bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Structure-function analysis of the nsp14 N7-guanine methyltransferase reveals an essential role in *Betacoronavirus* replication

Natacha S. Ogando[1#], Priscila El Kazzi[2#], Jessika C. Zevenhoven-Dobbe[1], Brenda W. Bontes[1], Alice Decombe[2], Clara C. Posthuma[1], Volker Thiel[3,4], Bruno Canard[2], François Ferron[2,3#], Etienne Decroly[2£#] & Eric J. Snijder[1*#]

[1] Department of Medical Microbiology, Leiden University Medical Center, Leiden, The Netherlands, 2333ZA

[2] Architecture et Fonction des Macromolécules Biologiques, Centre National de la Recherche Scientifique, Aix-Marseille Université, Marseille, France, 13288

[3] European Virus Bioinformatics Center, Jena, Germany, 07743.

[4] Institute of Virology and Immunology (IVI), Bern, Switzerland,3350.

[5] Department of Infectious Diseases and Pathobiology, Vetsuisse Faculty, University of Bern, Bern, Switzerland, 3012.

[#] Authors contributed equally

* Eric J. Snijder; **Email:** E.J.Snijder@lumc.nl

[£] Etienne Decroly; **Email:** Etienne.decroly@univ-amu.fr

Author Contributions: N.O., P.E.K., J.C.Z., B.B., A.D., and F.F. performed experimental work. N.O., P.E.K., F.F., J.C.Z., B.B., B.C., E.D., and E.J.S. analyzed the data and formulated research hypothesis. C.C.P., V.T., B.C., E.D., and E.J.S. supervised research and contributed in scientific discussions. N.O., P.E.K., F.F., E.D., and E.J.S. wrote the paper. All authors approved the version submitted.

Competing Interest Statement: All authors declare no competing interests.

Classification: Biological Sciences, Microbiology

Keywords: SARS-CoV, MERS-CoV, SARS-CoV-2, RNA synthesis, mRNA capping

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

As coronaviruses (CoVs) replicate in the host cell cytoplasm, they rely on their own capping machinery to ensure the efficient translation of their mRNAs, protect them from degradation by cellular 5' exoribonucleases, and escape innate immune sensing. The CoV nonstructural protein 14 (nsp14) is a bi-functional replicase subunit harboring an N-terminal 3′-to-5′ exoribonuclease (ExoN) domain and a C-terminal (N7-guanine)-methyltransferase (N7-MTase) domain that is presumably involved in viral mRNA capping. Here, we aimed to integrate structural, biochemical, and virological data to assess the importance of conserved N7-MTase residues for nsp14's enzymatic activities and virus viability. We revisited the crystal structure of severe acute respiratory syndrome (SARS)-CoV nsp14 to perform an *in silico* comparative analysis between betacoronaviruses. We identified several residues likely involved in the formation of the N7-MTase catalytic pocket, which presents a fold distinct from the Rossmann fold observed in most known MTases. Next, for SARS-CoV and Middle East respiratory syndrome-CoV, site-directed mutagenesis of selected residues was used to assess their importance for *in vitro* enzymatic activity. Most of the engineered mutations abolished N7-MTase activity, while not affecting nsp14-ExoN activity. Upon reverse engineering of these mutations into different betacoronavirus genomes, we identified two substitutions (R310A and F426A in SARS-CoV nsp14) abrogating virus viability and one mutation (H424A) yielding a crippled phenotype across all viruses tested. Our results identify the N7-MTase as a critical enzyme for betacoronavirus replication and define key residues of its catalytic pocket that can be targeted to design inhibitors with a potential *pan*-coronaviral activity spectrum.

Significance Statement

The ongoing SARS-CoV-2 pandemic emphasizes the urgent need to develop efficient broad-spectrum anti-CoV drugs. The structure-function characterization of conserved CoV replicative enzymes is key to identifying the most suitable drug targets. Using a multidisciplinary comparative approach and different betacoronaviruses, we characterized the key conserved residues of the nsp14 (N7-guanine)-methyltransferase, a poorly defined subunit of the CoV mRNA-synthesizing machinery. Our study highlights the unique structural features of this enzyme and establishes its essential role in betacoronavirus replication, while identifying two residues that are critical for the replication of the four betacoronaviruses tested, including SARS-CoV-2.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Main Text

Introduction

At their 5′ end, all eukaryotic mRNAs carry an N7-methylguanosine cap that ensures their translation by mediating mRNA recognition during the formation of the ribosomal pre-initiation complex. The co-transcriptional capping of cellular pre-mRNAs occurs in the nucleus and is also critical for pre-mRNA splicing and nuclear export (reviewed in (1-3)). The mRNA cap consists of an N7-methylated 5' guanosine moiety that is linked to the first nucleotide of the transcript by a 5'-5' triphosphate bridge (4). Its synthesis requires (presumably) the consecutive involvement of triphosphatase, guanylyltransferase, and guanine-N7 methyltransferase activities to produce a cap-0 structure. The first nucleotides of mammalian mRNAs are then methylated on the 2'OH position to yield a cap-1 structure that identifies the transcript as "self" and prevents activation of innate immune sensors (reviewed in (2, 5)). Furthermore, the cap structure promotes mRNA stability by providing protection from cellular 5' exoribonucleases.

Viruses rely on host ribosomes for their gene expression and have adopted different strategies to ensure translation of their own mRNAs. These include using the canonical nuclear capping pathway, so-called 'cap-snatching' mechanisms, and replacement of the cap by a ribosome-recruiting RNA structure (reviewed in (2, 6, 7)). Various cytosolically replicating virus families have evolved their own capping machinery. The latter applies to the coronavirus (CoV) family, which includes the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of COVID-19 (8, 9), and a range of other CoVs infecting human or animal hosts (10, 11). This century alone, the CoV family has given rise to three major zoonotic introductions: SARS-CoV-2, the Middle East respiratory syndrome-CoV (MERS-CoV) discovered in 2012, and SARS-CoV, emerging in South East Asia in 2002. All three belong to the genus *Betacoronavirus*, which is abundantly represented among CoVs circulating in bat species (12-15). Despite their demonstrated potential to cross species barriers, prophylactic and therapeutic solutions for CoV infections to prevent or rapidly contain the current COVID-19 pandemic were not available.

The positive-sense CoV genome is unusually large (~30 kb) and its 5' proximal two-thirds encodes for two replicase polyproteins that are post-translationally cleaved into 16 nonstructural proteins (nsp) (16, 17). The CoV replicative enzymes, including the nsp12 RNA-dependent RNA polymerase (RdRp), assemble into a protein complex that is embedded within virus-induced replication organelles (18-20) and directs the synthesis and capping of newly made viral genomes as well as subgenomic mRNAs that serve to express additional CoV genes. Capping is thought to involve the successive action of multiple CoV enzymes: (i) the nsp13 RNA triphosphatase removing the γ phosphate from the nascent 5'-triphosphorylated RNA (21, 22); (ii) an RNA guanylyltransferase (GTAse) producing a GpppN cap by transferring guanosine monophosphate (GMP) to the RNA's dephosphorylated 5' end, a role recently attributed to the nsp12 nucleotidyltransferase (NiRAN) domain, but remaining to be confirmed (23-25); (iii) the nsp14 (N7-guanine)-methyltransferase (N7-MTase) methylating the N7 position of the cap while using S-adenosyl methionine (SAM) as methyl donor; (iv) the nsp16 ribose 2'-O-methyltransferase (2'-O-MTase) converting the cap-0 into a cap-1 structure (^{7m}GpppN$_{2'Om}$; (26, 27)) by performing additional methylation with the assistance of nsp10 as co-factor (26, 28, 29).

Over the past 15 years, the CoV capping machinery has mainly been analyzed *in vitro*, in particular for SARS-CoV, but its characterization in the context of the viral replication cycle has remained limited to a handful of studies. This applies in particular to the CoV N7-MTase domain, expressed as part of the ~60-kDa nsp14, a bi-functional replicase subunit also containing an N-terminal 3'-to-5' exoribonuclease domain implicated in promoting the fidelity of CoV replication (30, 31). Following the discovery of an N7-MTase activity associated with nsp14's C-terminal domain (27), the protein was found to methylate non-methylated cap analogues or guanosine triphosphate (GTP) substrates in the presence of SAM in biochemical assays (26, 32, 33). While the association of nsp10 with nsp14 enhances its ExoN activity, the *in vitro* N7-MTase activity does not depend on nsp10 co-factor (26, 34). Biochemical and structural characterization of the N7-MTase and ExoN domains demonstrated that the two domains are functionally distinct (35-38). Nevertheless, truncations and alanine substitutions in the ExoN domain can severely affect SAM binding and N7-

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

MTase activity (27, 33). The notion that the two enzymatic domains are structurally intertwined was also supported by the SARS-CoV nsp14 crystal structure (35, 36) which was found to be composed of (i) a flexible N-terminal sub-domain forming the nsp10 binding site (aa 1-58), (ii) the 3'-to-5' exoribonuclease (ExoN) domain (aa 1-291), (iii) a flexible hinge region consisting of a loop that connects the N- and C-terminal domains, and three strands protruding from the C-terminal domain (aa 285-300 and aa 407-430), and (iv) the C-terminal N7-MTase domain (aa 292-527) ((35, 36); Fig. 1A).

Interestingly, the structural analysis of the SARS-CoV-nsp14 N7-MTase revealed a non-Rossmann fold (36), distinguishing this enzyme from commonly known cellular and viral methyltransferases (39, 40). Despite the biochemical characterization of the CoV N7-MTase, the assessment of its importance for virus replication has remained limited to studies with a few point mutations introduced into nsp14 of murine hepatitis virus, a model betacoronavirus (41-44). These studies highlighted two motifs important for CoV replication: (i) the presumed SAM binding motif I (DxGxPxG/A, with x being any amino acid; Fig. 2C, motif III), first discovered by superimposition of a SARS-CoV nsp14 N7-MTase structure model with the crystal structures of cellular N7-MTases (27); (ii) nsp14 residues 420-428 (Fig. 2C, part of motif VI) that, based on the SARS-CoV crystal structure, seem to form a constricted pocket holding the cap's GTP moiety (35). Comparative analysis of N7-MTase domains revealed that a number of residues crucial for substrate and ligand binding are conserved among homologous enzymes in more distant CoV (Fig. 2A) and other nidoviruses (45-47).

Due to its conservation and unique structural features, the CoV N7-MTase constitutes an attractive target for antiviral drug development (48-50), to combat SARS-CoV-2 or future emerging CoV threats. Only a few compounds have been reported to inhibit nsp14 N7-MTase activity *in vitro* (26, 32, 48-50). Evaluation of their antiviral activity revealed limited inhibition of CoV replication in cell culture, suggesting poor bio-availability and/or specificity (48, 51). Structural, biochemical, and virological studies of CoV N7-MTase structure and function have not been integrated thus far. Here, we set out to define the catalytic pocket, characterize its involvement in enzymatic activity, and use these observations to probe the enzyme's importance for CoV replication. Using four different betacoronaviruses (SARS-CoV, MERS-CoV, MHV, and SARS-CoV-2), we identified conserved features and residues supporting N7-MTase activity and viral replication, thus providing a solid framework for future efforts to design broad-spectrum inhibitors of this critical CoV enzyme.

Results

Identification of key residues for RNA and SAM binding by the CoV N7-MTase. The previously resolved SARS-CoV nsp14 structure (35, 36) revealed how the ExoN and N7-MTase domains are structurally interconnected, with possible functional implications (Fig. 1). Thus far, a structure of nsp14 in complex with 5'-capped RNA is lacking. Due to some structural peculiarities, it was unclear which conserved residues may be mechanistically involved in N7-methylation and how important these may be for overall CoV replication. Therefore, we first revisited the core structure of the SARS-CoV N7-MTase, to guide a subsequent biochemical and virological comparison across multiple betacoronaviruses.

In the SARS-CoV nsp14 structure (35), the ExoN core presents a fold characteristic of the DED/EDh family of exonucleases (31, 52, 53). However, the N7-MTase domain does not exhibit the canonical 'Rossmann fold' that is common among RNA virus MTases, RNA cap-0 MTases at large, and all five classes of SAM-dependent MTases (54, 55). A hinge region that is highly conserved across CoVs is present at the interface of nsp14's ExoN and N7-MTase domains (Fig. 1A) and constitutes a unique structural feature of this bi-functional CoV protein. It not only connects the two domains, but also forms an extension that protrudes from the surface of the N7-MTase domain (Fig. 1B). Although, the overall structure suggests ExoN and N7-MTase to be separate domains, the successful expression and purification of truncated forms of the N7-MTase domain, with or without the hinge sub-domain, has not been reported (27, 56). This might be related to the hydrophobic nature of the hinge, which is likely important for protein stability and folding. Several studies reported that the replacement of ExoN catalytic residues does not impair the N7-MTase

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

activity, suggesting that the functional interplay between the two domains is limited (26, 27, 33, 37, 38, 48). Whereas the hinge region allows lateral and rotational movement of the two nsp14 domains, one side of the hinge also constitutes the 'ceiling' of the N7-MTase active site (Fig. 1B).

The structures of SARS-CoV nsp14 in complex with SAM and GpppA (PDB: 5C8S and 5C8T; (35)) have defined the enzyme's cap-binding pocket. However, the crystal packing profoundly constrained the structural characterization of the N7-MTase domain and the overall low resolution left uncertainties regarding the positioning of the RNA ligand. Therefore, we performed a thorough structural analysis of the enzyme's cavity, supported by CoV-wide nsp14 sequence comparisons, in order to define conserved N7-MTase residues that may be involved in enzymatic activity (Fig. S2). Several aspects were taken into consideration while delimiting the SAM and RNA binding sites: the general geometry of the cavity, its electrostatic properties, and the conservation of specific amino acid residues. We used Surfnet software (57) to define the volume corresponding to the ligand-binding cavity (Fig. 1C). This volume is shaped as a dual bulb, with the larger pocket accommodating the capped RNA and the smaller one forming the SAM binding site. An electrostatic surface analysis shows positive charges lining the wall of the putative RNA-binding cavity (Fig. 1D and Fig. S1), which would be consistent with its function. Likewise, positive charges that might accommodate the carbocyclic part of the methyl donor were identified in the SAM binding pocket (Fig. 1D). Additionally, conserved hydrophobic residues (Motif I; Fig. 2C) were mapped to a deep hydrophobic cavity, supposedly accommodating the SAM base by a stacking interaction with F426 (SARS-CoV numbering). Finally, the integration of the structural models with CoV-wide N7-MTase sequence comparisons (Table S1 and Fig. S2) allowed the identification of conserved potential key residues within each cavity (blue regions in Fig. 2A). Based on their conservation and positioning, six conserved motifs (I-VI) were defined, each containing a series of specific charged or aromatic residues that have their side chain pointing toward the cavity (Fig. 2B and 2C). Their features suggested they can facilitate the methyl transfer from SAM onto the cap's guanine residue at the 5' end of the RNA substrate, by stabilizing and/or correctly positioning the cap structure. The following potential key residues were identified (amino acid numbers matching those in SARS-CoV nsp14): Motif I, W292; Motif II, N306 and R310; Motif III, D331 and K336; Motif IV, D352; Motif V, N386; Motif VI, Y420, N422, H424, and F426 (Fig. 2B and 2C). To assess the possible impact of their replacement on nsp14 folding, we analyzed the predicted impact of single-site substitutions with alanine on the thermostability of SARS-CoV nsp14 (Table S2 and Fig. 5A)**.** Except for R310, all replacements yielded positive ΔΔG values, suggesting that these mutations may affect MTase stability by altering either its fold, or the cavity for SAM or RNA binding (Table S2 and Fig. 5A). Noticeably, mutations in Motifs I and VI, which are spatially close as part of the hinge and most likely involved in the binding of capped RNA, resulted in the largest ΔΔG gains. Similar observations were made when the impact of substitutions with other amino acids was evaluated for other betacoronaviruses (Table S3).

Identification of residues crucial for *in vitro* N7-MTase activity. To experimentally verify the outcome of our structural analysis (Fig. 1-2), we probed the functional importance of selected residues through targeted mutagenesis and *in vitro* N7-MTase assays. Based on their conservation, charge, position, and potential role for RNA or SAM binding in the catalytic pocket (Fig. 2B and 2C), eleven and nine N7-MTase residues were replaced with alanine in recombinant SARS-CoV and MERS-CoV nsp14, respectively. N-terminally H-tagged proteins were expressed in *E.* coli and purified using immobilized metal affinity chromatography (IMAC) followed by size exclusion chromatography (Fig. 3A and 3B).

We evaluated the N7-MTase activity of nsp14 mutants in an assay using a GpppACCCC capped RNA substrate and radiolabeled [^{3}H]SAM. The transfer of the [^{3}H]methyl group onto the RNA substrate was quantified using filter binding assays (Fig. 3C and 3D), as described previously (26, 34), and compared to the enzymatic activity of wild-type SARS-CoV or MERS-CoV nsp14. With the exception of N306A (30% residual activity), N422A (53% residual activity), and H424A (40% remaining), all SARS-CoV mutations tested almost completely abrogated nsp14 N7-MTase activity (Fig. 3C). In the case of MERS-CoV nsp14, only mutants N418A and F422A retained partial N7-MTase activity, 34% and 70%, respectively, while again all other mutations rendered the enzymatic activity barely detectable (Fig. 3D). In terms of residual activity, differences were

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

observed for some pairs of equivalent SARS-CoV and MERS-CoV mutants (e.g. the H and F in motif VI), but overall the results were fully in line with the outcome of our structural analysis. Thus, our data confirmed and extended a previous study (35), and showed that N7-MTase activity is affected by mutations that either may inhibit SAM binding (W292A, D331A, G333A, K336A, D352A in SARS-CoV) or likely interfere with RNA chain stabilization (N306A, R310A, Y420A, N422A, F426A) in the catalytic pocket.

Revisiting the interplay between the N7-MTase and ExoN domains of nsp14. Despite the notion that the ExoN and N7-MTase domains of CoV nsp14 may be functionally independent (27, 33, 35, 36), they are structurally interconnected by the hinge region (Fig. 1). Therefore, we evaluated the impact of all of our N7-MTase mutations on ExoN functionality, using an *in vitro* assay with 5'-radiolabeled RNA substrate H4 (34), a 22-nt RNA of which the largest part folds into a hairpin structure. Its degradation was monitored using denaturing polyacrylamide gel electrophoresis and autoradiography (Fig. 4). Nsp10 was added as a co-factor that importantly stimulates nsp14 ExoN activity (34, 35, 37), as again confirmed in the 'nsp14 only' control assay (Fig. 4). As expected, in time course experiments, we observed the progressive 3'-to-5' degradation of the RNA substrate by the wild-type nsp10-nsp14 pair of both SARS-CoV (Fig. 4A) and MERS-CoV (Fig. 4B). In the same assay, most of our N7-MTase mutations barely affected ExoN activity (Fig. 4A and 4B), also supporting the notion that these mutant proteins had folded correctly. In contrast, the ExoN activity of SARS-CoV mutants R310A, Y420A and H424A, and MERS-CoV mutant W292A was strongly or partially affected, as indicated by the reduced amount of hydrolysis products at the bottom of the gel. Meanwhile, incorporation of the H420A mutation completely abrogated MERS-CoV nsp14 ExoN activity. Three of the five mutations (Y420A and H424A in SARS-CoV and H420A in MERS-CoV) that affected ExoN activity mapped to motif VI in the hinge region (Fig. 2). Based on the structural analysis, we assume that these mutations affect either the overall nsp14 folding or – more likely - constrain the flexibility of the hinge subdomain with negative consequences for ExoN functionality (35, 36). Conversely, a MERS-CoV ExoN knockout mutant (D90A/E92A), which was included as a control, was found to modestly impact N7-MTase activity (Fig. 3D). Taken together, our data suggest that, although the N7-MTase sequence is well conserved among betacoronaviruses ((35, 37) and Fig. S2), the differences observed between SARS-CoV and MERS-CoV must be caused by a certain level of structural variability or differences in recombinant protein stability.

The nsp14 N7-MTase is critical for SARS-CoV viability. As summarized above, most prior biochemical and structural studies of the CoV N7-MTase were performed using SARS-CoV nsp14, whereas mutagenesis in the context of virus replication (using reverse genetics) was restricted to MHV studies in which, for different reasons, the conserved D and G residues in motif III and the Y residue in motif VI were targeted (41, 43, 58). To establish a connection between the biochemical and virological data on the N7-MTase, we first introduced twelve single N7-MTase mutations into the SARS-CoV genome, using a bacterial artificial chromosome-based reverse genetics system. Each mutant was engineered in duplicate and launched by *in vitro* transcribing full-length RNA that was electroporated into BHK-21 cells. To propagate viral progeny, if released, transfected BHK-21 cells were mixed with Vero E6 cells and incubated up to 6 days. Each mutant was launched at least four times, using RNA from 2 independent clones in two independent experiments, and mutant phenotypes are summarized in Fig. 5A.

In line with the biochemical data, the non-viable phenotype of six of the twelve SARS-CoV mutants (Fig. 5B) provided clear support for the importance of key residues in N7-MTase motifs II (R310), III (D331 and G333), V (N386), and VI (Y420 and F426). As anticipated, mutations in the canonical SAM binding motif III (DxGxPxG/A) completely abrogated SARS-CoV replication (Fig. 5A), apparently confirming the critical role of D331, which was postulated to be a key residue for methylation upon the discovery of the CoV N7-MTase (27). On the other hand, D331A was the only non-viable SARS-CoV mutant for which reversion to wild-type was occasionally observed,

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

suggesting that a very low level of viral RNA synthesis remained possible in spite of this mutation (see also below).

Remarkably, SARS-CoV mutations N306A, K336A, and N422A in motifs II, III, and VI, respectively, were found to yield viruses with plaque phenotypes and progeny titers similar to those of the wild-type control (Fig. 5), despite the major impact of these mutations on *in vitro* N7-MTase activity (Fig. 3C). Likewise, the viable but severely crippled (small-plaque) virus phenotypes of motif-I mutant W292A and motif VI-mutant H424A were surprising (Fig. 5B), although for the latter the biochemical assays did reveal some activity when performed with an increased enzyme concentration (Fig. 3C and (35)). Interestingly, mutant D352A yielded a mixed-size plaque phenotype, suggesting rapid (pseudo)reversion in a minor fraction of this mutant's progeny (Fig. 5B). For all six viable mutants, the presence of the original mutation in the viral progeny was confirmed by sequence analysis of the full-length nsp14-coding region of the viral genome. No other mutations were detected in this region of the genome. For non-viable mutants, transfected cells were incubated and monitored for 6 days and absence of viral activity was also confirmed by immunofluorescence microscopy with antibodies specific for double-stranded RNA and SARS-CoV nsp4.

In general, our data demonstrated the importance of the N7-MTase domain for SARS-CoV viability and confirmed the importance of the motifs and key residues identified using structural biology and biochemical approaches (summary presented in Fig. 5A). Nevertheless, for several mutants the data from different types of assays did not readily align, which prompted us to expand the reverse genetics efforts to other betacoronaviruses.

Phenotypic differences between betacoronaviruses N7-MTase mutants suggest complex structure-function relationships. Even when targeting highly conserved viral functions, the introduction of equivalent mutations in closely related viruses can sometimes yield remarkably different mutant phenotypes. A recent example is the inactivation of the nsp14 ExoN, which is tolerated by MHV and SARS-CoV, but not by MERS-CoV and SARS-CoV-2, the latter virus having an nsp14 sequence that is 95% identical to that of SARS-CoV (37). To expand our understanding of the impact of N7-MTase mutagenesis, we engineered, launched, and analyzed a set of MERS-CoV and MHV mutants, using technical procedures similar to those described above for SARS-CoV (see Methods). In this case, the production of viable progeny was facilitated by co-culturing transfected BHK-21 cells with host cells appropriate for the amplification of MHV (17clone1 cells) or MERS-CoV (Huh7 cells). Again, each mutant was launched at least four times (from duplicate full-length cDNA clones) and the results are summarized in Fig. 5.

The mutations tested for MERS-CoV and MHV had a large predicted impact in our folding free energy analysis (Table S2 and Table S3) and/or yielded a non-viable or crippled phenotype in our SARS-CoV study (Fig. 5A). We evaluated whether these residues were equally critical for the replication of other betacoronaviruses. For clarity, in the text below we will refer to the conserved key residues of each motif instead of using nsp14 amino acid numbers, which are slightly different when comparing SARS-CoV, MERS-CoV and MHV (see Fig. 5A).

In contrast to the SARS-CoV result, the replacement of the W in SAM binding site motif I was lethal for both MERS-CoV and MHV. Strikingly, mutagenesis of the D and G in motif III (SAM binding site) yielded the opposite outcome: both were not tolerated in SARS-CoV, but resulted in crippled but viable or even wild type-like phenotypes for MERS-CoV and MHV, respectively (Fig. 5B). These results again indicated that CoV N7-MTase active site mutants can be (partially) viable, even in the absence of detectable *in vitro* enzymatic activity (Fig. 3D). Similar to our observations for SARS-CoV, the replacement of the D in motif IV and the N in motif VI had moderate or no impact, respectively, on the production of MERS-CoV progeny (Fig. 5). Replacement of the conserved H in motif VI (RNA binding site) consistently crippled replication across SARS-CoV, MERS-CoV, and MHV (Fig. 5B), while replacement of the conserved Y in the same motif was partially tolerated by MERS-CoV, but not by SARS-CoV and MHV.

Our betacoronavirus comparison identified only two N7-MTase mutations that consistently abrogated the replication of all three viruses tested: the R-to-A in motif II and the F-to-A in Motif VI, which both map to the putative RNA binding site. This was surprising in the case of MERS-CoV, given the fact that this mutation (F422A in MERS-CoV) allowed substantial N7-MTase activity in

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the *in vitro* assay (Fig. 3D). When SARS-CoV-2 emerged during the course of this study, the three mutations that produced a similar phenotype across SARS-CoV, MERS-CoV and MHV (R310A, H424A, and F426A; using SARS-CoV numbering) were also engineered for this newly discovered coronavirus. Again, the R310A and F426A replacements were found to fully abrogate virus replication, while H424A yielded a crippled phenotype in SARS-CoV-2 (Fig. 5).

Discussion

Most viral MTases belong to the Rossmann-fold family (55, 59), a ubiquitous higher-order structure among dinucleotide-binding enzymes (55, 60). The CoV nsp14 N7-MTase was the first identified example of a non-Rossmann fold viral MTase (35, 36, 45), and the only one thus far for which some structural and functional information had been gathered. While some viral N7-MTase crystal structures have been resolved (35, 36, 61-63), their biochemical properties and signature sequences critical for RNA binding or enzymatic activity remain poorly defined compared to *e.g.* the 2′-O-MTases, an example of which is found in CoV nsp16 (reviewed in(6)). Likewise, the biological role and relevance of the CoV N7-MTase have not been explored in much detail. In recent studies and reviews, often related to SARS-CoV-2, the enzyme is widely assumed to secure the translation of CoV subgenomic mRNAs and genome, which obviously is a critical step for any positive-stranded RNA virus. However, direct biochemical evidence showing that CoV mRNAs indeed carry an N7-methylated cap at their 5' end is still lacking. The presence of such a cap on CoV RNAs was first postulated following RNase T1 and T2 digestion studies with ^{32}P-labeled MHV RNA, 40 years ago (64). Additional support came from immunoprecipitation experiments using a cap-specific monoclonal antibody (recognizing both the rare nucleoside 2,2,7-trimethylguanosine and 7-methylguanosine (m7G) cap structures) that brought down the mRNAs of equine torovirus (65), a distant CoV relative for which – perhaps strikingly – an N7-MTase domain still remains to be identified(45). The presence of enzymes required for capping in CoVs and many of their relatives (6, 17, 45, 47, 66) and the *in vitro* activity profile of recombinant CoV nsp14 (26, 27, 32, 33, 37, 38) lend additional credibility to CoV capping and cap methylation, but do not exclude the possibility that the CoV N7-MTase may target other substrates as well.

To enhance our overall understanding of nsp14 N7-MTase structure and function, also in the light of its emergence as an important drug target in the battle against SARS-CoV-2 (50, 67-69), we now revisited the SARS-CoV nsp14 X-ray structure to define the most likely residues involved in N7-MTase substrate binding and catalysis. Instead of a βαβ architecture (a seven-stranded β-sheet surrounded by six α-helices) and the canonical MTase motifs, the CoV N7-MTase incorporates twelve β-strands and five α-helices that form a five-stranded β-sheet core (36, 45). The overall nsp14 structure reveals two domains interconnected by a hinge that may confer the flexibility needed to orchestrate the different functions of the protein during CoV replication (36). Furthermore, the protein binds to nsp10, a critical co-factor for nsp14's ExoN activity (34, 70). The conversion of a 5'-terminal GMP cap (GpppN) into a cap-0 structure (^{7m}GpppN) involves multiple steps: stabilization of the RNA chain, SAM binding, methyl transfer to the N7 position of the cap, release of the methylated RNA substrate, and SAH release. Our structural analysis identified several residues with their side chains pointing towards the catalytic pocket, which could be classified as likely RNA- or SAM-binding motifs (Fig. 2B and 2C). Taking into account the amino acid sequence conservation between MHV, SARS-CoV, SARS-CoV-2, and MERS-CoV (Fig. 2A and alignment in Fig. S2), and the structures available to date (35, 36, 71, 72), we surmised these CoV N7-MTases to have an overall similar fold and structural organization. The impact of alanine substitutions of selected key residues in these motifs was then evaluated both *in vitro*, using SARS-CoV or MERS-CoV recombinant nsp14, and in the context of the viral replication cycle, by engineering the corresponding virus mutants in different betacoronaviruses.

Although the biochemical and virological data presented in this study clearly provide support for the predictions derived from our structural analysis, the overall interpretation of the data set undeniably is much more complex than anticipated (Fig. 5A). Replacement of conserved SARS-CoV and MERS-CoV N7-MTase residues largely or completely abrogated enzymatic activity *in vitro* (Fig. 3C and 3D), supporting their identification as key residues for the enzyme's functionality when

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the protein is expressed alone (N7-MTase activity) or when tested in complex with nsp10 (ExoN activity). However, for several SARS-CoV and MERS-CoV mutations the data on enzymatic activity *in vitro* and virus mutant viability appeared to be at odds with each other (Fig. 5A). One possible interpretation is that (very) low levels of N7-MTase activity may still suffice to support viral replication in cell culture models. Alternatively, the *in vitro* N7-MTase assays may have suffered from technical complications, such as suboptimal or incorrect (mutant) N7-MTase domain folding. This could be different for nsp14 expressed in the context of the virus-infected cell and in the presence of its natural interaction partners, in particular other members of the viral replication and transcription complex. It is conceivable that the impact of nsp14 mutations on the fold and/or critical protein-protein or protein-RNA interactions of the N7-MTase domain could fluctuate between different assay systems. This might explain a stronger (*e.g.*, MERS-CoV mutant F422A) or less dramatic effect in the virus-infected cell compared to what is observed in enzymatic assays (Fig. 5A). Mutations mapping to motif VI (hinge region) yielded inconsistent results in comparison to prior *in vitro* studies (26, 27, 32-35), which might be attributed (in part) to different *in vitro* assay conditions. Such technical explanations, however, do not apply when introducing equivalent substitutions in different betacoronaviruses and evaluating them in the context of the viral replication cycle. Also here apparent inconsistencies were observed in terms of the variable impact of certain mutations on the overall replication of virus mutants. The results obtained with mutations in motif III (the presumed SAM binding motif DxGxPxG/A) were a striking example: the viral phenotype for the D-to-A mutant (D331A in SARS-CoV and MERS-CoV, D330A in MHV) ranged from non-viable for SARS-CoV, via severely crippled for MERS-CoV to wild type-like for MHV (Fig. 5). SARS-CoV residue D331 was first identified as important for N7-MTase activity by the superimposition of nsp14 with cellular N7-MTase structures (27). However, a previous MHV study (43) had already documented that replacement of the corresponding residue D330 did not affect MHV replication, and pointed to G332 as a more important residue in motif III, which was confirmed in this study (Fig. 5). These results are consistent with the SARS-CoV nsp14 crystal structure showing that residue G333 in the DxG motif (G332 in MHV) is in direct contact with the SAM methyl donor (35), although apparently its replacement is not sufficient to render all betacoronaviruses non-viable. These results stress the importance to achieve a series of high-resolution structures of these different proteins in order to determine the subtle mechanistic differences.

The only other N7-MTase position probed by reverse genetics so far was the conserved tyrosine in motif VI (Fig. 2C; Y414 in MHV). This residue attracted attention by the intriguing serendipitous finding that its replacement with histidine did not affect replication of MHV strain A59 in cell culture, but strongly reduced replication and virulence in mice (41). Also, an Y414A substitution was tolerated in MHV-A59 (44, 58), but in our study Y414A prevented the recovery of infectious progeny for MHV strain JHM, which exhibits less robust RNA synthesis and overall replication than MHV-A59. The results for the corresponding SARS-CoV (non-viable) and MERS-CoV (crippled) mutants were also variable, adding to the complexity of the overall picture.

A substantial set of N7-MTase mutations was monitored for 'side effects' at the level of *in vitro* ExoN activity (Fig. 4), although for SARS-CoV and MHV these would unlikely explain a lack of viability as ExoN knock-out mutants for both these viruses are only mildly crippled (42, 58, 73). Strikingly, for MERS-CoV, which does not tolerate ExoN inactivation (37), two of the N7-MTase mutations (G333A in motif III and H420A in motif VI) abolished detectable ExoN activity *in vitro* (Fig. 4B), but still allowed a certain level of virus replication (small-plaque phenotype), an observation that clearly warrants further investigation. In more general terms, the ExoN biochemical assay (Fig. 4) suggested that the functional separation between the two enzyme domains may be less strict than previously concluded, as also recently hypothesized following an *in silico* and biochemical analysis using SARS-CoV-2 nsp14 ExoN domain (72, 73). Alternatively, structural variation may explain the discrepancies observed. The impact of SARS-CoV N7-MTase motif-VI mutations on ExoN activity was major, highlighting the peculiar structural organization of nsp14, in which part of the N7-MTase substrate-binding cavity maps to the hinge that connects the N7-MTase and ExoN domains (Fig. 1). For other N7-MTase motifs probed, the functional separation from ExoN was confirmed, as also deduced from previous studies (27, 33, 35, 38).

In our reverse genetics studies with four betacoronaviruses, a consistent phenotype was observed only for N7-MTase mutants carrying replacements of the conserved R in motif II (non-

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

viable) and the conserved H and F in motif VI (crippled and non-viable, respectively). SARS-CoV residue R310 was previously reported to play a role in SAM binding (33), whereas F426 was proposed to entrench and stabilize the guanosine's purine moiety in the proximity of SAM (35). Our analysis (Fig. 2) redefined both residues as part of putative RNA binding site motifs II and VI, respectively, and they were found to be essential for *in vitro* N7-MTase activity in SARS-CoV. Our results highlight the importance of the nsp14 N7-MTase for CoV replication, but the variable impact of the replacement of several conserved residues suggests a substantial degree of conformational or functional flexibility in the enzyme's active site. Other factors, such as interactions of nsp14 with other replicase subunits, may also contribute to the observed phenotypic differences between equivalent N7-MTase mutants of different betacoronaviruses. Likewise, the translation of *in vitro* N7-MTase activity to virus viability is not straightforward and suggests complex structure-function relationships for the structurally unique CoV N7-MTase. Given both its essential role in CoV replication and its emerging status as a target for antiviral drug development efforts, it will be important to further expand the integrated biochemical and virological analysis to support the rational design of broad-spectrum inhibitors of the CoV N7-MTase.

Materials and Methods

Bioinformatics analysis. Forty-seven CoV nsp14 sequences were retrieved (a complete list is provided in Table S1) and aligned using MAFFT . Delineation of motif I to VI was done manually using Seaview and WebLogo (74, 75). Structure analysis (PDB: 5NFY; (36)), volume estimation, cavity determination and sequence conservation was plotted onto the structure using UCSF Chimera (76). Electrostatic surface calculations were done using APBS (77). Predicting the structural impact of mutations was done using the PoPMuSiC server (http://dezyme.com/en/software) (78). This program introduces single-site mutations into a protein's structure and estimates the change in ΔΔGs values of such mutations. In the next step, all possible single-site mutations (4731 mutations) were sorted by their ΔΔGs, but only those in the conserved motifs in the vicinity of the catalytic pocket were used for further studies. PopMuSic predictions were cross-validated with SNAP2 to assess the impact of single amino acid substitutions on protein function (79).

Recombinant protein expression and purification. Recombinant SARS- and MERS-CoV nsp10 and nsp14 were expressed in *E. coli* and purified as described previously (26), MERS-CoV-nsp14 (37, 49) and MERS-nsp10 (29, 80). Vectors for mutant nsp14 expression were generated by QuikChange site-directed mutagenesis using Accuzyme DNA polymerase (Bioline) and verified by sequence analysis. For each recombinant protein used, two batches were produced and tested in enzymatic assays.

***In vitro* nsp14 N7-MTase activity assay.** Reaction mixtures contained 50 or 200 nM of SARS-CoV or MERS-CoV recombinant nsp14, 7 nM GpppACCCC synthetic RNA substrate, 40 mM Tris-HCl (pH 8.0), 10 mM DTT, 5 mM $MgCl_2$, 1.9 µM SAM, 0.1 µM 3H-SAM (Perkin Elmer). After a 30-min incubation at 30°C, the assay was stopped by addition of a 10-fold volume of ice-cold 100 µM S-adenosyl-homocysteine (SAH; Thermo Fisher). Samples were spotted on DEAE filter mats (PerkinElmer) and washed twice with 10 mM ammonium formate (Sigma-Aldrich) (pH 8.0), twice with MilliQ water, and once with absolute ethanol (Sigma-Aldrich) (26), and MTase activity was quantified using a Wallac scintillation counter. To determine relative enzyme activities, the incorporation measurements for mutant proteins were normalized to values obtained with wild-type nsp14. Samples were measured in triplicate in each experiment.

***In vitro* nsp14 ExoN assay.** Synthetic RNA substrate H4 (34) was radiolabeled at its 5' end using T4 polynucleotide kinase (Epicentre) and [γ-^{32}P]ATP (Perkin Elmer) and used as substrate in ExoN activity assays. To this end, recombinant SARS-CoV or MERS-CoV nsp14 and nsp10 were mixed in a 1:4 concentration ratio of nsp14:nsp10 as indicated in Fig. 4. The proteins were added to 500 nM radiolabeled substrate in reaction buffer (40 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 1 mM DTT).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The protein mix was left for 10 min at room temperature to allow the formation of the complex. Assays were performed at 37°C and stopped by addition of a 3x volume of loading buffer containing 96% formamide and 10 mM EDTA. Samples were analyzed on 7 M urea-containing 14% (wt/vol) polyacrylamide gels (acrylamide/bisacrylamide ratio, 19:1) buffered with 0.5xTris-taurine-EDTA and run at high voltage (1,600 V). Results were visualized by phosphorimaging using a Typhoon-9410 variable-mode scanner (GE Healthcare).

Cell culture. Baby hamster kidney cells (BHK-21; ATCC CCL10), Vero E6 (ATCC; CCL-81), HuH7 cells and mouse 17 Cl1 cells were grown as described previously (19, 37, 81, 82). In order to amplify viral progeny and titrate recombinant CoVs by plaque assay, Vero E6 cells were used for SARS-CoV and SARS-CoV-2, HuH7 cells for MERS-CoV, and 17Cl1 cells for MHV. Cells were cultured in Eagle's minimal essential medium (EMEM; Lonza) with 8% fetal calf serum (FCS; Bodinco) supplemented with 100 IU/ml of penicillin and 100 µg/ml of streptomycin (Sigma) and 2 mM L-Glutamine (PAA Laboratories). After infection, complete EMEM medium containing 2% FCS was used.

Viruses and reverse genetics. Mutations in the nsp14-coding region were engineered by two-step *en passant* recombineering in *E. coli* (83) using a bacterial artificial chromosome (BAC) vector with a full-length cDNA copy of a ß-CoV genome. Virus isolates used were MERS-CoV strain EMC/2012 (84, 85)), SARS-CoV Frankfurt-1 (86), MHV-JHM-IA (87), and SARS-CoV-2 BetaCoV/Wuhan/IVDC-HB-01/2019 (88). When designing mutations, additional translationally silent marker mutations was introduced near the site of mutagenesis, in order to analyze possible reversion and rule out potential contaminations with parental virus. For each mutant, two independent BAC clones were obtained, verified by sequencing of the full-length nsp14-coding region, and used for *in vitro* transcription (mMessage-mMachine T7 Kit; Ambion) and virus launching. Transfections with full-length RNA transcripts were performed as described before (37). Briefly, 5 µg RNA was electroporated into BHK-21 cells using an Amaxa nucleofector 2b (program A-031) and Nucleofection T solution kit (Lonza). Transfected BHK-21 cells were mixed in a 1:1 ratio with cells susceptible to CoV infection: Vero E6 cells (for SARS-CoV and SARS-CoV-2), HuH7 cells for MERS-CoV, or 17Cl1 cells (for MHV). Cell culture supernatants were collected when full cytopathic effect was observed, or at 6 days post transfection and progeny virus titers were determined by plaque assay (89). Viral replication was also monitored by immunofluorescence microscopy using antibodies recognizing double-stranded RNA (dsRNA;(90)) and non-structural or structural CoV proteins (37, 82, 91). To confirm the presence of the original mutations in viral progeny, supernatant from transfected cells was used to infect fresh cells, after which intracellular RNA was isolated with TriPure isolation reagent (Roche Applied Science). Next, the nsp14-coding region was amplified using standard RT-PCR methods and the purified amplicon was sequenced by Sanger sequencing. All work with live (recombinant) class-3 CoVs was done in a biosafety level 3 laboratory at Leiden University Medical Center.

Data Availability

All study data are included in the article and SI Appendix.

Acknowledgments

N.S.O. was supported by the Marie Skłodowska-Curie ETN European Training Network 'ANTIVIRALS' (EU Grant Agreement No. 642434). P.E.K. was the recipient of a scholarship from the Fundation "Méditerranée Infection". This work was supported by the Fondation pour la Recherche Médicale (Aide aux Équipes) and the SCORE project (EU Horizon 2020 research and innovation program, grant agreement 101003627). We thank LUMC colleagues Tessa Nelemans for excellent technical support and Alexander Gorbalenya for scientific discussions and critical reading of the manuscript.

References

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1. A. Ghosh, C. D. Lima, Enzymology of RNA cap synthesis. *Wiley Interdiscip Rev RNA* **1**, 152-172 (2010).
2. A. Ramanathan, G. B. Robb, S. H. Chan, mRNA capping: biological functions and applications. *Nucleic Acids Res* **44**, 7511-7526 (2016).
3. F. Ferron, E. Decroly, B. Selisko, B. Canard, The viral RNA capping machinery as a target for antiviral drugs. *Antiviral Res* **96**, 21-31 (2012).
4. A. J. Shatkin, Capping of eucaryotic mRNAs. *Cell* **9**, 645-653 (1976).
5. E. Kindler, V. Thiel, To sense or not to sense viral RNA--essentials of coronavirus innate immune evasion. *Curr Opin Microbiol* **20**, 69-75 (2014).
6. E. Decroly, F. Ferron, J. Lescar, B. Canard, Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol* **10**, 51-65 (2011).
7. E. V. Koonin, B. Moss, Viruses know more than one way to don a cap. *Proc Natl Acad Sci U S A* **107**, 3283-3284 (2010).
8. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).
9. A. Gorbalenya *et al.*, The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* **5**, 536-544 (2020).
10. L. van der Hoek, Human coronaviruses: what do they cause? *Antivir Ther* **12**, 651-658 (2007).
11. Y. Wang, M. Grunewald, S. Perlman, Coronaviruses: An Updated Overview of Their Replication and Pathogenesis. *Methods Mol Biol* **2203**, 1-29 (2020).
12. X. Y. Ge *et al.*, Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* **503**, 535-538 (2013).
13. V. D. Menachery *et al.*, A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat Med* **21**, 1508-1513 (2015).
14. B. Hu *et al.*, Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog* **13**, e1006698 (2017).
15. J. Cui, F. Li, Z. L. Shi, Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol* **17**, 181-192 (2019).
16. E. J. Snijder *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).
17. E. J. Snijder, E. Decroly, J. Ziebuhr, The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing. *Adv Virus Res* **96**, 59-126 (2016).
18. K. Knoops *et al.*, SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biol* **6**, e226 (2008).
19. E. J. Snijder *et al.*, A unifying structural and functional model of the coronavirus replication organelle: Tracking down RNA synthesis. *PLoS Biol* **18**, e3000715 (2020).
20. S. Klein *et al.*, SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. *Nat Commun* **11**, 5885 (2020).
21. A. Seybert, A. Hegyi, S. G. Siddell, J. Ziebuhr, The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5'-to-3' polarity. *RNA* **6**, 1056-1068 (2000).
22. K. A. Ivanov, J. Ziebuhr, Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities. *J Virol* **78**, 7833-7838 (2004).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

23. K. C. Lehmann *et al.*, Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. *Nucleic Acids Res* **43**, 8416-8434 (2015).
24. L. Yan *et al.*, Cryo-EM Structure of an Extended SARS-CoV-2 Replication and Transcription Complex Reveals an Intermediate State in Cap Synthesis. *Cell* 10.1016/j.cell.2020.11.016 (2020).
25. A. Shannon *et al.*, Protein-primed RNA synthesis in SARS-CoVs and structural basis for inhibition by AT-527. *bioRxiv* 10.1101/2021.03.23.436564, 2021.2003.2023.436564 (2021).
26. M. Bouvet *et al.*, In vitro reconstitution of SARS-coronavirus mRNA cap methylation. *PLoS Pathog* **6**, e1000863 (2010).
27. Y. Chen *et al.*, Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc Natl Acad Sci U S A* **106**, 3484-3489 (2009).
28. E. Decroly *et al.*, Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2'O)-methyltransferase activity. *J Virol* **82**, 8071-8084 (2008).
29. Y. Chen *et al.*, Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. *PLoS Pathog* **7**, e1002294 (2011).
30. M. R. Denison, R. L. Graham, E. F. Donaldson, L. D. Eckerle, R. S. Baric, Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity. *RNA Biol* **8**, 270-279 (2011).
31. N. S. Ogando *et al.*, The curious case of the nidovirus exoribonuclease: its role in RNA synthesis and replication fidelity. *Front Microbiol* **10**, 1813 (2019).
32. X. Jin *et al.*, Characterization of the guanine-N7 methyltransferase activity of coronavirus nsp14 on nucleotide GTP. *Virus Res* **176**, 45-52 (2013).
33. Y. Chen *et al.*, Structure-function analysis of severe acute respiratory syndrome coronavirus RNA cap guanine-N7-methyltransferase. *J Virol* **87**, 6296-6305 (2013).
34. M. Bouvet *et al.*, RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proc Natl Acad Sci U S A* **109**, 9372-9377 (2012).
35. Y. Ma *et al.*, Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. *Proc Natl Acad Sci U S A* **112**, 9436-9441 (2015).
36. F. Ferron *et al.*, Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *Proc Natl Acad Sci U S A* **115**, E162-E171 (2018).
37. N. S. Ogando *et al.*, The enzymatic activity of the nsp14 exoribonuclease is critical for replication of MERS-CoV and SARS-CoV-2. *J Virol* 10.1128/JVI.01246-20 (2020).
38. M. Saramago *et al.*, New targets for drug design: importance of nsp14/nsp10 complex formation for the 3'-5' exoribonucleolytic activity on SARS-CoV-2. *FEBS J* **288**, 5130-5147 (2021).
39. L. Xie, P. E. Bourne, Detecting evolutionary relationships across existing fold space, using sequence order-independent profile-profile alignments. *Proc Natl Acad Sci U S A* **105**, 5441-5446 (2008).
40. R. Gana, S. Rao, H. Huang, C. Wu, S. Vasudevan, Structural and functional studies of S-adenosyl-L-methionine binding proteins: a ligand-centric approach. *BMC Struct Biol* **13**, 6 (2013).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

41. S. M. Sperry *et al.*, Single-amino-acid substitutions in open reading frame (ORF) 1b-nsp14 and ORF 2a proteins of the coronavirus mouse hepatitis virus are attenuating in mice. *J Virol* **79**, 3391-3400 (2005).
42. L. D. Eckerle, X. Lu, S. M. Sperry, L. Choi, M. R. Denison, High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J Virol* **81**, 12135-12144 (2007).
43. J. B. Case, A. W. Ashbrook, T. S. Dermody, M. R. Denison, Mutagenesis of S-adenosyl-l-methionine-binding residues in coronavirus nsp14 N7-methyltransferase demonstrates differing requirements for genome translation and resistance to innate immunity. *J Virol* **90**, 7248-7256 (2016).
44. Z. Zhang *et al.*, Live attenuated coronavirus vaccines deficient in N7-Methyltransferase activity induce both humoral and cellular immune responses in mice. *Emerg Microbes Infect* **10**, 1626-1637 (2021).
45. F. Ferron, H. J. Debat, A. Shannon, E. Decroly, B. Canard, A N7-guanine RNA cap methyltransferase signature-sequence as a genetic marker of large genome, non-mammalian Tobaniviridae. *NAR Genom Bioinform* **2**, lqz022 (2020).
46. P. T. Nga *et al.*, Discovery of the first insect nidovirus, a missing evolutionary link in the emergence of the largest RNA virus genomes. *PLoS Pathog* **7**, e1002215 (2011).
47. A. Saberi, A. A. Gulyaeva, J. L. Brubacher, P. A. Newmark, A. E. Gorbalenya, A planarian nidovirus expands the limits of RNA genome size. *PLoS Pathog* **14**, e1007314 (2018).
48. Y. Sun *et al.*, Yeast-based assays for the high-throughput screening of inhibitors of coronavirus RNA cap guanine-N7-methyltransferase. *Antiviral Res* **104**, 156-164 (2014).
49. W. Aouadi *et al.*, Toward the identification of viral cap-methyltransferase inhibitors by fluorescence screening assay. *Antiviral Res* **144**, 330-339 (2017).
50. R. Ahmed-Belkacem *et al.*, Synthesis of adenine dinucleosides SAM analogs as specific inhibitors of SARS-CoV nsp14 RNA cap guanine-N7-methyltransferase. *Eur J Med Chem* **201**, 112557 (2020).
51. R. He *et al.*, Potent and selective inhibition of SARS coronavirus replication by aurintricarboxylic acid. *Biochem Biophys Res Commun* **320**, 1199-1203 (2004).
52. M. H. Barnes, P. Spacciapoli, D. H. Li, N. C. Brown, The 3'-5' exonuclease site of DNA polymerase III from gram-positive bacteria: definition of a novel motif structure. *Gene* **165**, 45-50 (1995).
53. Y. Zuo, M. P. Deutscher, Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res* **29**, 1017-1026 (2001).
54. M. Byszewska, M. Smietanski, E. Purta, J. M. Bujnicki, RNA methyltransferases involved in 5' cap biosynthesis. *RNA Biol* **11**, 1597-1607 (2014).
55. B. P. S. Chouhan, S. Maimaiti, M. Gade, P. Laurino, Rossmann-Fold Methyltransferases: Taking a "beta-Turn" around Their Cofactor, S-Adenosylmethionine. *Biochemistry* **58**, 166-170 (2019).
56. N. H. Moeller *et al.*, Structure and dynamics of SARS-CoV-2 proofreading exoribonuclease ExoN. *bioRxiv* 10.1101/2021.04.02.438274, 2021.2004.2002.438274 (2021).
57. R. A. Laskowski, SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J Mol Graph* **13**, 323-330, 307-328 (1995).
58. L. D. Eckerle, S. M. Brockway, S. M. Sperry, X. Lu, M. R. Denison, Effects of mutagenesis of murine hepatitis virus nsp1 and nsp14 on replication in culture. *Adv Exp Med Biol* **581**, 55-60 (2006).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

59. K. E. Medvedev, L. N. Kinch, N. V. Grishin, Functional and evolutionary analysis of viral proteins containing a Rossmann-like fold. *Protein Sci* **27**, 1450-1463 (2018).
60. S. T. Rao, M. G. Rossmann, Comparison of super-secondary structures in proteins. *J Mol Biol* **76**, 241-256 (1973).
61. G. Sutton, J. M. Grimes, D. I. Stuart, P. Roy, Bluetongue virus VP4 is an RNA-capping assembly line. *Nat Struct Mol Biol* **14**, 449-451 (2007).
62. Y. Tao, D. L. Farsetta, M. L. Nibert, S. C. Harrison, RNA synthesis in a cage--structural studies of reovirus polymerase lambda3. *Cell* **111**, 733-745 (2002).
63. M. P. Egloff, D. Benarroch, B. Selisko, J. L. Romette, B. Canard, An RNA cap (nucleoside-2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. *EMBO J* **21**, 2757-2768 (2002).
64. M. M. Lai, S. A. Stohlman, Comparative analysis of RNA genomes of mouse hepatitis viruses. *J Virol* **38**, 661-670 (1981).
65. A. L. van Vliet, S. L. Smits, P. J. Rottier, R. J. de Groot, Discontinuous and non-discontinuous subgenomic RNA transcription in a nidovirus. *EMBO J* **21**, 6571-6580 (2002).
66. A. Seybert *et al.*, A complex zinc finger controls the enzymatic activities of nidovirus helicases. *J Virol* **79**, 696-704 (2005).
67. L. A. Pearson *et al.*, Development of a High-Throughput Screening Assay to Identify Inhibitors of the SARS-CoV-2 Guanine-N7-Methyltransferase Using RapidFire Mass Spectrometry. *SLAS Discov* 10.1177/24725552211000652, 24725552211000652 (2021).
68. K. Devkota *et al.*, Probing the SAM Binding Site of SARS-CoV-2 Nsp14 In Vitro Using SAM Competitive Inhibitors Guides Developing Selective Bisubstrate Inhibitors. *SLAS Discov* **26**, 1200-1211 (2021).
69. S. Basu *et al.*, Identifying SARS-CoV-2 antiviral compounds by screening for small molecule inhibitors of Nsp14 RNA cap methyltransferase. *Biochem J* **478**, 2481-2497 (2021).
70. M. Bouvet *et al.*, Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes. *J Biol Chem* **289**, 25783-25796 (2014).
71. L. Yan *et al.*, Coupling of N7-methyltransferase and 3'-5' exoribonuclease with SARS-CoV-2 polymerase reveals mechanisms for capping and proofreading. *Cell* **184**, 3474-3485 e3411 (2021).
72. C. Liu *et al.*, Structural basis of mismatch recognition by a SARS-CoV-2 proofreading enzyme. *Science* **373**, 1142-1146 (2021).
73. E. C. Smith, H. Blanc, M. C. Surdel, M. Vignuzzi, M. R. Denison, Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. *PLoS Pathog* **9**, e1003565 (2013).
74. M. Gouy, S. Guindon, O. Gascuel, SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**, 221-224 (2010).
75. G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, WebLogo: a sequence logo generator. *Genome Res* **14**, 1188-1190 (2004).
76. E. F. Pettersen *et al.*, UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004).
77. E. Jurrus *et al.*, Improvements to the APBS biomolecular solvation software suite. *Protein Sci* **27**, 112-128 (2018).
78. Y. Dehouck, J. M. Kwasigroch, D. Gilis, M. Rooman, PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. *BMC Bioinformatics* **12**, 151 (2011).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

79. M. Hecht, Y. Bromberg, B. Rost, Better prediction of functional effects for sequence variants. *BMC Genomics* **16 Suppl 8**, S1 (2015).
80. Y. Wang *et al.*, Coronavirus nsp10/nsp16 methyltransferase can be targeted by nsp10-derived peptide in vitro and in vivo to reduce replication and pathogenesis. *J Virol* **89**, 8416-8427 (2015).
81. D. D. Nedialkova, A. E. Gorbalenya, E. J. Snijder, Arterivirus Nsp1 modulates the accumulation of minus-strand templates to control the relative abundance of viral mRNAs. *PLoS Pathog* **6**, e1000772 (2010).
82. A. H. de Wilde *et al.*, MERS-coronavirus replication induces severe in vitro cytopathology and is strongly inhibited by cyclosporin A or interferon-alpha treatment. *J Gen Virol* **94**, 1749-1760 (2013).
83. B. K. Tischer, G. A. Smith, N. Osterrieder, En passant mutagenesis: a two step markerless red recombination system. *Methods Mol Biol* **634**, 421-430 (2010).
84. F. Almazan *et al.*, Engineering a replication-competent, propagation-defective Middle East respiratory syndrome coronavirus as a vaccine candidate. *mBio* **4**, e00650-00613 (2013).
85. H. H. Rabouw *et al.*, Middle East respiratory coronavirus accessory protein 4a inhibits PKR-mediated antiviral stress responses. *PLoS Pathog* **12**, e1005982 (2016).
86. S. Pfefferle *et al.*, Reverse genetic characterization of the natural genomic deletion in SARS-Coronavirus strain Frankfurt-1 open reading frame 7b reveals an attenuating function of the 7b protein in-vitro and in-vivo. *Virol J* **6**, 131 (2009).
87. A. R. Fehr *et al.*, The nsp3 macrodomain promotes virulence in mice with coronavirus-induced encephalitis. *J Virol* **89**, 1523-1536 (2015).
88. T. T. N. Thao *et al.*, Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. *Nature* 10.1038/s41586-020-2294-9 (2020).
89. S. H. van den Worm *et al.*, Reverse genetics of SARS-related coronavirus using vaccinia virus-based recombination. *PLoS One* **7**, e32857 (2012).
90. F. Weber, V. Wagner, S. B. Rasmussen, R. Hartmann, S. R. Paludan, Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. *J Virol* **80**, 5059-5064 (2006).
91. M. J. van Hemert *et al.*, SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity in vitro. *PLoS Pathog* **4**, e1000054 (2008).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figures and Tables

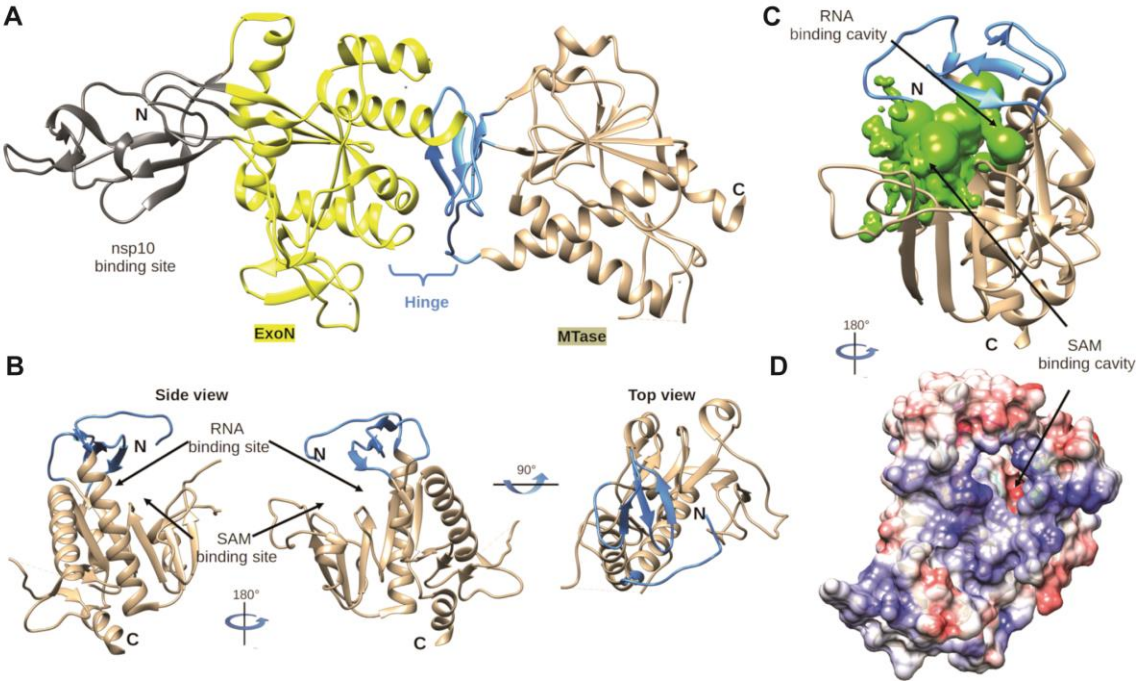

Figure 1. Global architecture of coronavirus nsp14. (A) Architecture of SARS-CoV nsp14 (PDB 5NFY) showing the nsp10 binding site (grey), N-terminal ExoN domain (yellow), hinge subdomain (blue), and C-terminal N7-MTase domain (brown). (B) Side and top view of the hinge region and N7-MTase domain. The three strands of the hinge (blue) protrude from the N7-MTase domain (brown). (C) Analysis of the volume of the N7-MTase active site, with the cavity highlighted in green and hinge subdomain in blue. (D) Electrostatic surface representation of the CoV nsp14 hinge region and N7-MTase domain. Surface electrostatic potential calculated by Adaptive Poisson-Boltzmann Solver, from - 10 (red) to + 10 (blue) kT/e.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

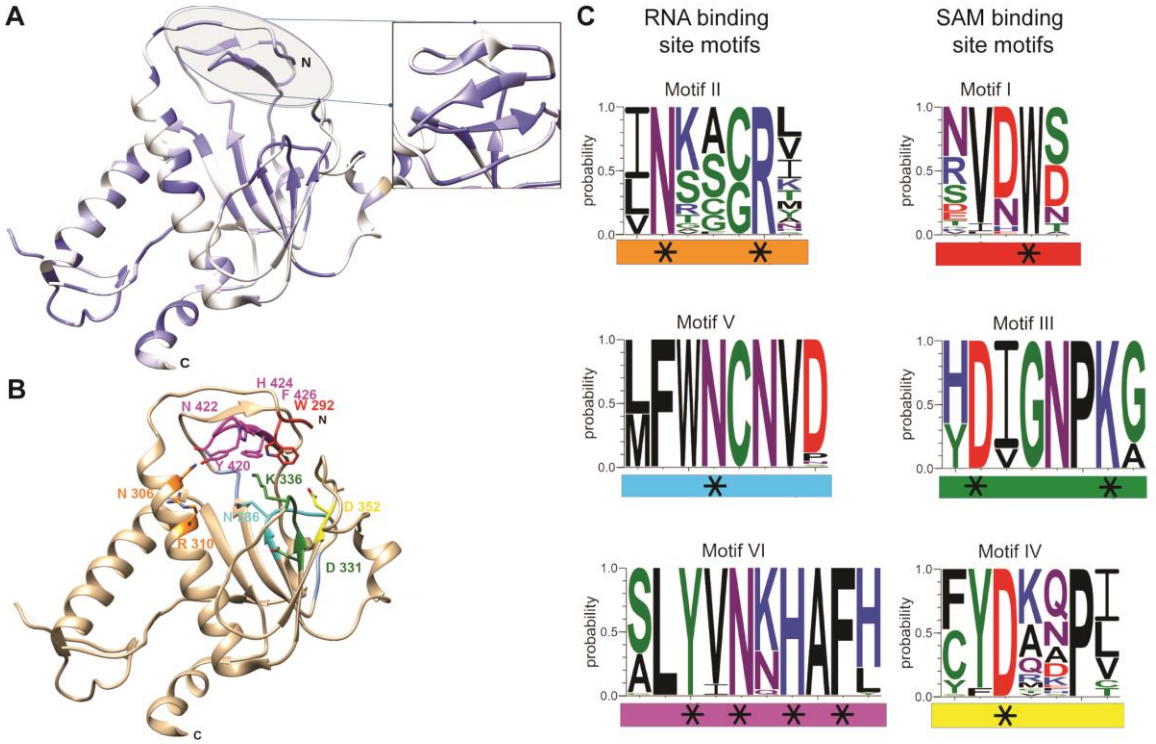

Figure 2. Coronavirus-wide nsp14 N7-MTase conservation and structural analysis. (A) CoV nsp14 amino acid sequence conservation plotted on the structure (PDB 5NFY) of the SARS-CoV hinge region and N7-MTase domain (dark blue to white shading representing 100% to less than 50% sequence identity). A list of sequences used for this comparison is presented in Table S1. (B) Close-up of identified conserved motifs and residues in the N7-MTase catalytic pocket. (C) WebLogo representation of 6 conserved motifs (I-VI) identified in the N7-MTase catalytic pocket. Each motif is highlighted with a specific color (matching that in panel B) and categorized as a proposed SAM- or RNA-binding motif. Black stars highlight charged or aromatic residues most likely involved in ligand binding or catalytic activity.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

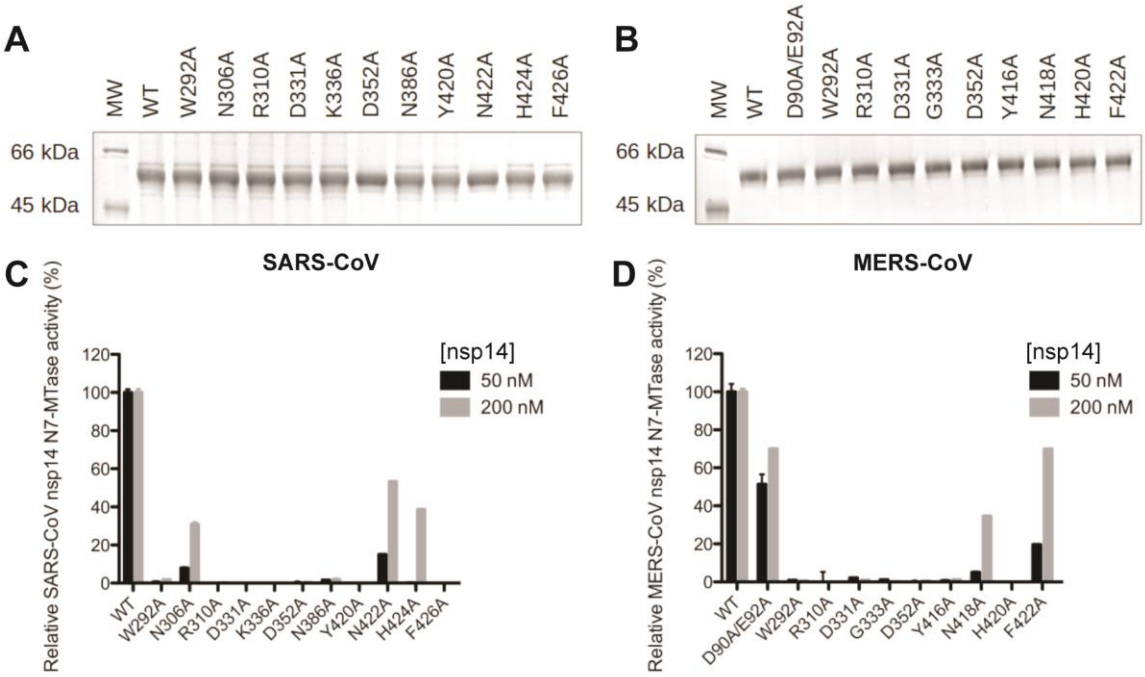

Figure 3. Expression and *in vitro* N7-MTase activity of SARS-CoV and MERS-CoV nsp14 mutants. Recombinant SARS-CoV (A) and MERS-CoV (B) wild-type and mutant nsp14 proteins were expressed in *E. coli* and purified. Proteins were loaded (2 µg and 1 µg for SARS-CoV and MERS-CoV, respectively) and analyzed using 10% SDS-PAGE gels stained with Coomassie blue. The *in vitro* N7-MTase activity of SARS-CoV (C) and MERS-CoV (D) nsp14 mutants was determined using an assay with a GpppACCCC synthetic RNA substrate and radiolabeled SAM as methyl donor. Nsp14 concentrations of 50 and 200 nM were used, as indicated. N7-MTase activities were compared to those of the respective wild-type nsp14 controls. For MERS-CoV, ExoN knockout mutant D90A/E92A was included as a control.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

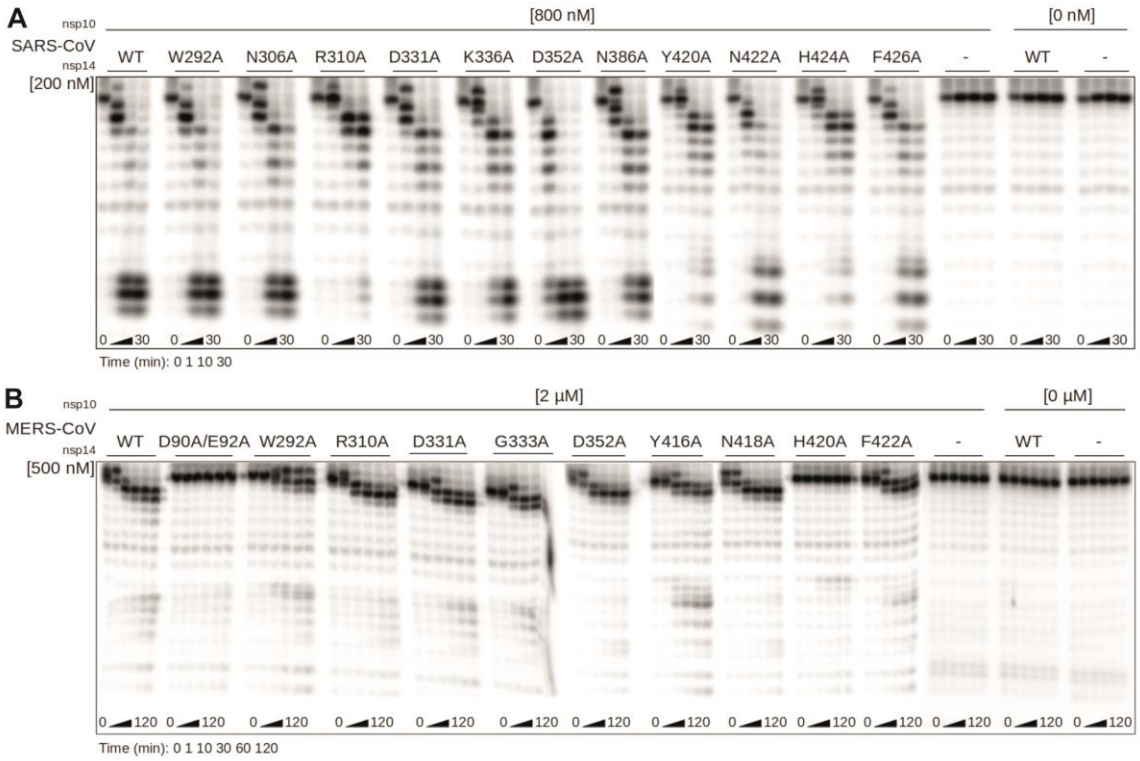

Figure 4. *In vitro* exoribonuclease activity of SARS-CoV and MERS-CoV N7-MTase mutants. The *in vitro* ExoN activity of SARS-CoV (A) and MERS-CoV (B) mutant nsp14 proteins (Fig. 3) was determined by monitoring the degradation of a 5' radiolabeled RNA substrate (see Methods). An nsp14 concentration of 200 or 500 nM was used (as indicated) and a fourfold molar excess of the corresponding nsp10 was added. A time course assay was performed using time points 0, 1, 10, and 30 min for SARS-CoV, and 0, 1, 10, 30, 60, and 120 min for MERS-CoV nsp14. Reaction products were analyzed by denaturing gel electrophoresis and autoradiography.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.17.444407; this version posted November 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A

Coronavirus N7-MTase motif	I	II		III			IV	V	VI			
Conserved residue replaced with Alanine	W	N	R	D	G	K	D	N	Y	N	H	F
nsp14 aa # SARS-CoV and SARS-CoV-2	**292**	**306**	**310**	**331**	**333**	**336**	**352**	**386**	**420**	**422**	**424**	**426**
nsp14 aa # MERS-CoV	292		310	331	333		352		416	418	420	422
nsp14 aa # MHV	291		309	330	332				414		418	420
Predicted impact of Ala substitution on **SARS-CoV nsp14 fold** (ΔΔG in kcal/mol)†	1.8	0.72	0	0.46		0.38	1.12	0.99	3.35	0.92	2.21	1.57
Predicted to interact with	SAM	RNA		SAM			SAM	RNA	RNA			
Mutant *in vitro* N7-MTase activity SARS-CoV*	--	+/-	--	--		--	--	--	--	+	+/-	--
Mutant *in vitro* N7-MTase activity MERS-CoV*	--		--	--	--		--		--	+/-	--	+
Mutant *in vitro* ExoN activity SARS-CoV**	+	+	+/-	+		+	+	+	+/-	+	+/-	+
Mutant *in vitro* ExoN activity MERS-CoV**	+/-		+	+	+		+		+	+	--	+
Virus mutant viability SARS-CoV§	+/-	++	--	--	--	++/-	+/-	--	--	++	+/-	--
Virus mutant viability MERS-CoV§	--		--	+/-	+/-		++/-		+/-	++	+/-	--
Virus mutant viability MHV§	--		--	++	--				--		++/-	--
Virus mutant viability SARS-CoV-2§			--								++/-	--

B

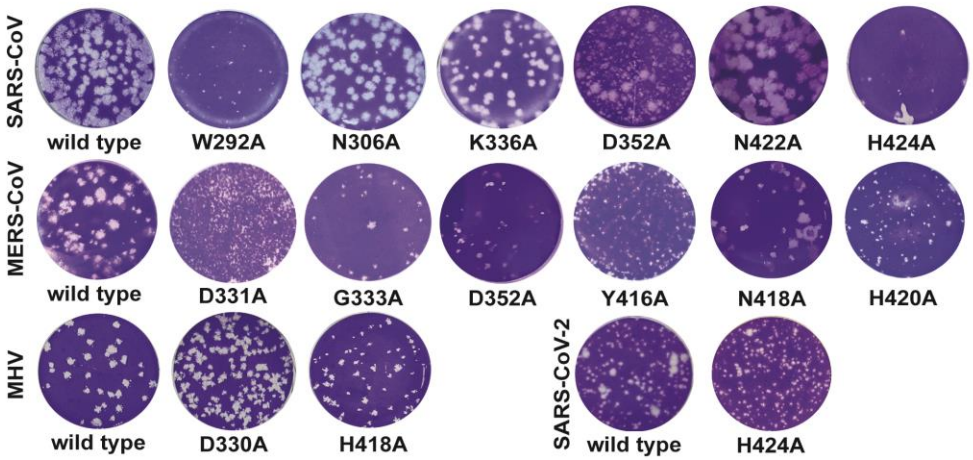

Figure 5. Virological characterization of betacoronavirus N7-MTase mutants. (A) Summary of results obtained from *in silico*, biochemical and virological studies of CoV mutants. † Values presented correspond to ΔΔG values from Table S2. * N7-MTase activity of each mutant was compared to the wild-type control enzyme and scored +, +/-, or -- when >50%, between 10 and 50%, or <10%, respectively. Values used here correspond to results obtained using the high enzyme concentration (Fig. 3B and 3D). ** ExoN activity of each mutant was evaluated relative to the wild-type control enzyme and scored +, +/-, or -- for equal, reduced, and abolished ExoN activity, respectively. § Mutant virus phenotypes, as deduced from plaque assays, were scored as: --, non-viable; +/-, severely crippled; ++/-, mildly crippled; ++, similar to the wild-type control. Empty cells indicate mutants that were not generated. (B) Plaque phenotype of the progeny of viable N7-MTase mutants. Plaque assays were performed using supernatants harvested from transfected cells at 3 (MERS-CoV in HuH7 cells, SARS-CoV and SARS-CoV-2 in Vero E6 cells) or 4 days (MHV in 17Cl1) post transfection.