, I. *et al.* A second, non-canonical RNA-dependent RNA polymerase in SARS Coronavirus. *EMBO J.* **25**, 4933–4942 (2006).
44. Johnson, K. A., Simpson, Z. B. & Blom, T. Global Kinetic Explorer: A new computer program for dynamic simulation and fitting of kinetic data. *Anal. Biochem.* **387**, 20–29 (2009).
45. Touret, F. *et al.* In vitro screening of a FDA approved chemical library reveals potential inhibitors of SARS-CoV-2 replication. *bioRxiv* 2020.04.03.023846 (2020) doi:10.1101/2020.04.03.023846.
46. Touret, F. *et al.* Phylogenetically based establishment of a dengue virus panel, representing all available genotypes, as a tool in dengue drug discovery. *Antiviral Res.* **168**, 109–113 (2019).
47. Delang, L. *et al.* Mutations in the chikungunya virus non-structural proteins cause resistance to favipiravir (T-705), a broad-spectrum antiviral. *J. Antimicrob. Chemother.* **69**, 2770–2784 (2014).

Acknowledgments: We thank Magali Gilles for excellent technical support as well as Prs. C. Drosten and F. Drexler for providing the SARS-CoV-2 through EVA-GLOBAL (European Union's Horizon 2020 programme, GA 871029). This work was supported by the Fondation pour la Recherche Médicale (Aide aux équipes), the SCORE project H2020 SC1-PHE-Coronavirus-2020 (grant#101003627) to BCa, Inserm through the REACTing initiative (REsearch and ACTION targeting emerging infectious diseases) to BCa and BCo, National Institutes of Health grant AI059130 to OP, and a grant from DZIF (German Center for Infection Research) to JH and CM.

Author contributions: BS/NTTL/VF performed experiments and analyzed data, JH/CM performed experiments, FT/GP/BCo performed experiments and analyzed data, FF/ED analyzed data, AS/OP/BCa designed experiment, performed experiments, analyzed data and wrote manuscript.

Competing interests: Authors declare no competing interests.

Materials & Correspondence: Correspondence to: either Olve.Peersen@colostate.edu or bruno.canard@afmb.univ-mrs.fr

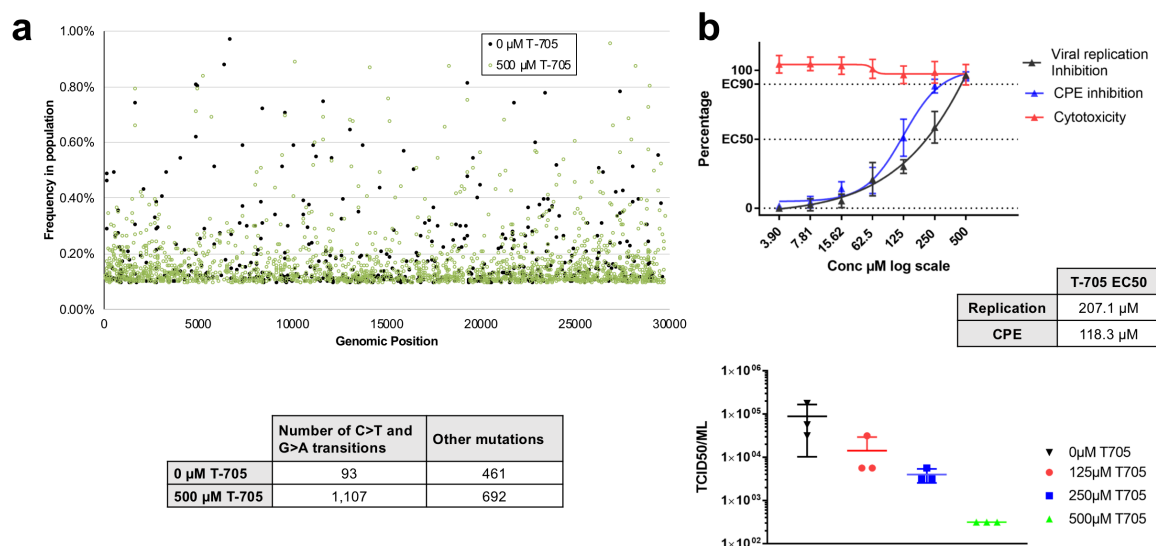


Fig. 1: Antiviral effects of T-705 on SARS-CoV-2. **a**, *In vitro* effects of T-705 on SARS-CoV-2. Distribution of the mutations along the SARS-CoV-2 genome and number of mutations observed in presence or absence of T-705. A 3-fold increase in the presence of the drug is observed ($P < 0.001$, Pearson's Chi-squared test with Yates' continuity correction). **b**, Quantification of the antiviral effect of T-705 by genome copy number, virus-mediated CPE, and number of infectious particles.

12

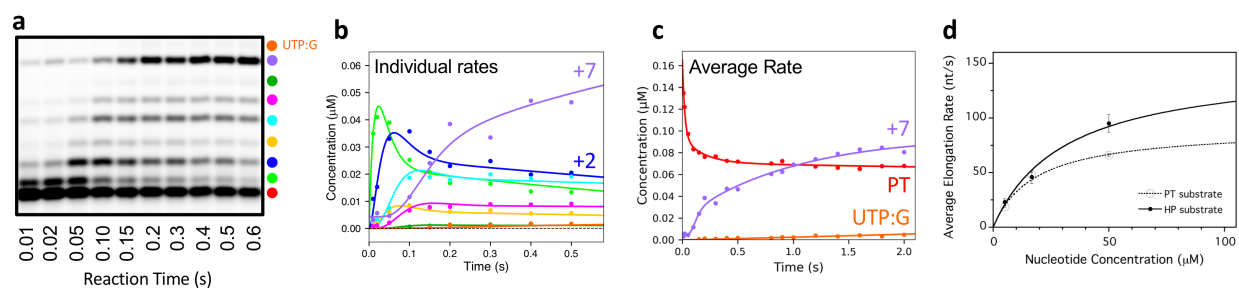


Fig. 3. Pre steady-state elongation by nsp12-7L8-8 complex. **a**, EDTA quench-flow data showing rapid elongation to +7 products on the PT10/20 substrate in the absence of CTP. **b,c**, Fitting of the quench-flow data to models with discrete rates for each elongation step (**b**, color key in **a**) or with a single average incorporation rate for each step (**c**). **d**, Nucleotide concentration dependence yielding maximal average elongation rates of 150 ± 30 nt/s on the HP substrate and 91 ± 4 nt/s on the PT substrate.

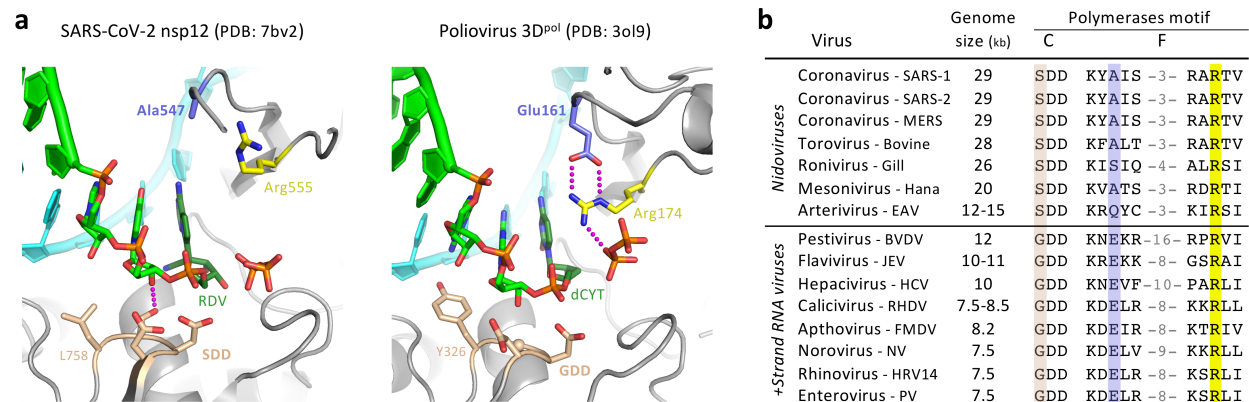


Fig. 4. Comparison of viral polymerase active site structures and sequences. **a**, Analogous post-catalysis, pre-translocation structures of CoV nsp12 and poliovirus 3D^{pol} after incorporation of remdesivir (RDV) and deoxycytosine (dCYT), respectively. The CoV polymerases have replaced a motif F glutamate residue (Glu161) with an alanine (Ala547), removing a highly conserved interaction that positions the motif F arginine for interactions with the NTP and pyrophosphate. **b**, Alignment of representative viral RdRp sequences showing the large genome nidoviruses have a SDD instead of GDD sequence in the palm domain motif C (tan) and alanine, serine, or glutamine in place of the aforementioned glutamate (blue) found in other positive strand RNA viruses. The NTP-interacting arginine (yellow) is conserved, but the overall length of the motif F loop is shorter in the nidoviruses (numbers reflect omitted residues).

Materials and Methods

Nucleotide substrates

T-1105-ribose-TP (RTP), the 5'-triphosphate of the non-fluorinated derivative of T-705-ribose, was synthesized as described previously²¹ following the iterative route from mono- via di- to triphosphate with minor modifications. To overcome low solubility of T-1105-ribonucleotides in organic solvent, the 2'- and 3'-acetyl groups were kept until the final TP-synthesis and were then cleaved by treatment with base. This reaction sequence resulted in an increased yield of up to 90% (for di-) and 46% (for triphosphate), nearly double of previously reported. T-705-RTP was obtained from Toronto Research Chemicals. Lyophilized aliquots were resuspended in TE buffer pH 7.0 and used the same day. Stability can be verified by measuring Abs at 372 nm and 350 nm for T-705-RTP and T-1105-RTP respectively. Other NTPs were purchased from GE Healthcare.

Synthetic Oligonucleotides

Primer-template (PT) pairs were purchased from Biomers (HPLC grade). RNA oligonucleotides ST20 (20-mer) and LS1 (40-mer), corresponding to the 3'-end of the SARS-CoV genome (excluding the poly A tail) were used as templates, annealed to a 5' cy5 labeled SP10 primer (10-mer) corresponding to the 5'-end of the anti-genome. For simplicity, these were named PT10/20 and PT10/40 throughout the manuscript. Annealing was performed by denaturing primer:template pairs (molar ratio of 1:1.5) in 110 mM KCl at 70° C for 10 min, then cooling slowly to room temperature over several hours. Hairpin RNAs were synthesized by Integrated DNA Technologies (Coralville, IA), resuspended at 100 uM concentration in 50 mM NaCl, 5 mM MgCl₂, 10 mM Tris (pH 8.0) and heated to 95°C for 15 minutes before snap cooling on ice.

Expression and Purification of SARS-CoV proteins

All SARS-CoV proteins used in this study were expressed in *Escherichia coli* (*E. coli*), under the control of T5 promoters. Cofactors nsp7L8 and nsp8 alone were expressed from pQE30 vectors with C-terminal and N-terminal hexa-histidine tags respectively. TEV cleavage site sequences were included for His-tag removal following expression. The nsp7L8 fusion protein was generated by inserting a GSGSGS linker between nsp7- and nsp8-coding sequences. Cofactors were expressed in NEB Express C2523 (New England Biolabs) cells carrying the pRare2LacI (Novagen) plasmid in the presence of Ampicillin (100 µM/mL) and Chloramphenicol (17 µg/mL). Protein expression was induced with 100 µM IPTG once the OD₆₀₀ = 0.5–0.6, and expressed overnight at 17°C. Protein was purified first through affinity chromatography with HisPur Cobalt resin (Thermo Scientific), with a lysis buffer containing 50 mM Tris-HCl pH 8, 300 mM NaCl, 10 mM Imidazole, supplemented with 20 mM MgSO₄, 0.25 mg/mL Lysozyme, 10 µg/mL DNase, 1 mM PMSF, with lysis buffer supplemented with 250 mM imidazole. Eluted protein was concentrated and dialyzed overnight in the presence of histidine labeled TEV protease (1:10 w/w ratio to TEV:protein) for removal of the protein tag. Cleaved protein was purified through a second cobalt column and protein was purified through size exclusion chromatography (GE, Superdex S200) in gel filtration buffer (25 mM HEPES pH 8, 150 mM NaCl, 5mM MgCl₂, 5 mM TCEP). Concentrated aliquots of protein were flash-frozen in liquid nitrogen and stored at -80 °C. A synthetic, codon-optimized SARS-CoV nsp12 gene (DNA 2.0) bearing C-terminal 8His-tag preceded by a TEV protease cleavage site was expressed from a pJ404 vector (DNA 2.0) in *E. coli* strain BL21/pG-Tf2 (Takara). Cells were grown at

37°C in the presence of Ampicillin and Chloramphenicol until OD₆₀₀ reached 2. Cultures were induced with 250 µM IPTG and protein expressed at 17°C overnight. Purification was performed as above in lysis buffer supplemented with 1% CHAPS. Two additional wash steps were performed prior to elution, with buffer supplemented with 20 mM imidazole and 50 mM arginine for the first and second washes respectively. Polymerase was eluted using lysis buffer with 500 mM imidazole and concentrated protein was purified through gel filtration chromatography (GE, Superdex S200) in the same buffer as for nsp7L8. Collected fractions were concentrated and supplemented with 50% glycerol final concentration and stored at -20 °C.

Steady-state elongation complex reactions

Nsp12, nsp7L8 and nsp8 were mixed just prior to each experiment at a 1:3:3 molar ratio (unless otherwise stated) and preincubated on ice for 10 mins. The protein complex was subsequently mixed with the RNA pre-mix (20 mM HEPES pH 7.5, 50 mM NaCl, 5 mM MgCl₂) containing either a single RNA substrate, or both HP and PT RNAs at equimolar ratios. Reactions were initiated with 50 µM (final concentration) of all four NTPs, or without CTP for partial elongation reactions. Final reaction concentrations were 1 µM nsp12, 0.2 µM RNA. Reactions were quenched at indicated timepoints with 5X volume of FBD stop solution (formamide, 10 mM EDTA). To verify that activity was not due to either nsp7L8 or nsp8, which have been shown to contain primase-like noncanonical RdRp activities⁴⁰⁻⁴³ assays were run in the absence of nsp12. The NTP K_d was estimated using the same setup described here, but with final (equimolar) concentration of ATP, GTP and UTP ranging from 0.78 – 100 µM (2-fold dilution series). The time course of product formation was fit to single exponential equation for each concentration of NTP to give the observed rate constant (k_{obs}). Observed rates were subsequently plotted against NTP concentration, and the data was fit via hyperbolic regression to give the equilibrium dissociation constant (K_d) and the maximum rate constant for incorporation of NTPs (k_{pol}).

Formation of stalled elongation complex

Attempts to form a stalled elongation complex were performed with HP8-22 RNA with varied ratios of nsp12:nsp7L8:nsp8 (using a constant nsp12 concentration 0.8 µM). Protein, HP8-22 RNA (0.2 µM) and 50 µM final concentrations of GTP and ATP were incubated for 5 min at room temperature to allow formation of a +4-lock stalled complex. Reactions were diluted 1:2 in high salt to prevent RNA rebinding, and chased at various timepoints (1 – 50 minutes) with 50 µM all NTPs. Chase reactions were quenched after 30 sec in FBD. Stability half-life was calculated from the ratio of full-length (FL) product produced relative to the total amount of +4-lock + FL product at each timepoint. Half-life was obtained by fitting the data through a single exponential.

Pre-steady state quenched-flow kinetics

Experiments were carried out in a Bio-Logic QFM-300 rapid chemical quench-flow apparatus that controls reaction time by flow rate through a 2.5 uL chamber between the reaction mixer and the quencher mixer. The nsp12-nsp7L8-nsp8 complex was assembled by first preincubating the proteins on ice for 10 minutes at a 1:3:3 molar ratio, then adding this to SP10/20 and HP8-29 hairpin RNA in reaction buffer (20 mM HEPES pH 7.5, 5mM MgCl₂, 50 mM NaCl) and further incubated at room temperature for 15 minutes prior to beginning the rapid-quench experiments, which took ~10 minutes to complete a set of 22 different reaction

times. Reactions were initiated by mixing equal volumes (18 μ l each) of the protein-RNA complex with NTP solutions containing equimolar concentrations of ATP, GTP and UTP, and then quenched with 18 μ l of 200 mM EDTA. Final concentrations in the reactions were 1 μ M nsp12, 0.2 μ M primer-template RNA, 0.2 μ M hairpin RNA, and 50, 16.7 or 5.6 μ M NTP, and the final EDTA concentration after quenching was 67 mM.

Product Analysis

Quenched reactions were mixed 1:1 with FBD loading dye and heated for 10 min at 70°C, and cooled on ice for 2 min before analysis on 17 – 20% polyacrylamide, 7 M Urea TBE gels. Gels were run at a constant 65 W using Sequi-Gen GT Systems from Bio-Rad or vertical electrophoresis systems from CBS Scientific and visualized using an Amersham™ Typhoon™ Biomolecular Imager (GE Healthcare). The intensity of each band was quantified using the ImageQuant software (GE Healthcare)/Image Gauge (Fuji) and/or using ImageJ as implemented in the Fiji package, with background subtraction. Product yield was determined by dividing the intensity of the product by total intensity of the product + remaining primer, and multiplying by the input concentration of RNA. For the rapid quench data, the program PeakFit (Systat Software) was used to fit gel lane profiles to a set of gaussian peaks and the fractional area contained within each peak was multiplied by the RNA concentration in the experiment to calculate the amount of each elongation species as a function of reaction time.

Pre-steady state kinetic data analysis

Rapid quench data were analysed using the program KinTek Explorer⁴⁴ to model the reactions as a series of seven (PT substrate) or eight (HP substrate) irreversible nucleotide addition steps. Data from each NTP concentration series were fit independently (not globally) to obtain observed rates using either a model with i) a single average rate or ii) individual rates for each nucleotide addition step (fig S8). Modeling the remaining intermediate species due to low processivity elongation using formal RNA dissociation and rebinding steps (at rates comparable to non-burst phase primer utilization) was not successful, suggesting the nsp12-RNA complex exists in some form of inactive state. Low processivity elongation was instead accommodated in the kinetic model by adding a reversible inactivation/reactivation equilibrium step for each elongation product, allowing intermediate products to depart from the immediate processive pathway toward full length product formation, but then rejoin the pathway for further elongation. The rate constants for these steps were shared among all intermediates species as there was insufficient data to fit them individually.

The single average rate models were fit using data from primer loss and the expected full-length (+7 or +8) product, but not the intermediate species. Note that for the PT template we observed detectable amounts of additional +8 and +9 bands at the higher 16.7 and 50 μ M NTP concentrations and therefore included them in the model. This is presumably due to a UTP:G mismatch to yield the +8 product in the absence of CTP, followed by a cognate ATP:U addition that is slow because it is priming on a mismatched base pair. For the HP substrate, we were unable to definitively identify the gel migration bands for the +6 and +7 species and therefore did not model individual rates for these incorporation events. Data from the HP RNAs are challenging to analyze because the high thermodynamic stability of the folded RNA structure can make it difficult to fully denature the helix during electrophoresis. This leads to a mixture of species that transition from denatured single stranded RNA for the initial species to more

compact and faster migrating duplexes for longer elongation products. We were therefore unable to reliably analyze quench flow data from longer hairpin products.

NTP-analog incorporation

Enzyme mix (20 μ M nsp12, 60 nsp7L8, 60 μ M nsp8) in complex buffer (25 mM HEPES pH 7.5, 150 NaCl, 5 mM TCEP, 5 mM $MgCl_2$) was incubated 10 min on ice and then diluted with reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$) to 4 μ M nsp12 (12 nsp7L8 and 8). The resulting enzyme complex was mixed with the same volume of 0.8 μ M primer/ template (P/T) with or without 0.8 μ M hairpin (HP) carrying 6-FAM or Cy5 fluorescent labels at their 5' ends in reaction buffer, and incubated at 25°C (or 30°C as given in Figure legends) for 10 min. Reactions were then started with the same volume of 100 μ M NTPs (leaving out one or two as given in Figure legends) and TE buffer or NTP-analogs (T-1105-RTP or T-705-RTP) in reaction buffer. Final concentrations in the reactions were 1 μ M nsp12 (3 μ M nsp7L8 and 8), 0.2 μ M P/T (and 0.2 μ M HP), 50 μ M NTPs and the given concentrations of NTP-analogs. Samples of 8 μ l were taken at given time points and mixed with 40 μ l of formamide containing 10 mM EDTA. Ten- μ l samples were analyzed by denaturing PAGE (20 % acrylamide, 7 M urea, TBE buffer); and product profiles visualized by a fluorescence imager (Amersham Typhoon).

In vitro infection assays

Cell line: VeroE6 (ATCC CRL-1586) cells were grown in minimal essential medium (Life Technologies) with 7.5% heat-inactivated fetal calf serum (FCS), at 37 °C with 5 % CO_2 with 1 % penicillin/streptomycin (PS, 5000 U.mL⁻¹ and 5000 μ g.mL⁻¹ respectively; Life Technologies) and supplemented with 1 % non-essential amino acids (Life Technologies).

Virus strain: SARS-CoV-2 strain BavPat1 was obtained from Pr. Drosten through EVA GLOBAL (<https://www.european-virus-archive.com/>). Virus stock was prepared as previously described⁴⁵. All experiments were conducted in a BSL3 laboratory.

Antiviral experiments: For EC50 and CC50 determination, one day prior to infection, 5×10^4 VeroE6 cells were seeded in 100 μ L assay medium (containing 2.5 % FCS) in 96 well plates. The next day, seven 2-fold serial dilutions of T-705 (500 μ M to 3.9 μ M in triplicate) were added to the cells (25 μ L/well, in assay medium). Four virus control wells were supplemented with 25 μ L of assay medium. After 15 min, 25 μ L of a virus mix diluted in medium was added to the wells. The amount of virus working stock used was calibrated prior to the assay, based on a replication kinetics, so that the replication growth is still in the exponential growth phase for the readout as previously described^{46,47}. Four cell control wells (*i.e.* with no virus) were supplemented with 50 μ L of assay medium. Plates were incubated for 2 days at 37 °C prior to quantification of the viral genome by real-time RT-PCR. RNA extraction and viral RNA quantification was performed as previously described⁴⁵. The 50% and 90% effective concentrations (EC50, EC90; compound concentration required to inhibit viral RNA replication by 50% and 90%) were determined using logarithmic interpolation as previously described⁴⁶. For the evaluation of CC50 (the concentration that reduces the total cell number by 50%), the same culture conditions were set as for the determination of the EC50, without addition of the virus, then cell viability was measured using CellTiter Blue® (Promega). Cell supernatant media were discarded and CellTiter-Blue® reagent (Promega) was added following the manufacturer's instructions. Plates were incubated for 2 hours prior recording fluorescence (560/590nm) with a Tecan Infinite 200Pro machine. From the measured OD₅₉₀, CC50 was determined using

logarithmic interpolation. For EC50 determination using CPE inhibition, cells and T-705 were prepared as previously described (EC50 determination). Eight virus control wells were supplemented with 25 μ L of assay medium and eight cell control were supplemented with 50 μ L. After 15 min, 25 μ L of a virus mix diluted in 2.5 % FCS-containing medium was added to the wells at MOI 0.002. Three days after infection, CPE were assessed using CellTiter-Blue® reagent (Promega). For the infectivity test cells and compound were prepared as described above (EC50 determination) with only three 2-fold serial dilutions of T-705 (125 μ M to 500 μ M, in triplicate). Three virus control wells were supplemented with 25 μ L of assay medium. The experiment was conducted as previously described for the EC50 determination. At day 3 the supernatant was collected and each triplicate was titrated by measuring the 50 % tissue culture infectivity dose (TCID₅₀); briefly, 3 replicates were infected with 100 μ L of tenfold serial dilutions of the previous supernatant, and incubated for 3 days. CPE was measured by CellTiter-Blue® reagent (Promega) and TCID₅₀ was calculated and expressed as TCID₅₀/mL based on the Reed and Muench (43). Antiviral experiments data were analyzed using GraphPad Prism 7 software (Graph pad software). Graphical representations were also performed on GraphPad Prism 7 software.

Sequence analysis

Eight overlapping amplicons were produced from the extracted viral RNA using the SuperScript IV One-Step RT-PCR System (Thermo Fisher Scientific) and specific primers (supplemental data). PCR products were pooled at equimolar proportions. After Qubit quantification using Qubit® dsDNA HS Assay Kit and Qubit 2.0 fluorometer (ThermoFisher Scientific) amplicons were fragmented by sonication in ~200bp long fragments. Libraries were built adding barcodes for sample identification to the fragmented DNA using AB Library Builder System (ThermoFisher Scientific). To pool the barcoded samples at equimolar ratio a quantification step by real time PCR using Ion Library TaqMan™ Quantitation Kit (Thermo Fisher Scientific) was realised. An emulsion PCR of the pools and loading on 530 chip was realised using the automated Ion Chef instrument (ThermoFisher). Sequencing was performed using the S5 Ion torrent technology v5.12 (Thermo Fisher Scientific) following manufacturer's instructions. Consensus sequence was obtained after trimming of reads (reads with quality score <0.99, and length <100pb were removed and the 30 first and 30 last nucleotides were removed from the reads) mapping of the reads on a reference (determined following Blast of De Novo contigs) using CLC genomics workbench software v.20 (Qiagen). A *de novo* contig was also produced to ensure that the consensus sequence was not affected by the reference sequence. Quasi species with frequency over 0.1% were studied.