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

Natacha S Ogando, Priscila El Kazzi, Jessika C Zevenhoven-Dobbe, Brenda W Bontes, Alice Decombe, Clara C S Posthuma, Volker Thiel, Bruno Canard, Francois Ferron, Etienne Decroly, et al.

HAL Id: hal-03411659
https://amu.hal.science/hal-03411659

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

4

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

8   [1] Department of Medical Microbiology, Leiden University Medical Center, Leiden, The
9   Netherlands, 2333ZA

10   [2] Architecture et Fonction des Macromolécules Biologiques, Centre National de la Recherche
11   Scientifique, Aix-Marseille Université, Marseille, France, 13288

12   [3] European Virus Bioinformatics Center, Jena, Germany, 07743.

13   [4] Institute of Virology and Immunology (IVI), Bern, Switzerland,3350.

14   [5] Department of Infectious Diseases and Pathobiology, Vetsuisse Faculty, University of Bern,
15   Bern, Switzerland, 3012.

16   [#] Authors contributed equally

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

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

19   **Author Contributions:** N.O., P.E.K., J.C.Z., B.B., A.D., and F.F. performed experimental work.
20   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
21   hypothesis. C.C.P., V.T., B.C., E.D., and E.J.S. supervised research and contributed in scientific
22   discussions. N.O., P.E.K., F.F., E.D., and E.J.S. wrote the paper. All authors approved the
23   version submitted.

24   **Competing Interest Statement:** All authors declare no competing interests.

25   **Classification:** Biological Sciences, Microbiology

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

27

28   **Abstract**

29   As coronaviruses (CoVs) replicate in the host cell cytoplasm, they rely on their own capping
30   machinery to ensure the efficient translation of their mRNAs, protect them from degradation by
31   cellular 5' exoribonucleases, and escape innate immune sensing. The CoV nonstructural protein
32   14 (nsp14) is a bi-functional replicase subunit harboring an N-terminal 3′-to-5′ exoribonuclease
33   (ExoN) domain and a C-terminal (N7-guanine)-methyltransferase (N7-MTase) domain that is
34   presumably involved in viral mRNA capping. Here, we aimed to integrate structural, biochemical,
35   and virological data to assess the importance of conserved N7-MTase residues for nsp14's
36   enzymatic activities and virus viability. We revisited the crystal structure of severe acute respiratory
37   syndrome (SARS)-CoV nsp14 to perform an *in silico* comparative analysis between
38   betacoronaviruses. We identified several residues likely involved in the formation of the N7-MTase
39   catalytic pocket, which presents a fold distinct from the Rossmann fold observed in most known
40   MTases. Next, for SARS-CoV and Middle East respiratory syndrome-CoV, site-directed
41   mutagenesis of selected residues was used to assess their importance for *in vitro* enzymatic
42   activity. Most of the engineered mutations abolished N7-MTase activity, while not affecting nsp14-
43   ExoN activity. Upon reverse engineering of these mutations into different betacoronavirus
44   genomes, we identified two substitutions (R310A and F426A in SARS-CoV nsp14) abrogating virus
45   viability and one mutation (H424A) yielding a crippled phenotype across all viruses tested. Our
46   results identify the N7-MTase as a critical enzyme for betacoronavirus replication and define key
47   residues of its catalytic pocket that can be targeted to design inhibitors with a potential *pan-*
48   coronaviral activity spectrum.

49   **Significance Statement**
50   The ongoing SARS-CoV-2 pandemic emphasizes the urgent need to develop efficient broad-
51   spectrum anti-CoV drugs. The structure-function characterization of conserved CoV replicative
52   enzymes is key to identifying the most suitable drug targets. Using a multidisciplinary comparative
53   approach and different betacoronaviruses, we characterized the key conserved residues of the
54   nsp14 (N7-guanine)-methyltransferase, a poorly defined subunit of the CoV mRNA-synthesizing
55   machinery. Our study highlights the unique structural features of this enzyme and establishes its
56   essential role in betacoronavirus replication, while identifying two residues that are critical for the
57   replication of the four betacoronaviruses tested, including SARS-CoV-2.
58
59
60

2

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-

3

116   MTase activity (27, 33). The notion that the two enzymatic domains are structurally intertwined was
117   also supported by the SARS-CoV nsp14 crystal structure (35, 36) which was found to be composed
118   of (i) a flexible N-terminal sub-domain forming the nsp10 binding site (aa 1-58), (ii) the 3'-to-5'
119   exoribonuclease (ExoN) domain (aa 1-291), (iii) a flexible hinge region consisting of a loop that
120   connects the N- and C-terminal domains, and three strands protruding from the C-terminal domain
121   (aa 285-300 and aa 407-430), and (iv) the C-terminal N7-MTase domain (aa 292-527) ((35, 36);
122   Fig. 1A).
123   Interestingly, the structural analysis of the SARS-CoV-nsp14 N7-MTase revealed a non-
124   Rossmann fold (36), distinguishing this enzyme from commonly known cellular and viral
125   methyltransferases (39, 40). Despite the biochemical characterization of the CoV N7-MTase, the
126   assessment of its importance for virus replication has remained limited to studies with a few point
127   mutations introduced into nsp14 of murine hepatitis virus, a model betacoronavirus (41-44). These
128   studies highlighted two motifs important for CoV replication: (i) the presumed SAM binding motif I
129   (DxGxPxG/A, with x being any amino acid; Fig. 2C, motif III), first discovered by superimposition of
130   a SARS-CoV nsp14 N7-MTase structure model with the crystal structures of cellular N7-MTases
131   (27); (ii) nsp14 residues 420-428 (Fig. 2C, part of motif VI) that, based on the SARS-CoV crystal
132   structure, seem to form a constricted pocket holding the cap's GTP moiety (35). Comparative
133   analysis of N7-MTase domains revealed that a number of residues crucial for substrate and ligand
134   binding are conserved among homologous enzymes in more distant CoV (Fig. 2A) and other
135   nidoviruses (45-47).
136   Due to its conservation and unique structural features, the CoV N7-MTase constitutes an
137   attractive target for antiviral drug development (48-50), to combat SARS-CoV-2 or future emerging
138   CoV threats. Only a few compounds have been reported to inhibit nsp14 N7-MTase activity *in vitro*
139   (26, 32, 48-50). Evaluation of their antiviral activity revealed limited inhibition of CoV replication in
140   cell culture, suggesting poor bio-availability and/or specificity (48, 51). Structural, biochemical, and
141   virological studies of CoV N7-MTase structure and function have not been integrated thus far. Here,
142   we set out to define the catalytic pocket, characterize its involvement in enzymatic activity, and use
143   these observations to probe the enzyme's importance for CoV replication. Using four different
144   betacoronaviruses (SARS-CoV, MERS-CoV, MHV, and SARS-CoV-2), we identified conserved
145   features and residues supporting N7-MTase activity and viral replication, thus providing a solid
146   framework for future efforts to design broad-spectrum inhibitors of this critical CoV enzyme.
147
148
149   **Results**
150
151   **Identification of key residues for RNA and SAM binding by the CoV N7-MTase.** The previously
152   resolved SARS-CoV nsp14 structure (35, 36) revealed how the ExoN and N7-MTase domains are
153   structurally interconnected, with possible functional implications (Fig. 1). Thus far, a structure of
154   nsp14 in complex with 5'-capped RNA is lacking. Due to some structural peculiarities, it was unclear
155   which conserved residues may be mechanistically involved in N7-methylation and how important
156   these may be for overall CoV replication. Therefore, we first revisited the core structure of the
157   SARS-CoV N7-MTase, to guide a subsequent biochemical and virological comparison across
158   multiple betacoronaviruses.
159   In the SARS-CoV nsp14 structure (35), the ExoN core presents a fold characteristic of the
160   DED/EDh family of exonucleases (31, 52, 53). However, the N7-MTase domain does not exhibit
161   the canonical 'Rossmann fold' that is common among RNA virus MTases, RNA cap-0 MTases at
162   large, and all five classes of SAM-dependent MTases (54, 55). A hinge region that is highly
163   conserved across CoVs is present at the interface of nsp14's ExoN and N7-MTase domains (Fig.
164   1A) and constitutes a unique structural feature of this bi-functional CoV protein. It not only connects
165   the two domains, but also forms an extension that protrudes from the surface of the N7-MTase
166   domain (Fig. 1B). Although, the overall structure suggests ExoN and N7-MTase to be separate
167   domains, the successful expression and purification of truncated forms of the N7-MTase domain,
168   with or without the hinge sub-domain, has not been reported (27, 56). This might be related to the
169   hydrophobic nature of the hinge, which is likely important for protein stability and folding. Several
170   studies reported that the replacement of ExoN catalytic residues does not impair the N7-MTase

171 activity, suggesting that the functional interplay between the two domains is limited (26, 27, 33, 37,
172 38, 48). Whereas the hinge region allows lateral and rotational movement of the two nsp14
173 domains, one side of the hinge also constitutes the 'ceiling' of the N7-MTase active site (Fig. 1B).
174   The structures of SARS-CoV nsp14 in complex with SAM and GpppA (PDB: 5C8S and
175 5C8T; (35)) have defined the enzyme's cap-binding pocket. However, the crystal packing
176 profoundly constrained the structural characterization of the N7-MTase domain and the overall low
177 resolution left uncertainties regarding the positioning of the RNA ligand. Therefore, we performed
178 a thorough structural analysis of the enzyme's cavity, supported by CoV-wide nsp14 sequence
179 comparisons, in order to define conserved N7-MTase residues that may be involved in enzymatic
180 activity (Fig. S2). Several aspects were taken into consideration while delimiting the SAM and RNA
181 binding sites: the general geometry of the cavity, its electrostatic properties, and the conservation
182 of specific amino acid residues. We used Surfnet software (57) to define the volume corresponding
183 to the ligand-binding cavity (Fig. 1C). This volume is shaped as a dual bulb, with the larger pocket
184 accommodating the capped RNA and the smaller one forming the SAM binding site. An
185 electrostatic surface analysis shows positive charges lining the wall of the putative RNA-binding
186 cavity (Fig. 1D and Fig. S1), which would be consistent with its function. Likewise, positive charges
187 that might accommodate the carbocyclic part of the methyl donor were identified in the SAM binding
188 pocket (Fig. 1D). Additionally, conserved hydrophobic residues (Motif I; Fig. 2C) were mapped to a
189 deep hydrophobic cavity, supposedly accommodating the SAM base by a stacking interaction with
190 F426 (SARS-CoV numbering). Finally, the integration of the structural models with CoV-wide N7-
191 MTase sequence comparisons (Table S1 and Fig. S2) allowed the identification of conserved
192 potential key residues within each cavity (blue regions in Fig. 2A). Based on their conservation and
193 positioning, six conserved motifs (I-VI) were defined, each containing a series of specific charged
194 or aromatic residues that have their side chain pointing toward the cavity (Fig. 2B and 2C). Their
195 features suggested they can facilitate the methyl transfer from SAM onto the cap's guanine residue
196 at the 5' end of the RNA substrate, by stabilizing and/or correctly positioning the cap structure. The
197 following potential key residues were identified (amino acid numbers matching those in SARS-CoV
198 nsp14): Motif I, W292; Motif II, N306 and R310; Motif III, D331 and K336; Motif IV, D352; Motif V,
199 N386; Motif VI, Y420, N422, H424, and F426 (Fig. 2B and 2C). To assess the possible impact of
200 their replacement on nsp14 folding, we analyzed the predicted impact of single-site substitutions
201 with alanine on the thermostability of SARS-CoV nsp14 (Table S2 and Fig. 5A). Except for R310,
202 all replacements yielded positive ΔΔG values, suggesting that these mutations may affect MTase
203 stability by altering either its fold, or the cavity for SAM or RNA binding (Table S2 and Fig. 5A).
204 Noticeably, mutations in Motifs I and VI, which are spatially close as part of the hinge and most
205 likely involved in the binding of capped RNA, resulted in the largest ΔΔG gains. Similar observations
206 were made when the impact of substitutions with other amino acids was evaluated for other
207 betacoronaviruses (Table S3).
208
209 **Identification of residues crucial for *in vitro* N7-MTase activity.** To experimentally verify the
210 outcome of our structural analysis (Fig. 1-2), we probed the functional importance of selected
211 residues through targeted mutagenesis and *in vitro* N7-MTase assays. Based on their
212 conservation, charge, position, and potential role for RNA or SAM binding in the catalytic pocket
213 (Fig. 2B and 2C), eleven and nine N7-MTase residues were replaced with alanine in recombinant
214 SARS-CoV and MERS-CoV nsp14, respectively. N-terminally H-tagged proteins were expressed
215 in *E.* coli and purified using immobilized metal affinity chromatography (IMAC) followed by size
216 exclusion chromatography (Fig. 3A and 3B).
217   We evaluated the N7-MTase activity of nsp14 mutants in an assay using a GpppACCCC
218 capped RNA substrate and radiolabeled [³H]SAM. The transfer of the [³H]methyl group onto the
219 RNA substrate was quantified using filter binding assays (Fig. 3C and 3D), as described previously
220 (26, 34), and compared to the enzymatic activity of wild-type SARS-CoV or MERS-CoV nsp14.
221 With the exception of N306A (30% residual activity), N422A (53% residual activity), and H424A
222 (40% remaining), all SARS-CoV mutations tested almost completely abrogated nsp14 N7-MTase
223 activity (Fig. 3C). In the case of MERS-CoV nsp14, only mutants N418A and F422A retained partial
224 N7-MTase activity, 34% and 70%, respectively, while again all other mutations rendered the
225 enzymatic activity barely detectable (Fig. 3D). In terms of residual activity, differences were

226   observed for some pairs of equivalent SARS-CoV and MERS-CoV mutants (e.g. the H and F in
227   motif VI), but overall the results were fully in line with the outcome of our structural analysis. Thus,
228   our data confirmed and extended a previous study (35), and showed that N7-MTase activity is
229   affected by mutations that either may inhibit SAM binding (W292A, D331A, G333A, K336A, D352A
230   in SARS-CoV) or likely interfere with RNA chain stabilization (N306A, R310A, Y420A, N422A,
231   F426A) in the catalytic pocket.
232
233   **Revisiting the interplay between the N7-MTase and ExoN domains of nsp14.** Despite the
234   notion that the ExoN and N7-MTase domains of CoV nsp14 may be functionally independent (27,
235   33, 35, 36), they are structurally interconnected by the hinge region (Fig. 1). Therefore, we
236   evaluated the impact of all of our N7-MTase mutations on ExoN functionality, using an *in vitro* assay
237   with 5'-radiolabeled RNA substrate H4 (34), a 22-nt RNA of which the largest part folds into a
238   hairpin structure. Its degradation was monitored using denaturing polyacrylamide gel
239   electrophoresis and autoradiography (Fig. 4). Nsp10 was added as a co-factor that importantly
240   stimulates nsp14 ExoN activity (34, 35, 37), as again confirmed in the 'nsp14 only' control assay
241   (Fig. 4). As expected, in time course experiments, we observed the progressive 3'-to-5' degradation
242   of the RNA substrate by the wild-type nsp10-nsp14 pair of both SARS-CoV (Fig. 4A) and MERS-
243   CoV (Fig. 4B). In the same assay, most of our N7-MTase mutations barely affected ExoN activity
244   (Fig. 4A and 4B), also supporting the notion that these mutant proteins had folded correctly. In
245   contrast, the ExoN activity of SARS-CoV mutants R310A, Y420A and H424A, and MERS-CoV
246   mutant W292A was strongly or partially affected, as indicated by the reduced amount of hydrolysis
247   products at the bottom of the gel. Meanwhile, incorporation of the H420A mutation completely
248   abrogated MERS-CoV nsp14 ExoN activity. Three of the five mutations (Y420A and H424A in
249   SARS-CoV and H420A in MERS-CoV) that affected ExoN activity mapped to motif VI in the hinge
250   region (Fig. 2). Based on the structural analysis, we assume that these mutations affect either the
251   overall nsp14 folding or – more likely - constrain the flexibility of the hinge subdomain with negative
252   consequences for ExoN functionality (35, 36). Conversely, a MERS-CoV ExoN knockout mutant
253   (D90A/E92A), which was included as a control, was found to modestly impact N7-MTase activity
254   (Fig. 3D). Taken together, our data suggest that, although the N7-MTase sequence is well
255   conserved among betacoronaviruses ((35, 37) and Fig. S2), the differences observed between
256   SARS-CoV and MERS-CoV must be caused by a certain level of structural variability or differences
257   in recombinant protein stability.
258
259   **The nsp14 N7-MTase is critical for SARS-CoV viability**. As summarized above, most prior
260   biochemical and structural studies of the CoV N7-MTase were performed using SARS-CoV nsp14,
261   whereas mutagenesis in the context of virus replication (using reverse genetics) was restricted to
262   MHV studies in which, for different reasons, the conserved D and G residues in motif III and the Y
263   residue in motif VI were targeted (41, 43, 58). To establish a connection between the biochemical
264   and virological data on the N7-MTase, we first introduced twelve single N7-MTase mutations into
265   the SARS-CoV genome, using a bacterial artificial chromosome-based reverse genetics system.
266   Each mutant was engineered in duplicate and launched by *in vitro* transcribing full-length RNA that
267   was electroporated into BHK-21 cells. To propagate viral progeny, if released, transfected BHK-21
268   cells were mixed with Vero E6 cells and incubated up to 6 days. Each mutant was launched at least
269   four times, using RNA from 2 independent clones in two independent experiments, and mutant
270   phenotypes are summarized in Fig. 5A.
271   In line with the biochemical data, the non-viable phenotype of six of the twelve SARS-CoV
272   mutants (Fig. 5B) provided clear support for the importance of key residues in N7-MTase motifs II
273   (R310), III (D331 and G333), V (N386), and VI (Y420 and F426). As anticipated, mutations in the
274   canonical SAM binding motif III (DxGxPxG/A) completely abrogated SARS-CoV replication (Fig.
275   5A), apparently confirming the critical role of D331, which was postulated to be a key residue for
276   methylation upon the discovery of the CoV N7-MTase (27). On the other hand, D331A was the only
277   non-viable SARS-CoV mutant for which reversion to wild-type was occasionally observed,

278 suggesting that a very low level of viral RNA synthesis remained possible in spite of this mutation
279 (see also below).
280   Remarkably, SARS-CoV mutations N306A, K336A, and N422A in motifs II, III, and VI,
281 respectively, were found to yield viruses with plaque phenotypes and progeny titers similar to those
282 of the wild-type control (Fig. 5), despite the major impact of these mutations on *in vitro* N7-MTase
283 activity (Fig. 3C). Likewise, the viable but severely crippled (small-plaque) virus phenotypes of
284 motif-I mutant W292A and motif VI-mutant H424A were surprising (Fig. 5B), although for the latter
285 the biochemical assays did reveal some activity when performed with an increased enzyme
286 concentration (Fig. 3C and (35)). Interestingly, mutant D352A yielded a mixed-size plaque
287 phenotype, suggesting rapid (pseudo)reversion in a minor fraction of this mutant's progeny (Fig.
288 5B). For all six viable mutants, the presence of the original mutation in the viral progeny was
289 confirmed by sequence analysis of the full-length nsp14-coding region of the viral genome. No
290 other mutations were detected in this region of the genome. For non-viable mutants, transfected
291 cells were incubated and monitored for 6 days and absence of viral activity was also confirmed by
292 immunofluorescence microscopy with antibodies specific for double-stranded RNA and SARS-CoV
293 nsp4.
294   In general, our data demonstrated the importance of the N7-MTase domain for SARS-CoV
295 viability and confirmed the importance of the motifs and key residues identified using structural
296 biology and biochemical approaches (summary presented in Fig. 5A). Nevertheless, for several
297 mutants the data from different types of assays did not readily align, which prompted us to expand
298 the reverse genetics efforts to other betacoronaviruses.
299
300 **Phenotypic differences between betacoronaviruses N7-MTase mutants suggest complex**
301 **structure-function relationships.** Even when targeting highly conserved viral functions, the
302 introduction of equivalent mutations in closely related viruses can sometimes yield remarkably
303 different mutant phenotypes. A recent example is the inactivation of the nsp14 ExoN, which is
304 tolerated by MHV and SARS-CoV, but not by MERS-CoV and SARS-CoV-2, the latter virus having
305 an nsp14 sequence that is 95% identical to that of SARS-CoV (37). To expand our understanding
306 of the impact of N7-MTase mutagenesis, we engineered, launched, and analyzed a set of MERS-
307 CoV and MHV mutants, using technical procedures similar to those described above for SARS-
308 CoV (see Methods). In this case, the production of viable progeny was facilitated by co-culturing
309 transfected BHK-21 cells with host cells appropriate for the amplification of MHV (17clone1 cells)
310 or MERS-CoV (Huh7 cells). Again, each mutant was launched at least four times (from duplicate
311 full-length cDNA clones) and the results are summarized in Fig. 5.
312   The mutations tested for MERS-CoV and MHV had a large predicted impact in our folding
313 free energy analysis (Table S2 and Table S3) and/or yielded a non-viable or crippled phenotype in
314 our SARS-CoV study (Fig. 5A). We evaluated whether these residues were equally critical for the
315 replication of other betacoronaviruses. For clarity, in the text below we will refer to the conserved
316 key residues of each motif instead of using nsp14 amino acid numbers, which are slightly different
317 when comparing SARS-CoV, MERS-CoV and MHV (see Fig. 5A).
318   In contrast to the SARS-CoV result, the replacement of the W in SAM binding site motif I
319 was lethal for both MERS-CoV and MHV. Strikingly, mutagenesis of the D and G in motif III (SAM
320 binding site) yielded the opposite outcome: both were not tolerated in SARS-CoV, but resulted in
321 crippled but viable or even wild type-like phenotypes for MERS-CoV and MHV, respectively (Fig.
322 5B). These results again indicated that CoV N7-MTase active site mutants can be (partially) viable,
323 even in the absence of detectable *in vitro* enzymatic activity (Fig. 3D). Similar to our observations
324 for SARS-CoV, the replacement of the D in motif IV and the N in motif VI had moderate or no
325 impact, respectively, on the production of MERS-CoV progeny (Fig. 5). Replacement of the
326 conserved H in motif VI (RNA binding site) consistently crippled replication across SARS-CoV,
327 MERS-CoV, and MHV (Fig. 5B), while replacement of the conserved Y in the same motif was
328 partially tolerated by MERS-CoV, but not by SARS-CoV and MHV.
329   Our betacoronavirus comparison identified only two N7-MTase mutations that consistently
330 abrogated the replication of all three viruses tested: the R-to-A in motif II and the F-to-A in Motif VI,
331 which both map to the putative RNA binding site. This was surprising in the case of MERS-CoV,
332 given the fact that this mutation (F422A in MERS-CoV) allowed substantial N7-MTase activity in

7

333 the *in vitro* assay (Fig. 3D). When SARS-CoV-2 emerged during the course of this study, the three
334 mutations that produced a similar phenotype across SARS-CoV, MERS-CoV and MHV (R310A,
335 H424A, and F426A; using SARS-CoV numbering) were also engineered for this newly discovered
336 coronavirus. Again, the R310A and F426A replacements were found to fully abrogate virus
337 replication, while H424A yielded a crippled phenotype in SARS-CoV-2 (Fig. 5).
338
339
340 **Discussion**
341
342 Most viral MTases belong to the Rossmann-fold family (55, 59), a ubiquitous higher-order structure
343 among dinucleotide-binding enzymes (55, 60). The CoV nsp14 N7-MTase was the first identified
344 example of a non-Rossmann fold viral MTase (35, 36, 45), and the only one thus far for which some
345 structural and functional information had been gathered. While some viral N7-MTase crystal
346 structures have been resolved (35, 36, 61-63), their biochemical properties and signature
347 sequences critical for RNA binding or enzymatic activity remain poorly defined compared to *e.g.*
348 the 2′-O-MTases, an example of which is found in CoV nsp16 (reviewed in(6)). Likewise, the
349 biological role and relevance of the CoV N7-MTase have not been explored in much detail. In recent
350 studies and reviews, often related to SARS-CoV-2, the enzyme is widely assumed to secure the
351 translation of CoV subgenomic mRNAs and genome, which obviously is a critical step for any
352 positive-stranded RNA virus. However, direct biochemical evidence showing that CoV mRNAs
353 indeed carry an N7-methylated cap at their 5' end is still lacking. The presence of such a cap on
354 CoV RNAs was first postulated following RNase T1 and T2 digestion studies with ^{32}P-labeled MHV
355 RNA, 40 years ago (64). Additional support came from immunoprecipitation experiments using a
356 cap-specific monoclonal antibody (recognizing both the rare nucleoside 2,2,7-trimethylguanosine
357 and 7-methylguanosine (m7G) cap structures) that brought down the mRNAs of equine torovirus
358 (65), a distant CoV relative for which – perhaps strikingly – an N7-MTase domain still remains to
359 be identified(45). The presence of enzymes required for capping in CoVs and many of their relatives
360 (6, 17, 45, 47, 66) and the *in vitro* activity profile of recombinant CoV nsp14 (26, 27, 32, 33, 37, 38)
361 lend additional credibility to CoV capping and cap methylation, but do not exclude the possibility
362 that the CoV N7-MTase may target other substrates as well.
363   To enhance our overall understanding of nsp14 N7-MTase structure and function, also in
364 the light of its emergence as an important drug target in the battle against SARS-CoV-2 (50, 67-
365 69), we now revisited the SARS-CoV nsp14 X-ray structure to define the most likely residues
366 involved in N7-MTase substrate binding and catalysis. Instead of a βαβ architecture (a seven-
367 stranded β-sheet surrounded by six α-helices) and the canonical MTase motifs, the CoV N7-MTase
368 incorporates twelve β-strands and five α-helices that form a five-stranded β-sheet core (36, 45).
369 The overall nsp14 structure reveals two domains interconnected by a hinge that may confer the
370 flexibility needed to orchestrate the different functions of the protein during CoV replication (36).
371 Furthermore, the protein binds to nsp10, a critical co-factor for nsp14's ExoN activity (34, 70). The
372 conversion of a 5'-terminal GMP cap (GpppN) into a cap-0 structure (^{7m}GpppN) involves multiple
373 steps: stabilization of the RNA chain, SAM binding, methyl transfer to the N7 position of the cap,
374 release of the methylated RNA substrate, and SAH release. Our structural analysis identified
375 several residues with their side chains pointing towards the catalytic pocket, which could be
376 classified as likely RNA- or SAM-binding motifs (Fig. 2B and 2C). Taking into account the amino
377 acid sequence conservation between MHV, SARS-CoV, SARS-CoV-2, and MERS-CoV (Fig. 2A
378 and alignment in Fig. S2), and the structures available to date (35, 36, 71, 72), we surmised these
379 CoV N7-MTases to have an overall similar fold and structural organization. The impact of alanine
380 substitutions of selected key residues in these motifs was then evaluated both *in vitro*, using SARS-
381 CoV or MERS-CoV recombinant nsp14, and in the context of the viral replication cycle, by
382 engineering the corresponding virus mutants in different betacoronaviruses.
383   Although the biochemical and virological data presented in this study clearly provide
384 support for the predictions derived from our structural analysis, the overall interpretation of the data
385 set undeniably is much more complex than anticipated (Fig. 5A). Replacement of conserved SARS-
386 CoV and MERS-CoV N7-MTase residues largely or completely abrogated enzymatic activity *in vitro*
387 (Fig. 3C and 3D), supporting their identification as key residues for the enzyme's functionality when

388   the protein is expressed alone (N7-MTase activity) or when tested in complex with nsp10 (ExoN
389   activity). However, for several SARS-CoV and MERS-CoV mutations the data on enzymatic activity
390   *in vitro* and virus mutant viability appeared to be at odds with each other (Fig. 5A). One possible
391   interpretation is that (very) low levels of N7-MTase activity may still suffice to support viral
392   replication in cell culture models. Alternatively, the *in vitro* N7-MTase assays may have suffered
393   from technical complications, such as suboptimal or incorrect (mutant) N7-MTase domain folding.
394   This could be different for nsp14 expressed in the context of the virus-infected cell and in the
395   presence of its natural interaction partners, in particular other members of the viral replication and
396   transcription complex. It is conceivable that the impact of nsp14 mutations on the fold and/or critical
397   protein-protein or protein-RNA interactions of the N7-MTase domain could fluctuate between
398   different assay systems. This might explain a stronger (*e.g.*, MERS-CoV mutant F422A) or less
399   dramatic effect in the virus-infected cell compared to what is observed in enzymatic assays (Fig.
400   5A). Mutations mapping to motif VI (hinge region) yielded inconsistent results in comparison to prior
401   *in vitro* studies (26, 27, 32-35), which might be attributed (in part) to different *in vitro* assay
402   conditions. Such technical explanations, however, do not apply when introducing equivalent
403   substitutions in different betacoronaviruses and evaluating them in the context of the viral
404   replication cycle. Also here apparent inconsistencies were observed in terms of the variable impact
405   of certain mutations on the overall replication of virus mutants. The results obtained with mutations
406   in motif III (the presumed SAM binding motif DxGxPxG/A) were a striking example: the viral
407   phenotype for the D-to-A mutant (D331A in SARS-CoV and MERS-CoV, D330A in MHV) ranged
408   from non-viable for SARS-CoV, via severely crippled for MERS-CoV to wild type-like for MHV (Fig.
409   5). SARS-CoV residue D331 was first identified as important for N7-MTase activity by the
410   superimposition of nsp14 with cellular N7-MTase structures (27). However, a previous MHV study
411   (43) had already documented that replacement of the corresponding residue D330 did not affect
412   MHV replication, and pointed to G332 as a more important residue in motif III, which was confirmed
413   in this study (Fig. 5). These results are consistent with the SARS-CoV nsp14 crystal structure
414   showing that residue G333 in the DxG motif (G332 in MHV) is in direct contact with the SAM methyl
415   donor (35), although apparently its replacement is not sufficient to render all betacoronaviruses
416   non-viable. These results stress the importance to achieve a series of high-resolution structures of
417   these different proteins in order to determine the subtle mechanistic differences.
418   The only other N7-MTase position probed by reverse genetics so far was the conserved
419   tyrosine in motif VI (Fig. 2C; Y414 in MHV). This residue attracted attention by the intriguing
420   serendipitous finding that its replacement with histidine did not affect replication of MHV strain A59
421   in cell culture, but strongly reduced replication and virulence in mice (41). Also, an Y414A
422   substitution was tolerated in MHV-A59 (44, 58), but in our study Y414A prevented the recovery of
423   infectious progeny for MHV strain JHM, which exhibits less robust RNA synthesis and overall
424   replication than MHV-A59. The results for the corresponding SARS-CoV (non-viable) and MERS-
425   CoV (crippled) mutants were also variable, adding to the complexity of the overall picture.
426   A substantial set of N7-MTase mutations was monitored for 'side effects' at the level of *in*
427   *vitro* ExoN activity (Fig. 4), although for SARS-CoV and MHV these would unlikely explain a lack
428   of viability as ExoN knock-out mutants for both these viruses are only mildly crippled (42, 58, 73).
429   Strikingly, for MERS-CoV, which does not tolerate ExoN inactivation (37), two of the N7-MTase
430   mutations (G333A in motif III and H420A in motif VI) abolished detectable ExoN activity *in vitro*
431   (Fig. 4B), but still allowed a certain level of virus replication (small-plaque phenotype), an
432   observation that clearly warrants further investigation. In more general terms, the ExoN biochemical
433   assay (Fig. 4) suggested that the functional separation between the two enzyme domains may be
434   less strict than previously concluded, as also recently hypothesized following an *in silico* and
435   biochemical analysis using SARS-CoV-2 nsp14 ExoN domain (72, 73). Alternatively, structural
436   variation may explain the discrepancies observed. The impact of SARS-CoV N7-MTase motif-VI
437   mutations on ExoN activity was major, highlighting the peculiar structural organization of nsp14, in
438   which part of the N7-MTase substrate-binding cavity maps to the hinge that connects the N7-MTase
439   and ExoN domains (Fig. 1). For other N7-MTase motifs probed, the functional separation from
440   ExoN was confirmed, as also deduced from previous studies (27, 33, 35, 38).
441   In our reverse genetics studies with four betacoronaviruses, a consistent phenotype was
442   observed only for N7-MTase mutants carrying replacements of the conserved R in motif II (non-

443 viable) and the conserved H and F in motif VI (crippled and non-viable, respectively). SARS-CoV
444 residue R310 was previously reported to play a role in SAM binding (33), whereas F426 was
445 proposed to entrench and stabilize the guanosine's purine moiety in the proximity of SAM (35). Our
446 analysis (Fig. 2) redefined both residues as part of putative RNA binding site motifs II and VI,
447 respectively, and they were found to be essential for *in vitro* N7-MTase activity in SARS-CoV. Our
448 results highlight the importance of the nsp14 N7-MTase for CoV replication, but the variable impact
449 of the replacement of several conserved residues suggests a substantial degree of conformational
450 or functional flexibility in the enzyme's active site. Other factors, such as interactions of nsp14 with
451 other replicase subunits, may also contribute to the observed phenotypic differences between
452 equivalent N7-MTase mutants of different betacoronaviruses. Likewise, the translation of *in vitro*
453 N7-MTase activity to virus viability is not straightforward and suggests complex structure-function
454 relationships for the structurally unique CoV N7-MTase. Given both its essential role in CoV
455 replication and its emerging status as a target for antiviral drug development efforts, it will be
456 important to further expand the integrated biochemical and virological analysis to support the
457 rational design of broad-spectrum inhibitors of the CoV N7-MTase.
458
459
460 **Materials and Methods**
461
462 **Bioinformatics analysis.** Forty-seven CoV nsp14 sequences were retrieved (a complete list is
463 provided in Table S1) and aligned using MAFFT . Delineation of motif I to VI was done manually
464 using Seaview and WebLogo (74, 75). Structure analysis (PDB: 5NFY; (36)), volume estimation,
465 cavity determination and sequence conservation was plotted onto the structure using UCSF
466 Chimera (76). Electrostatic surface calculations were done using APBS (77). Predicting the
467 structural impact of mutations was done using the PoPMuSiC server
468 (http://dezyme.com/en/software) (78). This program introduces single-site mutations into a
469 protein's structure and estimates the change in $\Delta\Delta$Gs values of such mutations. In the next step,
470 all possible single-site mutations (4731 mutations) were sorted by their $\Delta\Delta$Gs, but only those in the
471 conserved motifs in the vicinity of the catalytic pocket were used for further studies. PopMuSic
472 predictions were cross-validated with SNAP2 to assess the impact of single amino acid
473 substitutions on protein function (79).
474
475 **Recombinant protein expression and purification.** Recombinant SARS- and MERS-CoV nsp10
476 and nsp14 were expressed in *E. coli* and purified as described previously (26), MERS-CoV-nsp14
477 (37, 49) and MERS-nsp10 (29, 80). Vectors for mutant nsp14 expression were generated by
478 QuikChange site-directed mutagenesis using Accuzyme DNA polymerase (Bioline) and verified by
479 sequence analysis. For each recombinant protein used, two batches were produced and tested in
480 enzymatic assays.
481
482 ***In vitro* nsp14 N7-MTase activity assay.** Reaction mixtures contained 50 or 200 nM of SARS-
483 CoV or MERS-CoV recombinant nsp14, 7 nM GpppACCCC synthetic RNA substrate, 40 mM Tris-
484 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-
485 min incubation at 30°C, the assay was stopped by addition of a 10-fold volume of ice-cold 100 µM
486 S-adenosyl-homocysteine (SAH; Thermo Fisher). Samples were spotted on DEAE filter mats
487 (PerkinElmer) and washed twice with 10 mM ammonium formate (Sigma-Aldrich) (pH 8.0), twice
488 with MilliQ water, and once with absolute ethanol (Sigma-Aldrich) (26), and MTase activity was
489 quantified using a Wallac scintillation counter. To determine relative enzyme activities, the
490 incorporation measurements for mutant proteins were normalized to values obtained with wild-type
491 nsp14. Samples were measured in triplicate in each experiment.
492
493 ***In vitro* nsp14 ExoN assay.** Synthetic RNA substrate H4 (34) was radiolabeled at its 5' end using
494 T4 polynucleotide kinase (Epicentre) and [γ-^{32}P]ATP (Perkin Elmer) and used as substrate in ExoN
495 activity assays. To this end, recombinant SARS-CoV or MERS-CoV nsp14 and nsp10 were mixed
496 in a 1:4 concentration ratio of nsp14:nsp10 as indicated in Fig. 4. The proteins were added to 500
497 nM radiolabeled substrate in reaction buffer (40 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 1 mM DTT).

498 The protein mix was left for 10 min at room temperature to allow the formation of the complex.
499 Assays were performed at 37°C and stopped by addition of a 3x volume of loading buffer containing
500 96% formamide and 10 mM EDTA. Samples were analyzed on 7 M urea-containing 14% (wt/vol)
501 polyacrylamide gels (acrylamide/bisacrylamide ratio, 19:1) buffered with 0.5xTris-taurine-EDTA
502 and run at high voltage (1,600 V). Results were visualized by phosphorimaging using a Typhoon-
503 9410 variable-mode scanner (GE Healthcare).
504
505 **Cell culture.** Baby hamster kidney cells (BHK-21; ATCC CCL10), Vero E6 (ATCC; CCL-81), HuH7
506 cells and mouse 17 Cl1 cells were grown as described previously (19, 37, 81, 82). In order to
507 amplify viral progeny and titrate recombinant CoVs by plaque assay, Vero E6 cells were used for
508 SARS-CoV and SARS-CoV-2, HuH7 cells for MERS-CoV, and 17Cl1 cells for MHV. Cells were
509 cultured in Eagle's minimal essential medium (EMEM; Lonza) with 8% fetal calf serum (FCS;
510 Bodinco) supplemented with 100 IU/ml of penicillin and 100 μg/ml of streptomycin (Sigma) and 2
511 mM L-Glutamine (PAA Laboratories). After infection, complete EMEM medium containing 2% FCS
512 was used.
513
514 **Viruses and reverse genetics.** Mutations in the nsp14-coding region were engineered by two-
515 step *en passant* recombineering in *E. coli* (83) using a bacterial artificial chromosome (BAC) vector
516 with a full-length cDNA copy of a ß-CoV genome. Virus isolates used were MERS-CoV strain
517 EMC/2012 (84, 85)), SARS-CoV Frankfurt-1 (86), MHV-JHM-IA (87), and SARS-CoV-2
518 BetaCoV/Wuhan/IVDC-HB-01/2019 (88). When designing mutations, additional translationally
519 silent marker mutations was introduced near the site of mutagenesis, in order to analyze possible
520 reversion and rule out potential contaminations with parental virus. For each mutant, two
521 independent BAC clones were obtained, verified by sequencing of the full-length nsp14-coding
522 region, and used for *in vitro* transcription (mMessage-mMachine T7 Kit; Ambion) and virus
523 launching. Transfections with full-length RNA transcripts were performed as described before (37).
524 Briefly, 5 μg RNA was electroporated into BHK-21 cells using an Amaxa nucleofector 2b (program
525 A-031) and Nucleofection T solution kit (Lonza). Transfected BHK-21 cells were mixed in a 1:1 ratio
526 with cells susceptible to CoV infection: Vero E6 cells (for SARS-CoV and SARS-CoV-2), HuH7
527 cells for MERS-CoV, or 17Cl1 cells (for MHV). Cell culture supernatants were collected when full
528 cytopathic effect was observed, or at 6 days post transfection and progeny virus titers were
529 determined by plaque assay (89). Viral replication was also monitored by immunofluorescence
530 microscopy using antibodies recognizing double-stranded RNA (dsRNA;(90)) and non-structural or
531 structural CoV proteins (37, 82, 91). To confirm the presence of the original mutations in viral
532 progeny, supernatant from transfected cells was used to infect fresh cells, after which intracellular
533 RNA was isolated with TriPure isolation reagent (Roche Applied Science). Next, the nsp14-coding
534 region was amplified using standard RT-PCR methods and the purified amplicon was sequenced
535 by Sanger sequencing. All work with live (recombinant) class-3 CoVs was done in a biosafety level
536 3 laboratory at Leiden University Medical Center.
537
538 **Data Availability**
539 All study data are included in the article and SI Appendix.
540
541 **Acknowledgments**
542
550
551
552 **References**

553   1.   A. Ghosh, C. D. Lima, Enzymology of RNA cap synthesis. *Wiley Interdiscip Rev RNA* **1**, 152-
554   172 (2010).
555   2.   A. Ramanathan, G. B. Robb, S. H. Chan, mRNA capping: biological functions and
556   applications. *Nucleic Acids Res* **44**, 7511-7526 (2016).
557   3.   F. Ferron, E. Decroly, B. Selisko, B. Canard, The viral RNA capping machinery as a target
558   for antiviral drugs. *Antiviral Res* **96**, 21-31 (2012).
559   4.   A. J. Shatkin, Capping of eucaryotic mRNAs. *Cell* **9**, 645-653 (1976).
560   5.   E. Kindler, V. Thiel, To sense or not to sense viral RNA--essentials of coronavirus innate
561   immune evasion. *Curr Opin Microbiol* **20**, 69-75 (2014).
562   6.   E. Decroly, F. Ferron, J. Lescar, B. Canard, Conventional and unconventional mechanisms
563   for capping viral mRNA. *Nat Rev Microbiol* **10**, 51-65 (2011).
564   7.   E. V. Koonin, B. Moss, Viruses know more than one way to don a cap. *Proc Natl Acad Sci*
565   *U S A* **107**, 3283-3284 (2010).
566   8.   P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat
567   origin. *Nature* **579**, 270-273 (2020).
568   9.   A. Gorbalenya *et al.*, The species Severe acute respiratory syndrome-related coronavirus:
569   classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* **5**, 536-544 (2020).
570   10.   L. van der Hoek, Human coronaviruses: what do they cause? *Antivir Ther* **12**, 651-658
571   (2007).
572   11.   Y. Wang, M. Grunewald, S. Perlman, Coronaviruses: An Updated Overview of Their
573   Replication and Pathogenesis. *Methods Mol Biol* **2203**, 1-29 (2020).
574   12.   X. Y. Ge *et al.*, Isolation and characterization of a bat SARS-like coronavirus that uses the
575   ACE2 receptor. *Nature* **503**, 535-538 (2013).
576   13.   V. D. Menachery *et al.*, A SARS-like cluster of circulating bat coronaviruses shows potential
577   for human emergence. *Nat Med* **21**, 1508-1513 (2015).
578   14.   B. Hu *et al.*, Discovery of a rich gene pool of bat SARS-related coronaviruses provides new
579   insights into the origin of SARS coronavirus. *PLoS Pathog* **13**, e1006698 (2017).
580   15.   J. Cui, F. Li, Z. L. Shi, Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol*
581   **17**, 181-192 (2019).
582   16.   E. J. Snijder *et al.*, Unique and conserved features of genome and proteome of SARS-
583   coronavirus, an early split-off from the coronavirus group 2 lineage. *J Mol Biol* **331**, 991-
584   1004 (2003).
585   17.   E. J. Snijder, E. Decroly, J. Ziebuhr, The Nonstructural Proteins Directing Coronavirus RNA
586   Synthesis and Processing. *Adv Virus Res* **96**, 59-126 (2016).
587   18.   K. Knoops *et al.*, SARS-coronavirus replication is supported by a reticulovesicular network
588   of modified endoplasmic reticulum. *PLoS Biol* **6**, e226 (2008).
589   19.   E. J. Snijder *et al.*, A unifying structural and functional model of the coronavirus replication
590   organelle: Tracking down RNA synthesis. *PLoS Biol* **18**, e3000715 (2020).
591   20.   S. Klein *et al.*, SARS-CoV-2 structure and replication characterized by in situ cryo-electron
592   tomography. *Nat Commun* **11**, 5885 (2020).
593   21.   A. Seybert, A. Hegyi, S. G. Siddell, J. Ziebuhr, The human coronavirus 229E superfamily 1
594   helicase has RNA and DNA duplex-unwinding activities with 5'-to-3' polarity. *RNA* **6**, 1056-
595   1068 (2000).
596   22.   K. A. Ivanov, J. Ziebuhr, Human coronavirus 229E nonstructural protein 13:
597   characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-
598   triphosphatase activities. *J Virol* **78**, 7833-7838 (2004).

599   23.   K. C. Lehmann *et al.*, Discovery of an essential nucleotidylating activity associated with a
600   newly delineated conserved domain in the RNA polymerase-containing protein of all
601   nidoviruses. *Nucleic Acids Res* **43**, 8416-8434 (2015).
602   24.   L. Yan *et al.*, Cryo-EM Structure of an Extended SARS-CoV-2 Replication and Transcription
603   Complex Reveals an Intermediate State in Cap Synthesis. *Cell* 10.1016/j.cell.2020.11.016
604   (2020).
605   25.   A. Shannon *et al.*, Protein-primed RNA synthesis in SARS-CoVs and structural basis for
606   inhibition by AT-527. *bioRxiv* 10.1101/2021.03.23.436564, 2021.2003.2023.436564
607   (2021).
608   26.   M. Bouvet *et al.*, In vitro reconstitution of SARS-coronavirus mRNA cap methylation. *PLoS*
609   *Pathog* **6**, e1000863 (2010).
610   27.   Y. Chen *et al.*, Functional screen reveals SARS coronavirus nonstructural protein nsp14 as
611   a novel cap N7 methyltransferase. *Proc Natl Acad Sci U S A* **106**, 3484-3489 (2009).
612   28.   E. Decroly *et al.*, Coronavirus nonstructural protein 16 is a cap-0 binding enzyme
613   possessing (nucleoside-2'O)-methyltransferase activity. *J Virol* **82**, 8071-8084 (2008).
614   29.   Y. Chen *et al.*, Biochemical and structural insights into the mechanisms of SARS
615   coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. *PLoS Pathog*
616   **7**, e1002294 (2011).
617   30.   M. R. Denison, R. L. Graham, E. F. Donaldson, L. D. Eckerle, R. S. Baric, Coronaviruses: an
618   RNA proofreading machine regulates replication fidelity and diversity. *RNA Biol* **8**, 270-
619   279 (2011).
620   31.   N. S. Ogando *et al.*, The curious case of the nidovirus exoribonuclease: its role in RNA
621   synthesis and replication fidelity. *Front Microbiol* **10**, 1813 (2019).
622   32.   X. Jin *et al.*, Characterization of the guanine-N7 methyltransferase activity of coronavirus
623   nsp14 on nucleotide GTP. *Virus Res* **176**, 45-52 (2013).
624   33.   Y. Chen *et al.*, Structure-function analysis of severe acute respiratory syndrome
625   coronavirus RNA cap guanine-N7-methyltransferase. *J Virol* **87**, 6296-6305 (2013).
626   34.   M. Bouvet *et al.*, RNA 3'-end mismatch excision by the severe acute respiratory syndrome
627   coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proc Natl Acad*
628   *Sci U S A* **109**, 9372-9377 (2012).
629   35.   Y. Ma *et al.*, Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10
630   complex. *Proc Natl Acad Sci U S A* **112**, 9436-9441 (2015).
631   36.   F. Ferron *et al.*, Structural and molecular basis of mismatch correction and ribavirin
632   excision from coronavirus RNA. *Proc Natl Acad Sci U S A* **115**, E162-E171 (2018).
633   37.   N. S. Ogando *et al.*, The enzymatic activity of the nsp14 exoribonuclease is critical for
634   replication of MERS-CoV and SARS-CoV-2. *J Virol* 10.1128/JVI.01246-20 (2020).
635   38.   M. Saramago *et al.*, New targets for drug design: importance of nsp14/nsp10 complex
636   formation for the 3'-5' exoribonucleolytic activity on SARS-CoV-2. *FEBS J* **288**, 5130-5147
637   (2021).
638   39.   L. Xie, P. E. Bourne, Detecting evolutionary relationships across existing fold space, using
639   sequence order-independent profile-profile alignments. *Proc Natl Acad Sci U S A* **105**,
640   5441-5446 (2008).
641   40.   R. Gana, S. Rao, H. Huang, C. Wu, S. Vasudevan, Structural and functional studies of S-
642   adenosyl-L-methionine binding proteins: a ligand-centric approach. *BMC Struct Biol* **13**, 6
643   (2013).

644 41. S. M. Sperry *et al.*, Single-amino-acid substitutions in open reading frame (ORF) 1b-nsp14
645   and ORF 2a proteins of the coronavirus mouse hepatitis virus are attenuating in mice. *J*
646   *Virol* **79**, 3391-3400 (2005).
647 42. L. D. Eckerle, X. Lu, S. M. Sperry, L. Choi, M. R. Denison, High fidelity of murine hepatitis
648   virus replication is decreased in nsp14 exoribonuclease mutants. *J Virol* **81**, 12135-12144
649   (2007).
650 43. J. B. Case, A. W. Ashbrook, T. S. Dermody, M. R. Denison, Mutagenesis of S-adenosyl-l-
651   methionine-binding residues in coronavirus nsp14 N7-methyltransferase demonstrates
652   differing requirements for genome translation and resistance to innate immunity. *J Virol*
653   **90**, 7248-7256 (2016).
654 44. Z. Zhang *et al.*, Live attenuated coronavirus vaccines deficient in N7-Methyltransferase
655   activity induce both humoral and cellular immune responses in mice. *Emerg Microbes*
656   *Infect* **10**, 1626-1637 (2021).
657 45. F. Ferron, H. J. Debat, A. Shannon, E. Decroly, B. Canard, A N7-guanine RNA cap
658   methyltransferase signature-sequence as a genetic marker of large genome, non-
659   mammalian Tobaniviridae. *NAR Genom Bioinform* **2**, lqz022 (2020).
660 46. P. T. Nga *et al.*, Discovery of the first insect nidovirus, a missing evolutionary link in the
661   emergence of the largest RNA virus genomes. *PLoS Pathog* **7**, e1002215 (2011).
662 47. A. Saberi, A. A. Gulyaeva, J. L. Brubacher, P. A. Newmark, A. E. Gorbalenya, A planarian
663   nidovirus expands the limits of RNA genome size. *PLoS Pathog* **14**, e1007314 (2018).
664 48. Y. Sun *et al.*, Yeast-based assays for the high-throughput screening of inhibitors of
665   coronavirus RNA cap guanine-N7-methyltransferase. *Antiviral Res* **104**, 156-164 (2014).
666 49. W. Aouadi *et al.*, Toward the identification of viral cap-methyltransferase inhibitors by
667   fluorescence screening assay. *Antiviral Res* **144**, 330-339 (2017).
668 50. R. Ahmed-Belkacem *et al.*, Synthesis of adenine dinucleosides SAM analogs as specific
669   inhibitors of SARS-CoV nsp14 RNA cap guanine-N7-methyltransferase. *Eur J Med Chem*
670   **201**, 112557 (2020).
671 51. R. He *et al.*, Potent and selective inhibition of SARS coronavirus replication by
672   aurintricarboxylic acid. *Biochem Biophys Res Commun* **320**, 1199-1203 (2004).
673 52. M. H. Barnes, P. Spacciapoli, D. H. Li, N. C. Brown, The 3'-5' exonuclease site of DNA
674   polymerase III from gram-positive bacteria: definition of a novel motif structure. *Gene*
675   **165**, 45-50 (1995).
676 53. Y. Zuo, M. P. Deutscher, Exoribonuclease superfamilies: structural analysis and
677   phylogenetic distribution. *Nucleic Acids Res* **29**, 1017-1026 (2001).
678 54. M. Byszewska, M. Smietanski, E. Purta, J. M. Bujnicki, RNA methyltransferases involved in
679   5' cap biosynthesis. *RNA Biol* **11**, 1597-1607 (2014).
680 55. B. P. S. Chouhan, S. Maimaiti, M. Gade, P. Laurino, Rossmann-Fold Methyltransferases:
681   Taking a "beta-Turn" around Their Cofactor, S-Adenosylmethionine. *Biochemistry* **58**, 166-
682   170 (2019).
683 56. N. H. Moeller *et al.*, Structure and dynamics of SARS-CoV-2 proofreading exoribonuclease
684   ExoN. *bioRxiv* 10.1101/2021.04.02.438274, 2021.2004.2002.438274 (2021).
685 57. R. A. Laskowski, SURFNET: a program for visualizing molecular surfaces, cavities, and
686   intermolecular interactions. *J Mol Graph* **13**, 323-330, 307-328 (1995).
687 58. L. D. Eckerle, S. M. Brockway, S. M. Sperry, X. Lu, M. R. Denison, Effects of mutagenesis of
688   murine hepatitis virus nsp1 and nsp14 on replication in culture. *Adv Exp Med Biol* **581**, 55-
689   60 (2006).

690 59. K. E. Medvedev, L. N. Kinch, N. V. Grishin, Functional and evolutionary analysis of viral
691   proteins containing a Rossmann-like fold. *Protein Sci* **27**, 1450-1463 (2018).
692 60. S. T. Rao, M. G. Rossmann, Comparison of super-secondary structures in proteins. *J Mol*
693   *Biol* **76**, 241-256 (1973).
694 61. G. Sutton, J. M. Grimes, D. I. Stuart, P. Roy, Bluetongue virus VP4 is an RNA-capping
695   assembly line. *Nat Struct Mol Biol* **14**, 449-451 (2007).
696 62. Y. Tao, D. L. Farsetta, M. L. Nibert, S. C. Harrison, RNA synthesis in a cage--structural
697   studies of reovirus polymerase lambda3. *Cell* **111**, 733-745 (2002).
698 63. M. P. Egloff, D. Benarroch, B. Selisko, J. L. Romette, B. Canard, An RNA cap (nucleoside-2'-
699   O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and
700   functional characterization. *EMBO J* **21**, 2757-2768 (2002).
701 64. M. M. Lai, S. A. Stohlman, Comparative analysis of RNA genomes of mouse hepatitis
702   viruses. *J Virol* **38**, 661-670 (1981).
703 65. A. L. van Vliet, S. L. Smits, P. J. Rottier, R. J. de Groot, Discontinuous and non-discontinuous
704   subgenomic RNA transcription in a nidovirus. *EMBO J* **21**, 6571-6580 (2002).
705 66. A. Seybert *et al.*, A complex zinc finger controls the enzymatic activities of nidovirus
706   helicases. *J Virol* **79**, 696-704 (2005).
707 67. L. A. Pearson *et al.*, Development of a High-Throughput Screening Assay to Identify
708   Inhibitors of the SARS-CoV-2 Guanine-N7-Methyltransferase Using RapidFire Mass
709   Spectrometry. *SLAS Discov* 10.1177/24725552211000652, 24725552211000652 (2021).
710 68. K. Devkota *et al.*, Probing the SAM Binding Site of SARS-CoV-2 Nsp14 In Vitro Using SAM
711   Competitive Inhibitors Guides Developing Selective Bisubstrate Inhibitors. *SLAS Discov* **26**,
712   1200-1211 (2021).
713 69. S. Basu *et al.*, Identifying SARS-CoV-2 antiviral compounds by screening for small molecule
714   inhibitors of Nsp14 RNA cap methyltransferase. *Biochem J* **478**, 2481-2497 (2021).
715 70. M. Bouvet *et al.*, Coronavirus Nsp10, a critical co-factor for activation of multiple
716   replicative enzymes. *J Biol Chem* **289**, 25783-25796 (2014).
717 71. L. Yan *et al.*, Coupling of N7-methyltransferase and 3'-5' exoribonuclease with SARS-CoV-
718   2 polymerase reveals mechanisms for capping and proofreading. *Cell* **184**, 3474-3485
719   e3411 (2021).
720 72. C. Liu *et al.*, Structural basis of mismatch recognition by a SARS-CoV-2 proofreading
721   enzyme. *Science* **373**, 1142-1146 (2021).
722 73. E. C. Smith, H. Blanc, M. C. Surdel, M. Vignuzzi, M. R. Denison, Coronaviruses lacking
723   exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading
724   and potential therapeutics. *PLoS Pathog* **9**, e1003565 (2013).
725 74. M. Gouy, S. Guindon, O. Gascuel, SeaView version 4: A multiplatform graphical user
726   interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**, 221-
727   224 (2010).
728 75. G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, WebLogo: a sequence logo generator.
729   *Genome Res* **14**, 1188-1190 (2004).
730 76. E. F. Pettersen *et al.*, UCSF Chimera--a visualization system for exploratory research and
731   analysis. *J Comput Chem* **25**, 1605-1612 (2004).
732 77. E. Jurrus *et al.*, Improvements to the APBS biomolecular solvation software suite. *Protein*
733   *Sci* **27**, 112-128 (2018).
734 78. Y. Dehouck, J. M. Kwasigroch, D. Gilis, M. Rooman, PoPMuSiC 2.1: a web server for the
735   estimation of protein stability changes upon mutation and sequence optimality. *BMC*
736   *Bioinformatics* **12**, 151 (2011).

737   79.   M. Hecht, Y. Bromberg, B. Rost, Better prediction of functional effects for sequence
738   variants. *BMC Genomics* **16 Suppl 8**, S1 (2015).
739   80.   Y. Wang *et al.*, Coronavirus nsp10/nsp16 methyltransferase can be targeted by nsp10-
740   derived peptide in vitro and in vivo to reduce replication and pathogenesis. *J Virol* **89**,
741   8416-8427 (2015).
742   81.   D. D. Nedialkova, A. E. Gorbalenya, E. J. Snijder, Arterivirus Nsp1 modulates the
743   accumulation of minus-strand templates to control the relative abundance of viral
744   mRNAs. *PLoS Pathog* **6**, e1000772 (2010).
745   82.   A. H. de Wilde *et al.*, MERS-coronavirus replication induces severe in vitro cytopathology
746   and is strongly inhibited by cyclosporin A or interferon-alpha treatment. *J Gen Virol* **94**,
747   1749-1760 (2013).
748   83.   B. K. Tischer, G. A. Smith, N. Osterrieder, En passant mutagenesis: a two step markerless
749   red recombination system. *Methods Mol Biol* **634**, 421-430 (2010).
750   84.   F. Almazan *et al.*, Engineering a replication-competent, propagation-defective Middle
751   East respiratory syndrome coronavirus as a vaccine candidate. *mBio* **4**, e00650-00613
752   (2013).
753   85.   H. H. Rabouw *et al.*, Middle East respiratory coronavirus accessory protein 4a inhibits PKR-
754   mediated antiviral stress responses. *PLoS Pathog* **12**, e1005982 (2016).
755   86.   S. Pfefferle *et al.*, Reverse genetic characterization of the natural genomic deletion in
756   SARS-Coronavirus strain Frankfurt-1 open reading frame 7b reveals an attenuating
757   function of the 7b protein in-vitro and in-vivo. *Virol J* **6**, 131 (2009).
758   87.   A. R. Fehr *et al.*, The nsp3 macrodomain promotes virulence in mice with coronavirus-
759   induced encephalitis. *J Virol* **89**, 1523-1536 (2015).
760   88.   T. T. N. Thao *et al.*, Rapid reconstruction of SARS-CoV-2 using a synthetic genomics
761   platform. *Nature* 10.1038/s41586-020-2294-9 (2020).
762   89.   S. H. van den Worm *et al.*, Reverse genetics of SARS-related coronavirus using vaccinia
763   virus-based recombination. *PLoS One* **7**, e32857 (2012).
764   90.   F. Weber, V. Wagner, S. B. Rasmussen, R. Hartmann, S. R. Paludan, Double-stranded RNA
765   is produced by positive-strand RNA viruses and DNA viruses but not in detectable
766   amounts by negative-strand RNA viruses. *J Virol* **80**, 5059-5064 (2006).
767   91.   M. J. van Hemert *et al.*, SARS-coronavirus replication/transcription complexes are
768   membrane-protected and need a host factor for activity in vitro. *PLoS Pathog* **4**, e1000054
769   (2008).

770
771

772 **Figures and Tables**
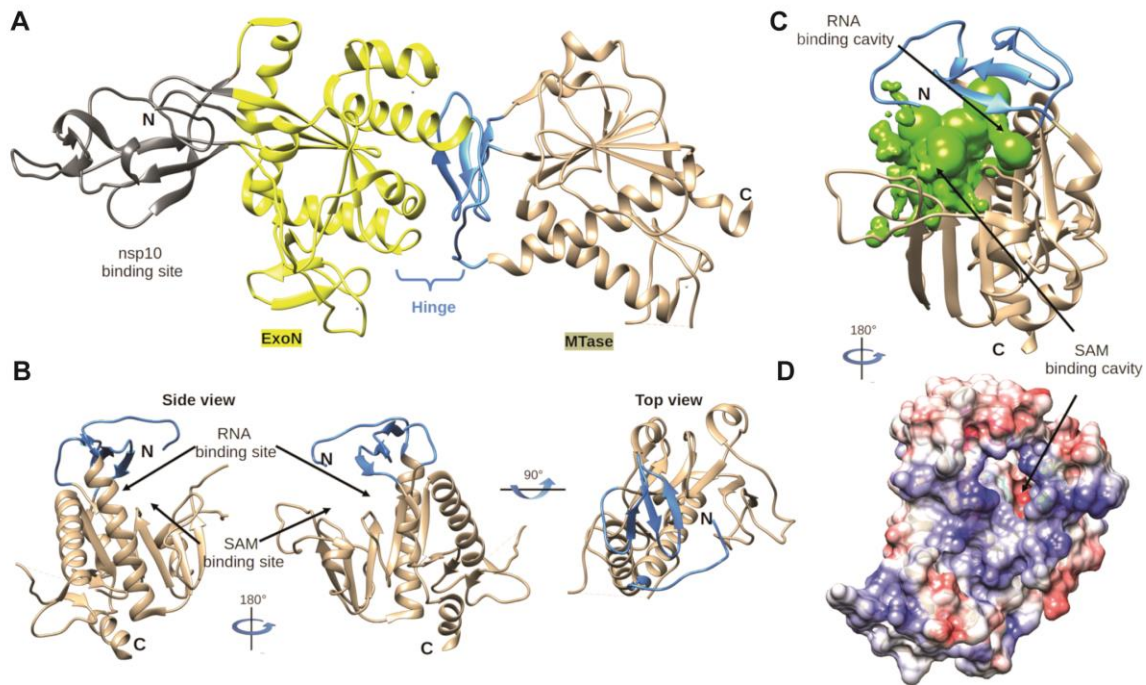
773
774 **Figure 1. Global architecture of coronavirus nsp14.** (A) Architecture of SARS-CoV nsp14 (PDB
775 5NFY) showing the nsp10 binding site (grey), N-terminal ExoN domain (yellow), hinge subdomain
776 (blue), and C-terminal N7-MTase domain (brown). (B) Side and top view of the hinge region and
777 N7-MTase domain. The three strands of the hinge (blue) protrude from the N7-MTase domain
778 (brown). (C) Analysis of the volume of the N7-MTase active site, with the cavity highlighted in green
779 and hinge subdomain in blue. (D) Electrostatic surface representation of the CoV nsp14 hinge
780 region and N7-MTase domain. Surface electrostatic potential calculated by Adaptive Poisson-
781 Boltzmann Solver, from - 10 (red) to + 10 (blue) kT/e.
782
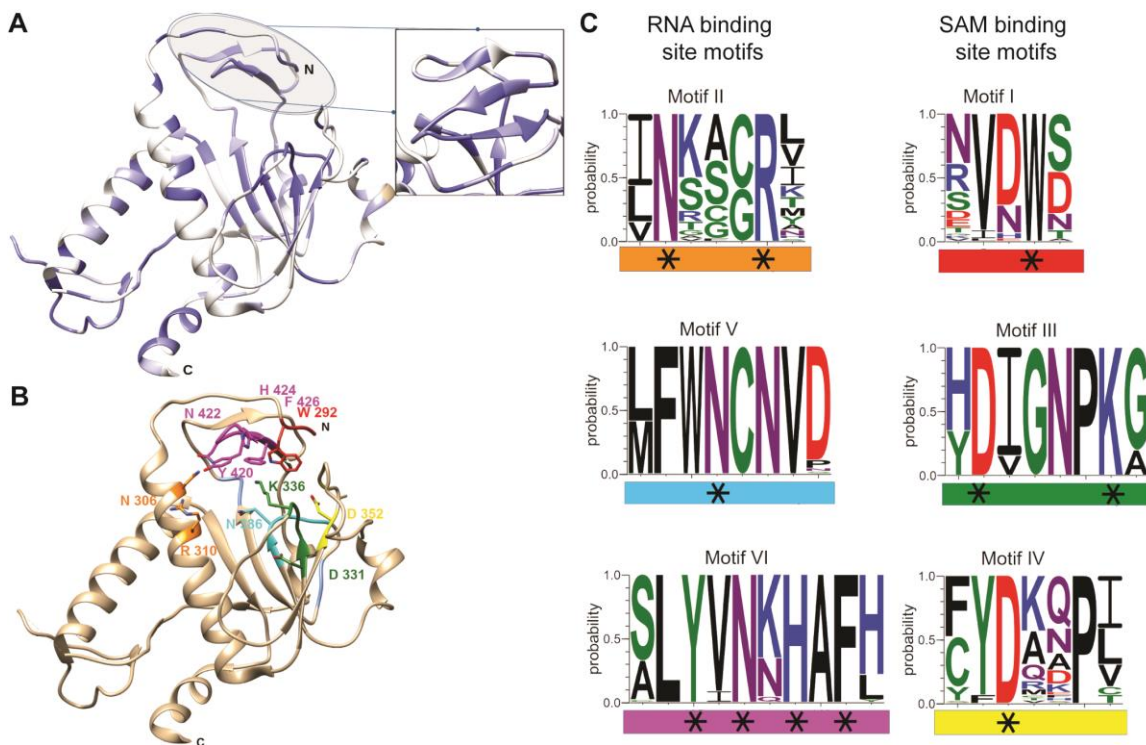
783
Figure 2. Coronavirus-wide nsp14 N7-MTase conservation and structural analysis. (A) CoV nsp14 amino acid sequence conservation plotted on the structure (PDB 5NFY) of the SARS-CoV hinge region and N7-MTase domain (dark blue to white shading representing 100% to less than 50% sequence identity). A 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.

793
794
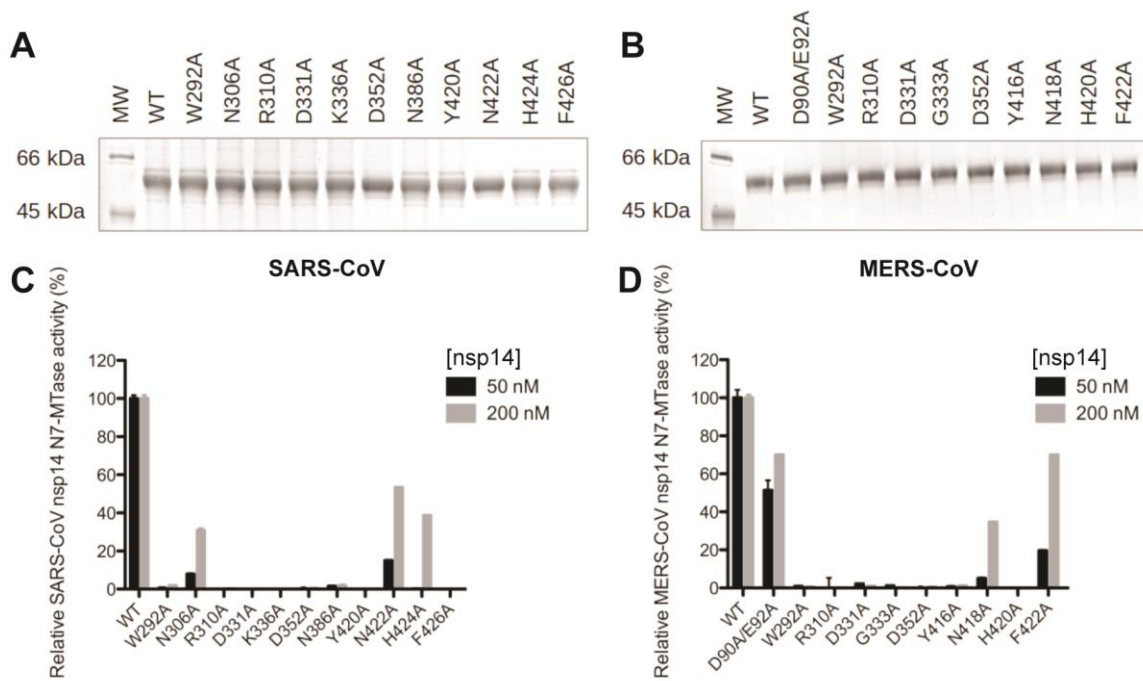
Figure 3. Expression and *in vitro* N7-MTase activity of SARS-CoV and MERS-CoV nsp14 mutants. Recombinant SARS-CoV (A) and MERS-CoV (B) wild-type and mutant nsp14 proteins were expressed in *E. coli* and purified. Proteins were loaded (2 µg and 1 µg for SARS-CoV and MERS-CoV, respectively) and analyzed 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.
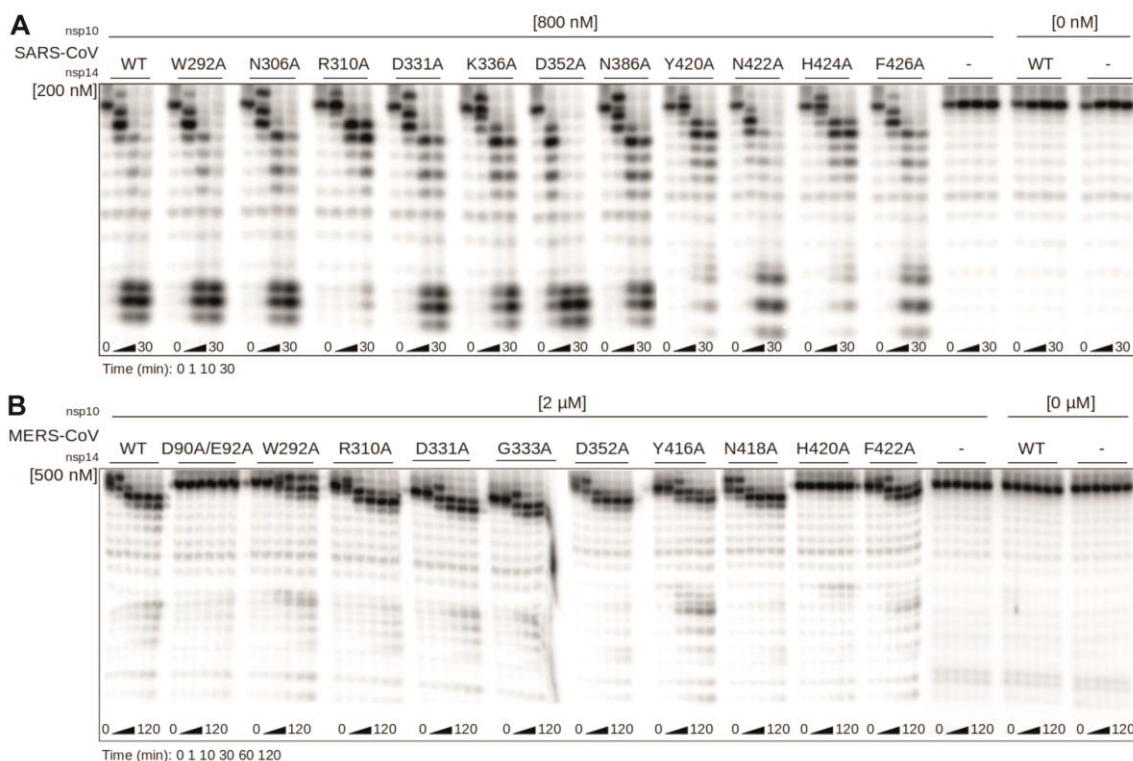
Figure 4. *In vitro* **exoribonuclease activity of SARS-CoV and MERS-CoV N7-MTase mutants**.
The *in vitro* ExoN activity of SARS-CoV (A) and MERS-CoV (B) mutant nsp14 proteins (Fig. 3) was
determined by monitoring the degradation of a 5' radiolabeled RNA substrate (see Methods). An
nsp14 concentration 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.

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§			--								++/-	--

817
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.