8822–8835 *Nucleic Acids Research, 2021, Vol. 49, No. 15* *Published online 5 August 2021*
https://doi.org/10.1093/nar/gkab677

NMPylation and de-NMPylation of SARS-CoV-2 nsp9 by the NiRAN domain

Bing Wang [1,2], Dmitri Svetlov [3] and Irina Artsimovitch [1,2,*]

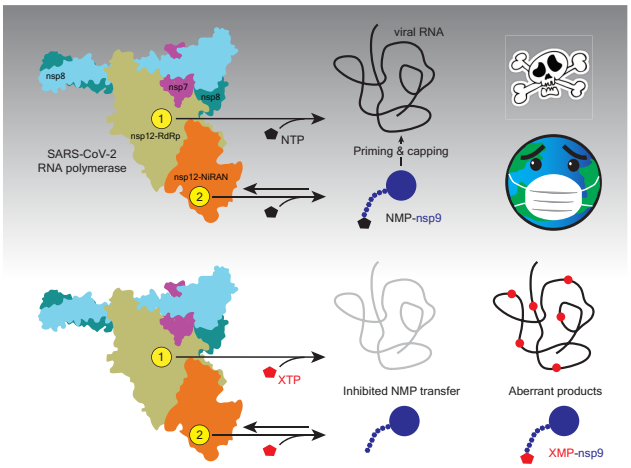
[1]Department of Microbiology, The Ohio State University, Columbus, OH 43210, USA, [2]The Center for RNA Biology, The Ohio State University, Columbus, OH 43210, USA and [3]Svetlov Scientific Software, Pasadena, CA 91106, USA

Received June 13, 2021; Revised July 15, 2021; Editorial Decision July 16, 2021; Accepted July 26, 2021

ABSTRACT

The catalytic subunit of SARS-CoV-2 RNA-dependent RNA polymerase (RdRp) contains two active sites that catalyze nucleotidyl-monophosphate transfer (NMPylation). Mechanistic studies and drug discovery have focused on RNA synthesis by the highly conserved RdRp. The second active site, which resides in a Nidovirus RdRp-Associated Nucleotidyl transferase (NiRAN) domain, is poorly characterized, but both catalytic reactions are essential for viral replication. One study showed that NiRAN transfers NMP to the first residue of RNA-binding protein nsp9; another reported a structure of nsp9 containing two additional N-terminal residues bound to the NiRAN active site but observed NMP transfer to RNA instead. We show that SARS-CoV-2 RdRp NMPylates the native but not the extended nsp9. Substitutions of the invariant NiRAN residues abolish NMPylation, whereas substitution of a catalytic RdRp Asp residue does not. NMPylation can utilize diverse nucleotide triphosphates, including remdesivir triphosphate, is reversible in the presence of pyrophosphate, and is inhibited by nucleotide analogs and bisphosphonates, suggesting a path for rational design of NiRAN inhibitors. We reconcile these and existing findings using a new model in which nsp9 remodels both active sites to alternately support initiation of RNA synthesis by RdRp or subsequent capping of the product RNA by the NiRAN domain.

GRAPHICAL ABSTRACT



INTRODUCTION

Coronaviruses (CoVs) are single-stranded positive-sense (+) RNA viruses that constitute the *Coronaviridae* family in the order *Nidovirales* (1). CoVs cause many respiratory and gastrointestinal infections in humans, from mild common colds to severe respiratory diseases, including the ongoing COVID-19 pandemic (2,3). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the etiological agent of COVID-19, is the third zoonotic CoV to have caused a major disease outbreak in humans in the last two decades (4). The prevalence of CoVs in animal reservoirs argues that future viral pandemics are all but certain (3,5) and makes advance preparations imperative. While the availability of effective vaccines against SARS-CoV-2 has been a game changer for the current COVID-19 pandemic, broad-spectrum antiviral drugs are needed to protect unvaccinated and immunocompromised individuals and to buy time needed for the development of new vaccines at the onset of the next viral epidemic.

CoVs have very large (approximately 30 kb) genomes that encode non-structural proteins (nsps) required for viral gene expression and replication. Upon infecting human

*To whom correspondence should be addressed. Tel: +1 614 292 6777; Email: artsimovitch.1@osu.edu

cells, the SARS-CoV-2 RNA genome is translated to produce a long polyprotein that is cleaved into nsps 1 through 16 by the viral protease nsp5 (6). The catalytic subunit of RNA-dependent RNA polymerase (RdRp), nsp12, stands out as the only protein that is present in all RNA viruses (7) and is therefore an attractive target for broad-spectrum antivirals (8). The high degree of conservation of the RdRp structure and its key catalytic elements (7,9) encouraged efforts to repurpose existing antivirals, such as remdesivir (10,11) and favipiravir (12), for the treatment of COVID-19 (13). However, these pursuits have not yet produced an effective clinical treatment, suggesting that a detailed mechanistic analysis of the viral replication cycle may be required to identify the best points for intervention.

SARS-CoV-2 RdRp holoenzyme is a four-subunit complex of the catalytic nsp12 and accessory nsp7 and nsp8 proteins (14–16), nsp12•7•8_2 (Figure 1A). The holoenzyme binds to two copies of the superfamily 1 helicase nsp13 (17) and associates with a proofreading exonuclease nsp14, capping enzymes, and other proteins to form large multi-subunit replication-transcription complex (RTC) that mediates synthesis and modification of viral RNAs (6). The nsp12 subunit also contains a second catalytic module, an N-terminal 250-residue Nidovirus RdRp-Associated Nucleotidyl transferase (NiRAN) domain (9). The NiRAN domain displays significant sequence divergence as compared to RdRp, with only four conserved motifs (preA_N, A_N, B_N and C_N) comprising the NiRAN signature (18). The NiRAN domain is present in all nidoviruses but has no homologs in other RNA viruses and, together with the nsp13 helicase (HELD) domain, is a genetic marker for the order *Nidovirales* (18).

As first shown with an equine arteritis virus (EAV) enzyme from the *Arteriviridae* family of nidoviruses, RdRp self-NMPylates *in vitro* with a clear preference for UTP as a substrate and the Mn^{2+} ion as a cofactor (18). Such an activity is common among AMPylases, which frequently transfer AMP to their autoinhibitory domains (19). Substitutions of several conserved NiRAN residues abolish nucleotidyl transfer *in vitro* and abrogate EAV replication in cell culture to the same extent as do substitutions of the catalytic RdRp residues (18). These results demonstrate that the NiRAN domain plays a critical role in the viral life cycle and thus is a valid target of antiviral drug discovery. Subsequent studies of two viruses from *Coronaviridae*, HCoV-229E and SARS-CoV-2, led to similar conclusions (20).

The location of the NiRAN nucleotidyl transfer site is well established, but the identity of the NMP acceptor remains debated. Single-particle cryogenic electron microscopy (cryoEM) studies of SARS-CoV-2 RdRp (17,21) revealed nucleotides bound to nsp12 residues shown to be required for self-NMPylation *in vitro* (18). A finding that NiRAN active site is structurally homologous to that of a protein pseudokinase, selenoprotein O/SelO (16,17,22), supported the proposed role of NiRAN in covalent NMPylation of protein targets (18). Consistently, Ziebuhr and colleagues recently showed that HCoV-229E and SARS-CoV-2 nsp12s efficiently transfer NMPs to nsp9 (20), a small (113 residues) RNA-binding protein that is essential for viral replication (23–25). nsp9 and nsp12 modifications shared the requirements for NTP substrates, metal cofactors, and NiRAN residues, arguing that both reactions utilize similar mechanisms. Mass spectrometry identified the primary amine of the N-terminal Asn, which is conserved among CoVs (Figure 1B), as a site of nsp9 modification (20). Mutational analysis revealed that (*i*) the Asn2 residue was critical for modification; (*ii*) Asn1 could be substituted with Ala or Ser with a modest loss of reactivity; and (*iii*) the presence of even one additional N-terminal Ala residue abolished nsp9 NMPylation (20). In support of the essential role of its NMPylation, these nsp9 substitutions had parallel effects on NMP transfer *in vitro* and on viral replication (20). Furthermore, in a cryoEM structure of nsp9 bound to the SARS-CoV-2 RdRp-helicase complex (21), the nsp9 Asn1 is adjacent to the NiRAN active site (Figure 1A). Strikingly, however, Yan *et al.* did not detect nsp9 modification and instead observed GMP transfer to RNA, which they proposed represents a key early step in the capping pathway (21). The lack of nsp9 reactivity is most likely explained by the presence of two additional, non-native residues, Gly and Ser, at the N-terminus of the recombinant nsp9 used to obtain the structure. While these residues were not modeled in PDB: 7CYQ, the GSNNELSPVALR tryptic peptide was identified by mass-spectrometry analysis and the density for Gly-2/Ser-1 residues is discernible in the EM map (Figure 1A). In the presence of these additional residues, the cognate NMPylation site, the N1 amine, is eliminated. Different metal ion cofactors, protein tags, or other reaction variables could also explain discrepancies in observed nsp12 catalytic properties.

Our findings that noncognate NTPs and nucleoside analogs modulate RdRp activity suggested that the RdRp and NiRAN active sites (thereafter referred to as AS1 and AS2, respectively) could be allosterically connected (26). Testing this hypothesis necessitates parallel assays of both nucleotidyl transfer activities and in turn requires using cognate NiRAN substrates. In agreement with HCoV-229E studies by Ziebuhr and colleagues (20), we show that SARS-CoV-2 nsp12 efficiently NMPylates nsp9 that has the native N-terminus, but not an nsp9 variant that bears two additional N-terminal residues. Substitutions of the invariant NiRAN residues abolished nsp9 NMPylation, whereas substitution of a catalytic RdRp residue, Asp760, did not. We found that NMPylation proceeds equally efficiently with Mg^{2+} and Mn^{2+}, is largely insensitive to the identity of the natural NTP and can utilize nucleotide analogs such as remdesivir triphosphate. We also show that NMPylation is reversible in the presence of pyrophosphate. Nucleotide analogs that lack the triphosphate moiety and pyrophosphate analogs bisphosphonates inhibit the forward reaction, suggesting a starting point for identification of NiRAN inhibitors.

MATERIALS AND METHODS

Construction of expression vectors

Plasmids used in this study are shown in Supplementary Table S1. The SARS-CoV-2 *nsp7/8/9/12* genes were codon-optimized for expression in *Escherichia coli* and synthesized by GenScript and subcloned into standard pET-derived expression vectors under control of the T7 gene 10 promoter and *lac* repressor, as described previously (26). The

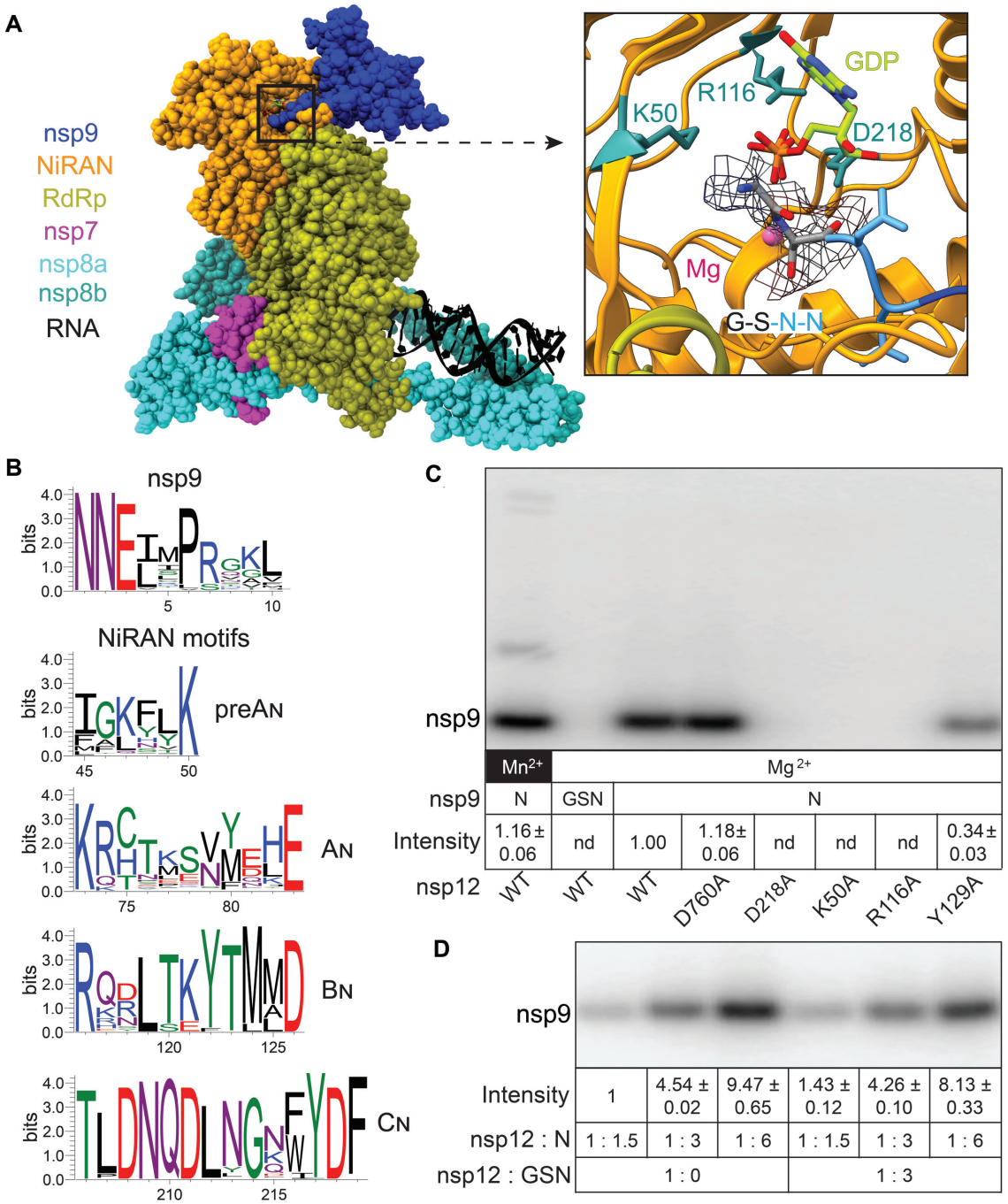


Figure 1. NiRAN-mediated NMPylation of nsp9. (**A**) Cryo-EM structure of the SARS-CoV-2 RdRp-helicase complex bound to a non-native nsp9 with a two-residue extension at the N-terminus (21). Left. Overall structure of the complex; PDB: 7CYQ; nsp13 helicase is not shown. Proteins are shown as colored molecular surfaces (as shown in the key) and RNA as black cartoon. The color coding corresponds to the figures throughout this manuscript unless otherwise specified. Right. Zoom in on the active site of the NiRAN domain (AS2) with GDP-Mg^{2+} (lime carbon atoms and magenta sphere, respectively). Side chains of key conserved residues from the pre-A_N, B_N and C_N motifs that were substituted in this work are shown as sticks. Four N-terminal residues of nsp9 (GSNN) are shown; the cryo-EM difference density for Gly and Ser residues from EMDB: 30504 is shown (gray mesh). Structural figures were prepared with Coot (61), UCSF ChimeraX 1.2 and PyMOL Molecular Graphics System, version 2.4.1, Schrodinger, LLC. (**B**) Conservation of residues at the N-terminus of nsp9 and four conserved NiRAN motifs in alpha-, beta-, gamma- and deltacoronavirus genera. (**C**) Mutations in the NiRAN active site and nsp9 N-terminal GS extension abolish NMP transfer, but Mn^{2+} is dispensable. NMPylation efficiency was compared to that observed with the wild-type nsp12 in the presence of 1 mM Mg^{2+} (set at 1) and is shown as mean ± SD ($n = 3$); nd, no signal detected above the background. (**D**) NMPylation of nsp9 is not inhibited by an excess of the artificially-extended ^{GSN}nsp9. NMPylation efficiency was compared to that observed with ^{N}nsp9 present at 1.5-fold molar excess over nsp12 in the absence of ^{GSN}nsp9 (set at 1) and is shown as mean ± SD ($n = 3$).

derivative plasmids were constructed by standard molecular biology approaches with restriction and modification enzymes from New England Biolabs. DNA oligonucleotides for vector construction and sequencing were obtained from Millipore Sigma, synthetic DNA fragments for Gibson Assembly—from IDT. The sequences of all plasmids were confirmed by Sanger sequencing at the Genomics Shared Resource Facility (The Ohio State University). All plasmids were deposited to Addgene.

Protein expression and purification

The expression and purification of nsp7/8/12 and nsp12 mutants are described in our previous study (26). All purification steps were carried out at 4°C. nsp9 variants were overexpressed in *E. coli* BL21 (DE3) cells (Novagen, Cat#69450). Cells were grown in lysogenic broth (LB) with kanamycin (50 μg/ml). Cells were cultured at 37°C to an OD_{600} of 0.6–0.8 and the temperature was lowered to 16°C. Expression was induced with 0.1 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG; Goldbio, Cat#I2481C25) for 18 hours. Induced cells were harvested by centrifugation (6000 × *g*), resuspended in lysis buffer (50 mM HEPES, pH 7.5, 300 mM NaCl, 5% glycerol (v/v), 1 mM Phenylmethylsulfonyl fluoride (PMSF; ACROS Organics, Cas#329-98-6), 5 mM β-ME, 10 mM imidazole), and lysed by sonication. The lysate was cleared by centrifugation (10 000 × *g*). The soluble protein was purified by absorption to Ni^{2+}-NTA resin (Cytiva, Cat#17531801), washed with Ni-buffer A (50 mM HEPES, pH 7.5, 300 mM NaCl, 5% glycerol, 5 mM β-ME, 50 mM imidazole), and eluted with Ni-buffer B (50 mM HEPES, pH 7.5, 300 mM NaCl, 5% glycerol, 5 mM β-ME, 500 mM imidazole). The eluted protein was further loaded onto a desalting column (Cytiva, Cat#17508701) in desalting buffer (50 mM HEPES, pH 7.5, 5% glycerol, 5 mM β-ME, 300 mM NaCl) to remove imidazole. The fusion protein was treated with TEV protease (for pIA1364) or SUMO protease (for pIA1414) at 4°C overnight. The sample was supplemented with 20 mM imidazole and passed through Ni^{2+}-NTA resin. The untagged protein was loaded onto a Sephacryl S-100 HR column (Cytiva, Cat#17116501) in desalting buffer. Peak fractions were assessed by SDS-PAGE and Coomassie staining. Purified protein was dialyzed into storage buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 45% glycerol, 5 mM β-ME), aliquoted and stored at −80°C.

Conservation analysis

A total of 75 viral genomes, representing alpha-, beta-, gamma- and deltacoronaviruses, were fetched from Uniprot reference Proteomes (version 2021_02) (27). Multiple sequence alignment (MSA) was done by MAFFT (version 7) (28). WebLogo (version 3) (29) was used to generate the sequence logos based on the MSA.

nsp9 NMPylation

All NMPylation assays were carried out at 37°C. For standard NMPylation assay, 0.5 μM nsp12 and 5 μM nsp9 were incubated in NMPylation buffer (25 mM HEPES, pH 7.5, 15 mM KCl, 5% glycerol, 2 mM $MgCl_2$, 2 mM DTT) for 5 min, then 25 μM GTP and 10 μCi [α^{32}P]-GTP (PerkinElmer, Cat#BLU006H250UC) were added to start the reaction. After a further 10 min incubation, the reaction was mixed with 4× NuPAGE™ LDS Sample Buffer (ThermoFisher, Cat#NP0007).

Competition assays

To assess the competition of ^{GSN}nsp9 with ^{N}nsp9, nsp12 was incubated with ^{GSN}nsp9 for 5 min. Then ^{N}nsp9 (at desired concentrations), 25 μM GTP, and 10 μCi [α^{32}P]-GTP were added to start the reaction. For nucleotide competition assays, NTPs, NTP analogs, and pyrophosphate were used at 0.5 mM. After incubation of nsp12/9 for 5 min in NMPylation buffer, the competitor NTPs (Cytiva, Cat#27202501), GDP (Sigma-Aldrich, Cat#G7127), GMP (Sigma-Aldrich, Cat#G8377), Inosine-5′-Triphosphate (ITP; TriLink Biotechnologies, Cat#N-1020), ppGpp (Trilink Biotechnologies, Cat#N-6001), GpCpp (Jena Bioscience, Cat#NU-405S), Remdesivir triphosphate (RTP; MedChemExpress, Cat#GS443902), and pyrophosphate (PP_i; Sigma-Aldrich, Cat#71515) was added together with 25 μM GTP and 10 μCi [α^{32}P]-GTP to start the reaction. After a further 10 min incubation, the reactions were stopped with LDS Sample Buffer as above.

De-NMPylation

To determine if PP_i can reverse the NMPylation reaction, 0.5 μM nsp12, 5 μM nsp9, 25 μM GTP and 10 μCi [α^{32}P]-GTP were incubated in NMPylation buffer for 15 min, then 0.5 mM PP_i was added. To determine which active site of nsp12 is responsible for the de-NMPylation activity, 0.5 μM His-tagged nsp12, 20 μM nsp9, 50 μM GTP, and 10 μCi [α^{32}P]-GTP were incubated in NMPylation buffer (2 mM DTT in the buffer was replaced by 2 mM β-mercaptoethanol) for 20 min. Dynabeads (ThermoFisher, Cat#10103D) were added to remove the His-tagged nsp12, followed by adding 0.5 μM nsp12 variants and 0.5 mM PP_i. Samples were quenched at indicated time points and analyzed by electrophoresis.

Inhibition by bisphosphonates

0.5 μM nsp12 and 5 μM nsp9 were incubated with different concentrations of Risedronate (Sigma-Aldrich, Cat#PHR1888) or Foscarnet (Sigma-Aldrich, Cat#PHR1436) in the NMPylation buffer for 5 min, then 25 μM GTP and 10 μCi [α^{32}P]-GTP were added to start the reaction. Reactions were performed for 10 min.

RNA extension and cleavage

An RNA oligonucleotide (5′-UUUUCAUGCUACGCG UAGUUUUCUACGCG-3′; 4N) with Cyanine 5.5 at the 5′-end was obtained from Millipore Sigma (USA); this RNA hairpin serves as both the primer and the template (15). The RNA scaffold was annealed in 20 mM HEPES, pH 7.5, 50 mM KCl by heating to 75°C and then gradually cooling to 4°C. To test RdRp activity, reactions were

carried out at 37°C with 500 nM nsp12 variants, 1 μM nsp7, 1.5 μM nsp8, 250 nM RNA and 250 μM NTPs in the transcription buffer (20 mM HEPES, pH 7.5, 15 mM KCl, 5% glycerol, 1 mM $MgCl_2$, 2 mM DTT) for 20 min at 37°C. For pyrophosphorolysis, holo RdRp was preincubated with the RNA scaffold at 37°C for 5 min in the transcription buffer; then the indicated combinations of PP_i and NTPs were added. Reactions were stopped by adding 2× stop buffer (8 M Urea, 20 mM EDTA, 1× TBE, 0.2% bromophenol blue).

Sample analysis

Protein samples were heated for 5 min at 95°C and separated by electrophoresis in NuPAGE™ 4–12% gels (ThermoFisher, Cat# NP0329BOX). RNA samples were heated for 2.5 min at 95°C and separated by electrophoresis in denaturing 9% acrylamide (19:1) gels (7 M Urea, 0.5× TBE). The gels were visualized and quantified using Typhoon FLA9000 (GE Healthcare) and ImageQuant. All assays were carried out in triplicates. The means and standard deviation (SD) were calculated by Excel (Microsoft).

RESULTS

NMPylation requires the native N-terminus of SARS-CoV-2 nsp9

Impressive progress in the structural studies of SARC-CoV-2 transcription machinery, reviewed in (9), far outpaces its functional analysis. The presence of two active sites that utilize the same NTP substrates in nsp12 complicates mechanistic analyses of RdRp, yet also provides means to assess overall 'quality' of a newly purified nsp12 variant (Supplementary Figure S1). A substitution that leads to large defects in both nucleotidyl transfer activities likely triggers gross protein misfolding because AS1 and AS2 are located very far apart and substitutions that abolish catalysis in one active site do not have reciprocal effects on the other (20,26). Using their cognate substrates is essential for the analysis of both nucleotidyl transfer activities and identical conditions should be employed if possible. In our experiments, we used standard solution conditions that support efficient RNA synthesis (Supplementary Figure S1), [α^{32}P]-GTP, which supports efficient NMPylation (18,20) as an NMP donor, and SARS-CoV-2 nsps containing native N- and C-termini (shown in Supplementary Figure S2).

Low efficiency of nsp12 self-NMPylation and its dependence on non-physiological concentrations of Mn^{2+} ion (18,20,30) suggested to us that this reaction may be fortuitous. Furthermore, a recent study identified multiple sites of NMPylation in nsp7 and nsp12 using mass spectrometry (30). By contrast, a nearly complete, although still Mn-dependent, modification of nsp9 and additional functional data reported by Slanina *et al.* strongly support a model in which nsp9 is a genuine NiRAN target (20). We first assayed NMP transfer of nsp12 alone using two nsp9 proteins: a variant with the native N-terminus (^{N}nsp9; confirmed by MS analysis; see below) and an extended nsp9 with two additional residues at –1 and –2 (^{GSN}nsp9), identical to that used in (21). These recombinant proteins were produced by cleavage of tagged nsp9 precursors by Ubiquitin-like-specific protease 1 (Ulp1) and Tobacco Etch Virus (TEV) proteases, respectively. We observed efficient GMP transfer to ^{N}nsp9 but not to ^{GSN}nsp9 by the wild-type (WT) nsp12 (Figure 1C). NMPylation was abrogated by substitutions of conserved NiRAN residues (K50A in $preA_N$, R116A in B_N and D218A in C_N; Figure 1B) that inhibit viral replication in cell culture (20), but not by the D760A substitution in AS1. As expected, NiRAN mutants did not affect RNA synthesis, whereas the D760A variant was inactive (Supplementary Figure S1). The Y129A substitution at the NiRAN/RdRp domain interface modestly reduced both activities (Figure 1C and Supplementary Figure S1), suggesting that the mutant protein could have an altered fold. We conclude that, as shown for the HCoV-229E system (20), SARS-CoV-2 nsp9 that has the native N terminus is NMPylated by AS2.

Under our reaction conditions (25 μM GTP, 10 min incubation at 37°C, 10-fold molar excess of nsp9), approximately 25% of nsp9 was GMPylated (Supplementary Figure S3A). This corresponds to 2.5 molecules of GMP-nsp9 per one molecule of nsp12, the expected ratio of proteins generated during translation of the viral genome (31). The modification efficiency could be affected in the presence of other RTC components and physiological solutes and is expected to increase dramatically at physiological NTP concentrations: the fraction of GMP labeling of SARS-CoV nsp7 + 8 (at 6 mM Mn^{2+}) increased from 0.01% at 0.2 μM GTP to 1% at 200 μM GTP (30).

Our findings and those of Slanina *et al.* (20) underscore the potential importance of native termini to protein function. The N- and C-termini are commonly modified to include purification tags, an approach that is justified when these ends are phylogenetically variable and make no functional interactions. However, in the context of CoV protein maturation, the 'correct' ends are generated upon proteolytic cleavage of the polyproteins by the viral protease. Given that the very first residue of nsp9 is the target of NMPylation, the potential importance of the identity of this residue is obvious, as is the prudence of preserving its native identity in experiments unless and until that identity is proven to be unimportant.

NMPylation occurs in the presence of nsp7/8 cofactors and does not require Mn^{2+}

To match the previously published conditions, we carried out GMPylation assays with nsp12 alone. To ascertain that this activity is preserved in context of the transcribing RdRp holoenzyme (nsp12•7•8_2), we repeated our assays in the presence of nsp7, nsp8, and an RNA scaffold under conditions that support robust RNA extension by SARS-CoV-RdRp (9,26). Our results demonstrate comparable nsp9 modification by nsp12 alone or as part of an active transcription complex (Supplementary Figure S3B).

Unlike that of its structural homolog SelO (22), the NiRAN domain's activity was thought to be dependent on the Mn^{2+} ion, at least for the EAV and HCoV-229E RdRps (18,20). Surprisingly, we observed equally efficient GMPylation in the presence of 1 mM Mg^{2+} or 1 mM Mn^{2+} (Figure 1C), concentrations that correspond to physiological

levels of Mg^{2+} but exceed those of Mn^{2+} (32). Mg^{2+} is the major cellular cofactor in electrophilic catalysis, in part due to its superior bioavailability and environmental abundance (32). Although Mn^{2+} can also function as the cofactor for the nucleotidyl transfer reaction for diverse nucleic acid polymerases, Mn^{2+} binding alters the active site geometry (32) to promote base misincorporation (33) and other inefficient reactions (34). The Mn^{2+} ion overrides a requirement for a canonical signal to fortuitously activate cyclic GMP-AMP [cGAMP] synthase (35) and can resuscitate a catalytically-compromised RNA polymerase II (36). Although Mn^{2+} and Mg^{2+} can support similar octahedral coordination in the active site (32), Mn^{2+} has also been observed to form a strikingly different network, independent of the catalytic triad residues (35). Consistent with the Mn^{2+}-induced gain-of-function, we observed GMP transfer to BSA in the presence of 1 mM Mn^{2+} (Supplementary Figure S4) and Mn^{2+}–dependent modifications of the accessory nsp7 and nsp8 subunits have been reported (30,37). These transfer reactions are very inefficient, $<<1\%$ at low GTP concentrations (30), when compared to NMPylation of nsp9 (Supplementary Figure S3A)and could be easily mistaken for the 'cognate' modification when observed in the absence of a real target. These large differences in transfer efficiencies could explain why we did not detect GMP transfer to nsp7 or nsp8 with either metal cofactor.

We do not know why only the Mn^{2+}-dependent NMP transfer was observed with EAV and HCoV-229E RdRps since we used very similar reaction conditions (18,20). We speculate that differences in RdRp folding could explain Mn^{2+} dependence. CoV RdRps are highly dynamic enzymes that undergo large conformational changes during the transcription cycle (38) and can become misfolded during expression in heterologous hosts (26). In particular, the NiRAN domain has been captured in different conformational states in cryoEM structures and becomes more ordered upon ligand binding to the active site (9,14,17,39,40).

nsp9 modification is not required for its release from nsp12

Enzymes that mediate protein NMPylation frequently have low affinity for their targets, necessitating covalent linkage of enzyme:substrate complexes for structural analysis (41). The modified nsp9 appears to be released from SARS-CoV-2 RdRp (Supplementary Figure S3A) and multi-round NMPylation of HCoV-229E nsp9 has also been reported (20). The formation of an apparently stable complex between SARS-CoV-2 RdRp and the modification-resistant ^{GSN}nsp9 visualized by cryoEM (21) raises a possibility that NMPylation is a prerequisite for nsp9 release.

To test this idea, we used competition between the native ^{N}nsp9 and the extended ^{GSN}nsp9 variant (Figure 1C). We found that preincubation of nsp12 with a three-fold molar excess of ^{GSN}nsp9 only slightly inhibited modification of ^{N}nsp9 (Figure 1D). Although it is possible that NMPylation alters nsp9 affinity for RdRp, a possibility that we intend to evaluate in the future, we conclude that free nsp9 is in a dynamic equilibrium with the nsp9•12 complex regardless of the presence of the N-terminal nucleotide moiety.

The role(s) of nsp9 in the viral life cycle remains to be elucidated. Substitutions of nsp9 residues that interact with nsp12 (21) abolish viral replication (24), an effect that has been attributed to a loss of nsp9 dimerization observed in structural studies (23,25) but is at least equally likely to be due to the loss of nsp9 binding to RdRp. The essentiality of nsp9 modification for viral replication (20) makes NMPylation a valid target for inhibition. Our results indicate that peptidomimetic compounds that resemble the nsp9 N-terminus are unlikely to serve as efficient inhibitors of NMPylation. However, substrate analogs that bind to AS2 may either interfere with NMP transfer to nsp9 or lead to modified but non-functional nsp9.

SARS-CoV-2 NiRAN can bind diverse nucleotides

Previous studies of NMPylation revealed differences in substrate utilization by different RdRps, which could be expected based on significant sequence divergence of the NiRAN domains (Figure 1B) that reflects a long evolutionary history of *Nidovirales* (42). For example, the His75 residue in the A_N motif, which contacts ADP•AlF_3 in the SARS-CoV-2 RdRp-helicase structure (17), is represented by Cys in HCoV-229E and by Val in EAV NiRAN domains. EAV RdRp displayed a strong preference for UTP, followed by GTP, and ATP and CTP were barely used (18), while HCoV-229E RdRp utilized all NTPs with preference for UTP (20). Structures of SARS-CoV-2 transcription complexes with NiRAN-bound nucleotides do not reveal any base-specific contacts (Figure 2A), suggesting that all NTPs would be used as substrates for NMPylation, and direct transfer of GMP and UMP to protein has been demonstrated by mass spectrometry (30). Consistently, competition experiments in which [α^{32}P]-GMP transfer to nsp9 was assayed in the presence of cold NTPs show that while GTP and UTP are marginally more effective competitors, the differences among all NTPs are small (Figure 2B). Thus, unlike EAV and HCoV-229E RdRp, the SARS-CoV-2 NiRAN does not appear to have a strong substrate preference. The assay design may also contribute to the observed discrepancies. Commercial radiolabeled NTP preparations contain impurities that compromise some sensitive assays, in contrast to high-purity NTPs (see Methods) that we use for all *in vitro* transcription experiments. Using competition of highly purified NTPs against the same radiolabeled NTP substrate minimizes concerns about variable purity of four different [α^{32}P]-NTPs and also reduces the cost.

These results suggest that nsp9 modification *in vivo* will be controlled by the relative abundance of natural NTPs and stabilities of the NMP adducts. Furthermore, it is possible that synthetic nucleoside triphosphates, such as the ATP analog remdesivir triphosphate (RTP) or the GTP analog AT-9010 that binds to AS2 (37), could transfer the NMP moiety to nsp9, whereas other analogs may act solely as competitive inhibitors. To test this idea, we used several nucleotide analogs as competitors of nsp9 GMPylation. We found that GDP, GMP, ITP (inosine triphosphate), and GMPCPP efficiently competed with [α^{32}P]-GMP transfer to nsp9, whereas ppGpp was less effective (Figure 2C). Unexpectedly, we also observed that the NMPylation reaction was strongly inhibited when inorganic pyrophosphate PP_i was present along with the GTP substrate (Figure 2C, last lane).

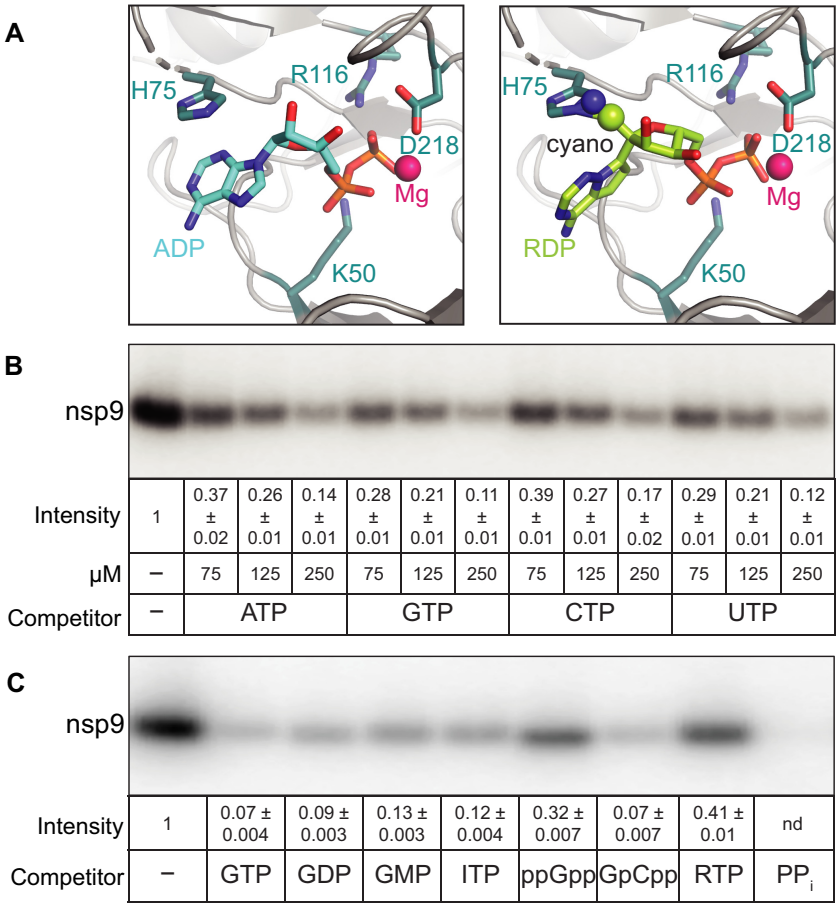


Figure 2. Effects of nucleotides on nsp9 modification. (**A**) Nucleotide binding to the NiRAN active site. The Mg^{2+} ion is shown as magenta sphere and key NiRAN residues - as sticks. Left, ADP (cyan carbon atoms) from PDB:6XEZ. Right, RDP (lime carbon atoms) modeled in place of ADP; the C1′ cyano group of RDP could clash with the side chain of nsp12 His75. (**B**) Unlabeled natural NTPs compete with [α^{32}P]-GTP for transfer to nsp9. (**C**) GTP analogs and PP_i inhibit NMPylation, but ppGpp and RTP do so less effectively; all nucleotides were present at 0.5 mM. In (B) and (C), NMPylation efficiency was compared to that observed in the absence of competitors, set at 1, and is shown as mean ± SD ($n = 3$).

Surprisingly, and in contrast to ATP (Figure 2B), we found that RTP was a relatively poor competitor (Figure 2C). Analysis of RNA synthesis by SARS-CoV-2 RdRp demonstrated that RTP binds to AS1 with much higher affinity than ATP and is a better substrate than ATP (43). Why does RTP fail to efficiently compete with GTP during NMPylation? In RTP, a cyano-group is attached to the 1′ position of the ATP ribose sugar; while the cyano-group does not interfere with RMP incorporation into the nascent RNA, it clashes with the Ser861 residue in nsp12 after RdRp adds three more nucleotides downstream of RMP, leading to a temporary stall during RNA chain extension (11,44). When remdesivir diphosphate is modeled in place of ADP into the structure of SARS-CoV-2 RdRp with the NiRAN-bound ADP•AlF_3 (Figure 2A, right), the cyano-group at the 1′ position clashes with His75, potentially explaining why RTP is a poor competitor of the NMPylation reaction.

In these experiments, an apparent reduction of [α^{32}P]-GMP transfer to nsp9 can be due to competitive inhibition of GTP binding (e.g. by GDP or GMP) or to nsp9 modification by an NTP analog (e.g. by ITP or RTP). To evaluate the second possibility, we used mass spectrometry. Our results show that, as reported by Slanina *et al.* (20), the N-terminus of nsp9 is modified by GMP (Supplementary Figure S5). We also observed NMPylation in the presence of ITP and RTP (Supplementary Figure S5). Although at present we cannot determine the efficiency of nsp9 modification by either nucleotide, our findings suggest that non-natural NTPs can be utilized as NiRAN substrates. In turn, this raises a possibility that antiviral nucleoside analogs have a potential to interfere with yet-to-be determined function of nsp9 in viral replication.

Pyrophosphate promotes the removal of GMP from nsp9

The nucleotidyl transfer reactions of AS1 and AS2 generate two products: PP_i, in each case, and either a one-nucleotide-extended RNA or NMP-nsp9, respectively. A reverse reaction, pyrophosphorolysis, is unfavorable at physiological concentrations of NTPs and PP_i, but is commonly used to evaluate the translocation register of multi-subunit DNA-dependent RNA polymerases (45), and has also been observed in RdRps (46,47). RNA polymerases behave as thermal ratchets that oscillate between the pre- and post-translocated registers on the template (48). This motion is

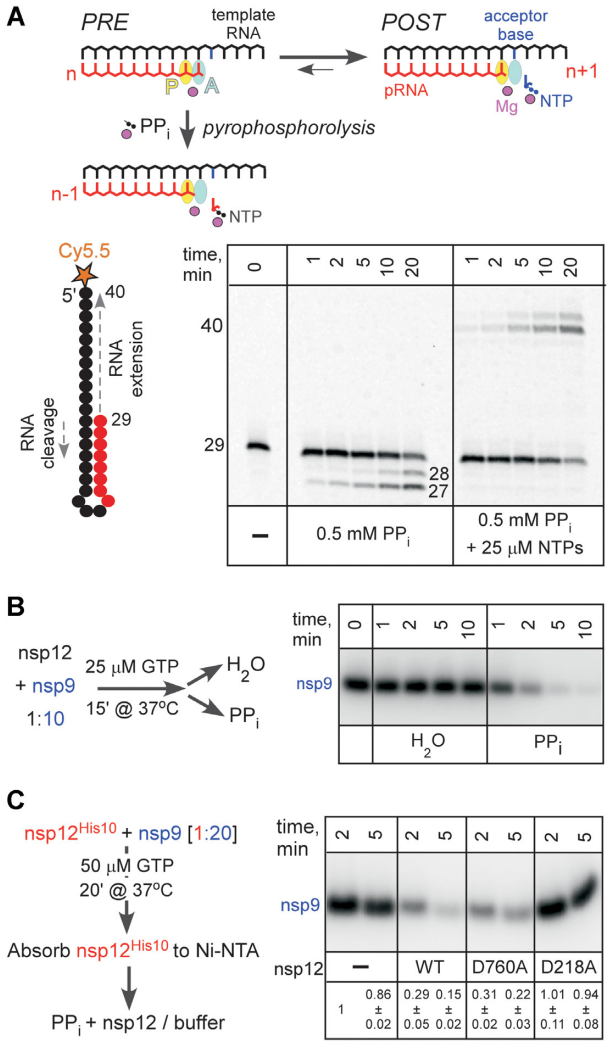


Figure 3. Removal of the GMP moiety in the presence of pyrophosphate. (**A**) Top: Pyrophosphorolysis of the nascent product RNA (pRNA; red). The active site (marked by the position of the catalytic Mg^{2+}; magenta sphere) consists of two sub-sites: the P-site (product; yellow) and the A-site (acceptor; cyan). Following nucleotide addition, the pRNA 3′ end is bound in the A-site in the pre-translocated state. Upon forward translocation, the 3′ end moves to the P-site and the incoming substrate NTP (blue) can bind to the A-site through base pairing with the acceptor base of the RNA template strand (black). The pre-translocated state is sensitive to pyrophosphorolysis, which generates the NTP product and a one-nt shortened pRNA. Bottom: SARS-CoV-2 RdRp transcription complex assembled on the 5′ Cy5.5-labeled hairpin, which comprises both the template and the product RNAs, is completely resistant to PP_i even in the absence of NTPs. (**B**) GMP removal from nsp9 in the presence of PP_i. (**C**) The hand-over assay in which pre-GMPylated nsp9 is incubated with WT or mutant nsp12 variants. Signal intensity was compared to that observed with nsp9 incubated with buffer (set at 1) and is shown as mean ± SD ($n = 3$).

rectified by binding of the incoming substrate NTP, which binds in the acceptor site (Figure 3A) and locks the post-translocated state, or of PP_i, which induces cleavage of the 3′-terminal nucleotide in the product site when the enzyme is in the pre-translocated register (Figure 3A). PP_i cleavage leads to shortening of the nascent RNA by one nucleotide and subsequent backward translocation, sometimes in several successive steps (49). The nascent RNA cleavage typically requires superphysiological concentrations of PP_i because the transcription elongation complex is biased toward the post-translocated state at most template positions, for bacterial RNA polymerases and SARS-CoV-2 RdRp alike (38,48). Consistently, we observed that scaffold-assembled SARS-CoV-2 complexes were relatively resistant to pyrophosphorolysis even in the absence of NTPs (Figure 3A), a result that is comparable to those obtained with hepatitis C virus (HCV) RdRp (46,47). Interestingly, we observed non-canonical PP_i-induced RNA cleavage by two nucleotides in a fraction of complexes, reminiscent of 'reverse pyrophosphorolysis' by noncognate NTP substrates in HCV RdRp that also generates a 2-nt cleavage product (46). Similar to the results obtained for the HCV enzyme (46), when PP_i was present in 200-fold molar excess over NTPs, polymerization reaction was favored and no cleavage was apparent (Figure 3A). Unlike some RNA polymerases, e.g. HCV RdRp (46) and *E. coli* RNA polymerase (34), which cleave the nascent RNA in the presence of noncognate NTPs, SARS-CoV-2 RdRp did not (Supplementary Figure S6).

However, we observed that PP_i efficiently inhibited nsp9 NMPylation (Figure 2C), consistent with the finding that PP_i binds to AS2 in SARS-CoV-2 RdRp/favipiravir complex (39). This inhibition could be due to PP_i competition with the substrate GTP, direct reversal of NMPylation reaction (pyrophosphorolysis), or hydrolysis assisted by the PP_i-bound Mg^{2+} ion. In cellular RNA polymerases, diverse small molecules and accessory proteins can deliver Mg^{2+} to the active site to stimulate the cleavage of the nascent RNA (48). To test if PP_i can de-NMPylate [^{32}P]-GMP-nsp9, we preincubated nsp9 with nsp12 prior to addition of PP_i (or water). In the presence of 0.5 mM PP_i, we observed rapid disappearance of the labeled nsp9 (Figure 3B), indicating that NMPylation is reversible; similar results were obtained with the nsp12•7•8_2 holoenzyme (Supplementary Figure S7).

In nsp12, two active sites mediate NMP transfer. A model in which NMPylated nsp9 serves as a primer for RNA synthesis implies that nsp9 binds to AS1 and positions the NMP for extension (20). Thus, both active sites could in principle mediate the PP_i-driven de-NMPylation. To evaluate the contribution of each active site, we carried out a 'hand-over' assay, in which histidine-tagged nsp12 used to NMPylate nsp9 was subsequently removed, and another, untagged nsp12 was added *post facto* (Figure 3C). We found that the WT and D760A nsp12s mediated de-NMPylation, whereas the D218A enzyme did not (Figure 3C), ruling out an essential contribution of AS1 to the reversal of nsp9 modification. Interestingly, while D760A is more efficient in NMPylating nsp9 (Figure 1C), it was slightly less efficient in the reverse direction. Thus, we cannot preclude the possibility of some involvement of AS1 in de-NMPylation, but the difference between the WT and D760A was barely significant ($P = 0.16$), necessitating a more detailed analysis with additional variants of AS1 and AS2 residues.

Only a few examples of de-AMPylation are known, and most utilize different catalytic domains, either in the same or in different proteins (19). An example in which the same Fic domain mediates AMPylation and de-AMPylation of BiP, a major ER chaperone required for protein homeosta-

sis in metazoans, has been recently reported (50). However, de-AMPylation releases AMP, not ATP, showing that FicD active site has both AMP transferase and phosphodiesterase activities (50). To elucidate the mechanism of the 'reverse' reaction catalyzed by the NiRAN domain, we analyzed the products of PP_i-induced de-GMPylation of Nsp9 using thin layer chromatography (Supplementary Figure S8). Our results suggest that de-GMPylation generates GTP and is therefore a true reversal of the forward reaction. However, the product pattern is complicated owing to an intrinsic nucleotide hydrolysis activity of nsp12 (Supplementary Figure S8), also observed by Yan *et al*. (21). Future experiments will be required to reveal the detailed mechanisms of all catalytic reactions catalyzed by RdRp.

Bisphosphonates inhibit nsp9 modification

Strong inhibition of NMPylation reaction by PP_i (Figure 3B) suggests that similar ligands that bind to AS2 (Figure 4A) may competitively inhibit nsp9 modification or trigger its reversal. To evaluate this possibility, we used chemically stable PP_i analogs bisphosphonates. We chose two FDA-approved compounds, Foscarnet (Fos) and Risedronate (Ris), as representative non-nitrogenous and nitrogenous bisphosphonates, respectively. Fos inhibits viral DNA polymerases, including HIV reverse transcriptase (51,52), and is used for treatment of infections caused by viruses in *Herpesviridae*. Ris is broadly used to treat diseases associated with bone loss, such as osteoporosis (53).

We show that Fos and Ris inhibit nsp9 NMPylation, although less efficiently than PP_i, (Figure 4B). While PP_i reduced NMPylation more than ten-fold when present at 50 μM, only 2-fold inhibition was achieved at 0.75 mM of either bisphosphonate (Figure 4B). These results suggest that while PP_i actively promotes de-NMPylation, bisphosphonates may act solely as competitive inhibitors of the forward reaction. Indeed, unlike PP_i, neither compound induced the removal of the GMP moiety from nsp9 (Figure 4C). We propose that bisphosphonates could be explored as inhibitors of NiRAN-mediated NMPylation; while neither of the two compounds tested was a potent inhibitor, many bisphosphonates are available or can be made to support structure-guided drug discovery.

DISCUSSION

Roles and targets of vital NiRAN NMPylation

The NiRAN domain is essential for replication of several human respiratory viruses, including the alphacoronavirus HCoV-229E, which causes the common cold, and betacoronaviruses SARS-CoV (18) and SARS-CoV-2 (20). Nucleotidylation activity of the NiRAN domain, which lacks any sequence homologs, was initially suggested by elegant bioinformatics analysis and confirmed by a proof-of-principle demonstration that EAV RdRp was capable of self-NMPylation (16–18). Structural similarities between SelO and NiRAN (16,17) further strengthened by identification of nsp9 as a NiRAN target among HCoV-229E proteins (20), argue that the NiRAN domain is a protein NMPylase.

In their pioneering study suggesting and confirming the existence of the NiRAN domain, Lehmann *et al*. postulated three potential roles of NiRAN-mediated NMPylation in the nidoviral replicative cycle (18). One possible role is that of an RNA ligase, although the identity of the substrates, and indeed the step itself, remains entirely hypothetical to date. Another is that of a guanylyltransferase (GTase) involved in 'capping' the 5′-end of transcribed RNA. Such capping is essential for viral replication and successful host infection, and all enzymes involved in the capping pathway, save the GTase, had already been identified years previously. The third possibility is that it serves a protein 'primer' of RNA synthesis, by covalently binding a nucleotide and, following its extension to a dinucleotide, delivering it to the 3′-end of the viral RNA template. Such priming is widely used across viral families (54).

In discussing each of these putative roles, the authors noted that the sum total of structural, functional, and phylogenetic evidence then available, including their own findings, could not be entirely reconciled with any single role, much less definitively preclude the two others. Several recent studies have been less hesitant, assigning to NiRAN exactly one of these roles.

First, Yan *et al*. posited that NiRAN performs a 'capping' role. In support of this assignment, they cited primarily structural arguments based on a cryo-EM snapshot of an RdRp-helicase complex in which the N-terminus of nsp9 was observed deep within the NiRAN active site, where it contacted a bound GDP molecule in a conformation stabilized by base-stacking with the His75 residue (21). They reasoned therefore that nsp9 must be either the target of NiRAN NMPylation or a competitive inhibitor of it, concluding the latter since their functional assays detected the formation of capped RNA but not the NMPylation of nsp9.

Second, Slanina *et al*. posited instead that NiRAN NMPylates nsp9, which then serves as a primer of RNA synthesis (20). Their functional studies provided direct evidence of NMPylation of nsp9 and NiRAN mediation thereof, with mutational and phylogenetic data supporting the additional conclusions that this NMPylation requires a free N terminus and allows little variation within the N-terminal tripeptide (Figure 1B). In particular, the indispensability both of Asn2 for nsp9 NMPylation *in vitro* and of NiRAN activity for viral replication provided a profound and elegant explanation why Asn2 is the only invariant residue across all nsp9 homologs (20).

Passing the baton: a speculative but integrative model

How can such findings be reconciled with one another, let alone with preceding or succeeding findings, including our own? Our results unequivocally demonstrate the importance of the native N-terminus of nsp9 for its NMPylation (Figure 1C), and thus we concur with Slanina *et al*. in arguing that the failure by Yan *et al*. to observe any such NMPylation is entirely due to their use of an artificially extended nsp9. The conclusion put forward by the latter – that NiRAN must therefore cap 5′ pRNA, and do so directly – is thus unfounded. However, if NiRAN is not the GTase 'missing link' in the capping pathway, no obvious candidate for this essential function remains.

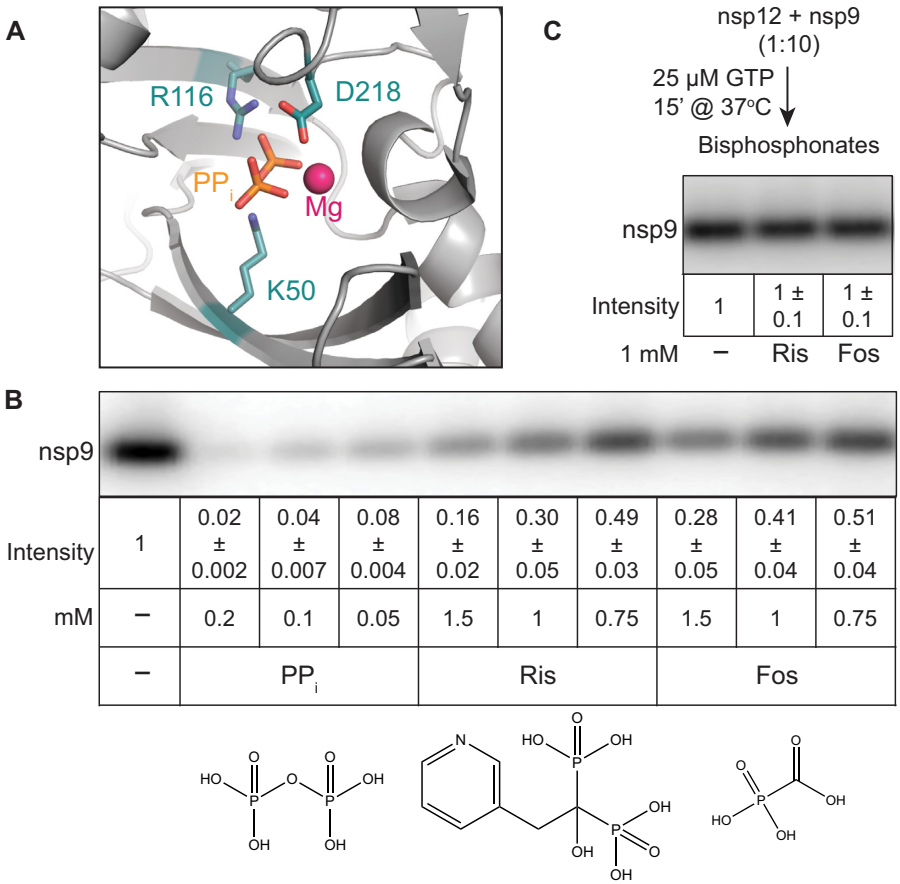


Figure 4. Effects of bisphosphonates on nsp9 modification. (**A**) Cryo-EM structure of the SARS-CoV-2 RdRp with PP_i bound to the NiRAN domain AS2. PDB: 7AAP (39). (**B**) Inhibition of NMPylation by PP_i and bisphosphonates; structures are shown below each compound. (**C**) Bisphosphonates do not promote the removal of NMP from modified nsp9. In (B) and (C), signal intensity was compared to that observed in the absence of added ligands (set at 1) and is shown as mean ± SD ($n = 3$).

As expected due to the lack of sequence-specific contacts between the nucleotide base and NiRAN residues (17,20), we found that all natural NTPs compete with GTP (Figure 2B), suggesting that nsp9 can be modified by diverse nucleotides (including remdesivir monophosphate, Supplementary Figure S5), and their respective cellular abundances will largely determine the identity of the adduct. However, it is possible that AS2 specificity may be 'tuned' in the presence of other RTC components.

We also show that, unlike the RNA chain synthesis, nsp9 modification is readily reversible in the presence of PPi (Figure 4B and Supplementary Figure S8) and that nsp9 interactions with AS2 are highly dynamic, i.e., NMPylated nsp9 released from nsp12 can be handed over to another enzyme for de-NMPylation (Figure 3C). Finally, we show that ligands that bind to AS2, including nucleoside mono- and diphosphates (Figure 2C) and bisphosphonates (Figure 4B), inhibit nsp9 NMPylation.

Taken together, these results strongly argue for NMPylation of nsp9 at NiRAN AS2. If so, to what end? nsp9 binds RNA, with no apparent sequence specificity (23,25), and nsp12 (21), but it is not clear how Asn1 modification would affect either interaction: residues thought to bind RNA are far away from Asn1 (23,25), and our results are inconsistent with any significant thermodynamic contribution of the NMPylation of nsp9 to its binding to nsp12 (Figure 1D). Rather, nsp9 appears to be ideally suited to deliver NMP to secondary acceptors: the NMP moiety is attached to the primary amine of N-terminal Asn1 (20) located at the end of a flexible N-terminal tail, and protein-N-NMP linkages are common in nucleotidyl transferases that catalyze ligation and capping reactions (55,56).

Therefore, we envision an essential role for NMPylated nsp9 in *both* priming and capping (Figure 5), perhaps as vital to the outcome as a baton passed between runners in a race. *First*, nsp9 is NMPylated by the NiRAN domain at AS2 and then dissociates from nsp12. *Second*, NMP-nsp9 binds to AS1 and serves as a primer for RNA synthesis; although nsp9 is not known to bind to specific RNA sequences (23,25), it is possible that, when bound to an RTC, NMP-nsp9 recognizes a specific sequence/structure in the viral RNA to direct precise initiation. It is likely that different RdRp complexes synthesize (+) and (-) RNA strands, complicating this analysis; priming of the (-) strand synthesis by NMP-nsp8 has been recently proposed (37). *Third*, as the nascent pRNA chain grows and is displaced from tRNA, pRNA-nsp9 rebinds to AS2 and a second nucleotidyl transfer reaction takes place to cap the

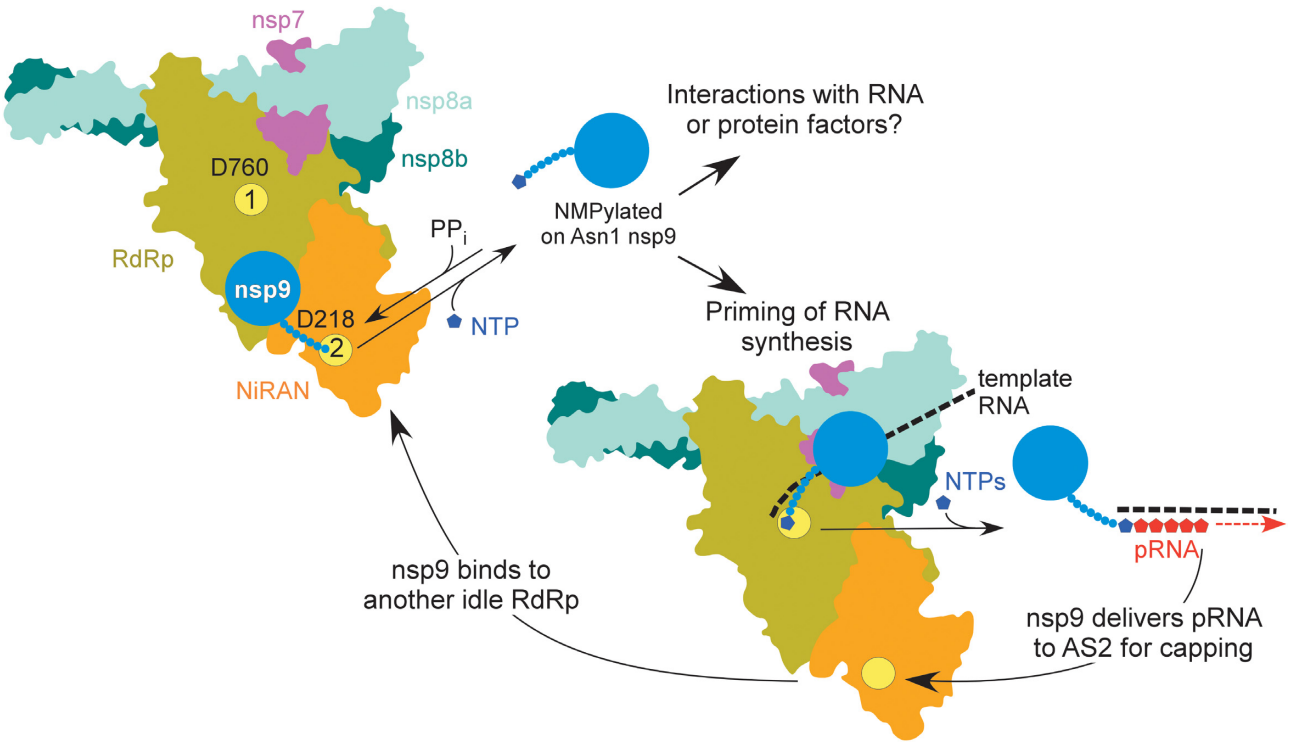


Figure 5. A model for nsp9 cycle; see text for details. AS1 and AS2 are shown as yellow circles, with the key catalytic residues indicated.

pRNA, releasing the unmodified nsp9 and resetting the cycle.

Avenues and implications for future research

We admit that this is a very speculative model and propose it to provoke investigation rather than to provide concrete answers. Mechanistic studies of SARS-CoV-2 RTCs are in their infancy, and future experiments will be needed to elucidate various aspects of its function and regulation. However, we argue that this model is a worthy starting point for several avenues of future research. Below we give several reasons for this claim and answer some anticipated objections.

First, such a capping mechanism is not unprecedented, for an analogous one has been described in rhabroviruses, such as vesicular stomatitis virus (VSV). VSV encodes a giant 2100-residue L protein, which contains RdRp, nucleotidyl transferase, and methyl transferase modules (54). Via a covalent (L-histidyl-$N^{\varepsilon 2}$)-pRNA intermediate, L transfers the pRNA moiety to GDP to yield GpppA-RNA (54). Can NiRAN use GDP as an acceptor? We show that GDP competes with GTP during nsp9 NMPylation (Figure 2C) and concentration of GDP in infected cells may be sufficient (57). While very little is known about the NiRAN catalytic mechanism, other AMPylases possess surprising catalytic diversity: in addition to NMP, Fic proteins can transfer phosphocholine and phosphate to their targets (58,59).

Second, nsp9 can be more than just a passive delivery vehicle for pRNA. Capping enzymes are composed of a nucleotidyl transfer domain fused to a distal OB-fold domain (55), suggesting that nsp9 OB-fold domain (23,25) may cooperate with the NiRAN domain during pRNA capping, remodeling AS2. For example, Gre and TFIIS transcription factors, which reactivate arrested RNA polymerases in all domains of life, deliver the second catalytic Mg^{2+} ion to the active site to switch it into an RNA cleavage mode (48). Such remodeling of AS2, possibly in partnership with other components of the RTC, might mediate sequence specificity of NMPylation of nsp9 in order to prime the initiation of positive- versus negative-polarity RNA.

Third, we also admit that neither we nor others have definitively precluded all other possible protein targets of NiRAN. Enzymes that catalyze post-translational protein modifications, including AMPylation, have broad specificities—SelO was found to transfer biotin-AMP to a number of targets, including common control substrates, and only some cellular targets of SelO are thought to be genuine (22). The high efficiency of NMP transfer to nsp9 (Supplementary Figure S3A), conservation of nsp9 N-terminus (Figure 1B), and the essentiality of N-terminal nsp9 residues for viral replication (19) all argue that nsp9 is a true protein target of NiRAN. In addition to nsp9, NiRAN could also modify some other viral or host proteins, complicating the extension of *in vitro* results to the viral replicative cycle or to the infection process as a whole. An interesting question, prompted by RdRp self-NMPylation observed in several studies (18,20,30), is whether the NiRAN domain is autoinhibited in the absence of a cognate substrate, a common feature among AMPylating enzymes (19). A conformational change upon substrate binding would trigger displacement of the autoinhibitory module, a target of self-NMPylation, making the active site accessible. Our observation that nsp12 GMPylates BSA only in the presence of

both Mn^{2+} and nsp9 (Supplementary Figure S4), together with the lack of BSA GMPylation reported by Conti *et al.* (30), is consistent with this idea.

Fourth, our model has at minimum the virtue of not merely reconciling various seemingly contradicting findings, but also suggesting how they might be integrated into a more holistic understanding of the role of NiRAN, in concert with RNA and protein factors, in the entire nidoviral replicative cycle. We recently showed that over-optimization of the SARS-CoV-2 RdRp coding sequence to replace rare codons in a heterologous expression platform can lead to an inactive enzyme (26). We also showed that both AS1 and AS2 not only share substrates and inhibitors but also 'cross-talk' via an allosteric pathway, and therefore that drug discovery and functional studies are myopic to focus exclusively on AS1 and unjustified in judging it to be the cause of all observed effects.

Similarly, a recent study found that whereas all existing cryo-EM structures of SARS-CoV-2 RdRp modeled nsp12 as chelating Zn centers, the physiological cofactors are in fact Fe–S clusters, which become replaced by Zn^{2+} ions in the aerobic conditions in which proteins are typically purified (60). Furthermore, such clusters were found to be essential, for their disassembly via oxidative degradation inhibited both RdRp activity and viral replication. This result, obtained using a well-characterized nitroxide, suggests a potentially rich vein of COVID-19 therapeutics that might have been completely overlooked had the suitability of the cryo-EM structural preparations not been properly questioned.

Both these previous results and those presented here clearly demonstrate the inherent dangers in using reductionist approaches to draw conclusions for more complex and holistic systems, such as viral replicative cycles and processes of host infection. Such approaches have advantages for quickly yielding insights into a narrow and well-defined question, and so it is wholly understandable why they are particularly attractive for research into systems like SARS-CoV-2 RdRp, where pressing concerns motivate researchers to obtain practical results as rapidly as possible. On the other hand, such frenetic research can easily outpace the self-correction normally occurring in science, suggesting that wherever possible, researchers should strive to holistically validate reductionist findings (e.g. verifying replication of mutant viruses in cell culture) and clearly communicate aspects of research methods that might be expected to restrict the applicability of their results.

DATA AVAILABILITY

All data that support the findings of this study are available from the corresponding author upon request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Georgi Belogurov, David Dulin, Markus Wahl, and three anonymous reviewers for helpful suggestions and to Liwen Zhang for mass spectrometry analysis.

FUNDING

Ohio State University Office of Research; National Institutes of Health [GM067153 to I.A.]; Mass spectrometry analysis was supported by NIH [P30 CA016058]; Fusion Orbitrap instrument was supported by NIH [S10 OD018056]. The open access publication charge for this paper has been waived by Oxford University Press - *NAR* Editorial Board members are entitled to one free paper per year in recognition of their work on behalf of the journal.
Conflict of interest statement. None declared.

REFERENCES

1. Gulyaeva,A.A. and Gorbalenya,A.E. (2021) A nidovirus perspective on SARS-CoV-2. *Biochem. Biophys. Res. Commun.*, **538**, 24–34.
2. Cui,J., Li,F. and Shi,Z.L. (2019) Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.*, **17**, 181–192.
3. Hu,B., Guo,H., Zhou,P. and Shi,Z.L. (2021) Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.*, **19**, 141–154.
4. Coronaviridae Study Group of the International Committee on Taxonomy of, V. (2020) The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.*, **5**, 536–544.
5. Munir,K., Ashraf,S., Munir,I., Khalid,H., Muneer,M.A., Mukhtar,N., Amin,S., Ashraf,S., Imran,M.A., Chaudhry,U. *et al.* (2020) Zoonotic and reverse zoonotic events of SARS-CoV-2 and their impact on global health. *Emerg. Microbes Infect.*, **9**, 2222–2235.
6. Snijder,E.J., Decroly,E. and Ziebuhr,J. (2016) The nonstructural proteins directing coronavirus RNA synthesis and processing. *Adv. Virus Res.*, **96**, 59–126.
7. Peersen,O.B. (2019) A comprehensive superposition of viral polymerase structures. *Viruses*, **11**, 745.
8. Jiang,Y., Yin,W. and Xu,H.E. (2021) RNA-dependent RNA polymerase: structure, mechanism, and drug discovery for COVID-19. *Biochem. Biophys. Res. Commun.*, **538**, 47–53.
9. Hillen,H.S. (2021) Structure and function of SARS-CoV-2 polymerase. *Curr. Opin. Virol.*, **48**, 82–90.
10. Gordon,C.J., Tchesnokov,E.P., Woolner,E., Perry,J.K., Feng,J.Y., Porter,D.P. and Gotte,M. (2020) Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. *J. Biol. Chem.*, **295**, 6785–6797.
11. Kokic,G., Hillen,H.S., Tegunov,D., Dienemann,C., Seitz,F., Schmitzova,J., Farnung,L., Siewert,A., Hobartner,C. and Cramer,P. (2021) Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. *Nat. Commun.*, **12**, 279.
12. Shannon,A., Selisko,B., Le,N.T., Huchting,J., Touret,F., Piorkowski,G., Fattorini,V., Ferron,F., Decroly,E., Meier,C. *et al.* (2020) Rapid incorporation of Favipiravir by the fast and permissive viral RNA polymerase complex results in SARS-CoV-2 lethal mutagenesis. *Nat. Commun.*, **11**, 4682.
13. Robson,F., Khan,K.S., Le,T.K., Paris,C., Demirbag,S., Barfuss,P., Rocchi,P. and Ng,W.L. (2020) Coronavirus RNA proofreading: molecular basis and therapeutic targeting. *Mol. Cell*, **80**, 1136–1138.
14. Gao,Y., Yan,L., Huang,Y., Liu,F., Zhao,Y., Cao,L., Wang,T., Sun,Q., Ming,Z., Zhang,L. *et al.* (2020) Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science*, **368**, 779–782.
15. Hillen,H.S., Kokic,G., Farnung,L., Dienemann,C., Tegunov,D. and Cramer,P. (2020) Structure of replicating SARS-CoV-2 polymerase. *Nature*, **584**, 154–156.
16. Kirchdoerfer,R.N. and Ward,A.B. (2019) Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat. Commun.*, **10**, 2342.
17. Chen,J., Malone,B., Llewellyn,E., Grasso,M., Shelton,P.M.M., Olinares,P.D.B., Maruthi,K., Eng,E.T., Vatandaslar,H., Chait,B.T. *et al.* (2020) Structural basis for helicase-polymerase coupling in the SARS-CoV-2 replication-transcription complex. *Cell*, **182**, 1560–1573.
18. Lehmann,K.C., Gulyaeva,A., Zevenhoven-Dobbe,J.C., Janssen,G.M., Ruben,M., Overkleeft,H.S., van Veelen,P.A., Samborskiy,D.V., Kravchenko,A.A., Leontovich,A.M. *et al.* (2015) 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.
19. Casey,A.K. and Orth,K. (2018) Enzymes Involved in AMPylation and deAMPylation. *Chem. Rev.*, **118**, 1199–1215.
20. Slanina,H., Madhugiri,R., Bylapudi,G., Schultheiss,K., Karl,N., Gulyaeva,A., Gorbalenya,A.E., Linne,U. and Ziebuhr,J. (2021) Coronavirus replication-transcription complex: Vital and selective NMPylation of a conserved site in nsp9 by the NiRAN-RdRp subunit. *Proc. Natl. Acad. Sci. U.S.A.*, **118**, e2022310118.
21. Yan,L., Ge,J., Zheng,L., Zhang,Y., Gao,Y., Wang,T., Huang,Y., Yang,Y., Gao,S., Li,M. *et al.* (2021) Cryo-EM structure of an extended SARS-CoV-2 replication and transcription complex reveals an intermediate state in cap synthesis. *Cell*, **184**, 184–193.
22. Sreelatha,A., Yee,S.S., Lopez,V.A., Park,B.C., Kinch,L.N., Pilch,S., Servage,K.A., Zhang,J., Jiou,J., Karasiewicz-Urbanska,M. *et al.* (2018) Protein AMPylation by an evolutionarily conserved pseudokinase. *Cell*, **175**, 809–821.
23. Egloff,M.P., Ferron,F., Campanacci,V., Longhi,S., Rancurel,C., Dutartre,H., Snijder,E.J., Gorbalenya,A.E., Cambillau,C. and Canard,B. (2004) The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 3792–3796.
24. Miknis,Z.J., Donaldson,E.F., Umland,T.C., Rimmer,R.A., Baric,R.S. and Schultz,L.W. (2009) Severe acute respiratory syndrome coronavirus nsp9 dimerization is essential for efficient viral growth. *J. Virol.*, **83**, 3007–3018.
25. Sutton,G., Fry,E., Carter,L., Sainsbury,S., Walter,T., Nettleship,J., Berrow,N., Owens,R., Gilbert,R., Davidson,A. *et al.* (2004) The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. *Structure*, **12**, 341–353.
26. Wang,B., Svetlov,V., Wolf,Y.I., Koonin,E.V., Nudler,E. and Artsimovitch,I. (2021) Allosteric activation of SARS-CoV-2 RdRp by remdesivir triphosphate and other phosphorylated nucleotides. *mBio*, **12**, e0142321.
27. UniProt, C. (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.*, **49**, D480–D489.
28. Katoh,K., Rozewicki,J. and Yamada,K.D. (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.*, **20**, 1160–1166.
29. Crooks,G.E., Hon,G., Chandonia,J.M. and Brenner,S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.
30. Conti,B.J., Leicht,A.S., Kirchdoerfer,R.N. and Sussman,M.R. (2021) Mass spectrometric based detection of protein nucleotidylation in the RNA polymerase of SARS-CoV-2. *Commun Chem*, **4**, 41.
31. Finkel,Y., Mizrahi,O., Nachshon,A., Weingarten-Gabbay,S., Morgenstern,D., Yahalom-Ronen,Y., Tamir,H., Achdout,H., Stein,D., Israeli,O. *et al.* (2020) The coding capacity of SARS-CoV-2. *Nature*, **589**, 125–130.
32. Gottesman,M.E., Chudaev,M. and Mustaev,A. (2020) Key features of magnesium that underpin its role as the major ion for electrophilic biocatalysis. *FEBS J.*, **287**, 5439–5463.
33. Vashishtha,A.K. and Konigsberg,W.H. (2016) Effect of different divalent cations on the kinetics and fidelity of RB69 DNA polymerase. *Biochemistry*, **55**, 2661–2670.
34. Sosunov,V., Sosunova,E., Mustaev,A., Bass,I., Nikiforov,V. and Goldfarb,A. (2003) Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J.*, **22**, 2234–2244.
35. Zhao,Z., Ma,Z., Wang,B., Guan,Y., Su,X.D. and Jiang,Z. (2020) Mn(2+) directly activates cGAS and structural analysis suggests Mn(2+) induces a noncanonical catalytic synthesis of 2′3′-cGAMP. *Cell Rep.*, **32**, 108053.
36. Cabart,P., Jin,H., Li,L. and Kaplan,C.D. (2014) Activation and reactivation of the RNA polymerase II trigger loop for intrinsic RNA cleavage and catalysis. *Transcription*, **5**, e28869.
37. Shannon,A., Fattorini,V., Sama,B., Selisko,B., Feracci,M., Falcou,C., Gauffre,P., El Kazzi,P., Decroly,E., Rabah,N. *et al.* (2021) Protein-primed RNA synthesis in SARS-CoVs and structural basis for inhibition by AT-527. bioRxiv doi: https://doi.org/10.1101/2021.03.23.436564, 23 March 2021, preprint: not peer reviewed.
38. Bera,S.C., Seifert,M., Kirchdoerfer,R.N., van Nies,P., Wubulikasimu,Y., Quack,S., Papini,F.S., Arnold,J.J., Canard,B., Cameron,C.E. *et al.* (2021) The nucleotide addition cycle of the SARS-CoV-2 polymerase. bioRxiv doi: https://doi.org/10.1101/2021.03.27.437309, 27 March 2021, preprint: not peer reviewed.
39. Naydenova,K., Muir,K.W., Wu,L.F., Zhang,Z., Coscia,F., Peet,M.J., Castro-Hartmann,P., Qian,P., Sader,K., Dent,K. *et al.* (2021) Structure of the SARS-CoV-2 RNA-dependent RNA polymerase in the presence of favipiravir-RTP. *Proc. Natl. Acad. Sci. U.S.A.*, **118**, e2021946118.
40. Yao,H., Song,Y., Chen,Y., Wu,N., Xu,J., Sun,C., Zhang,J., Weng,T., Zhang,Z., Wu,Z. *et al.* (2020) Molecular architecture of the SARS-CoV-2 Virus. *Cell*, **183**, 730–738.
41. Fauser,J., Gulen,B., Pogenberg,V., Pett,C., Pourjafar-Dehkordi,D., Krisp,C., Hopfner,D., Konig,G., Schluter,H., Feige,M.J. *et al.* (2021) Specificity of AMPylation of the human chaperone BiP is mediated by TPR motifs of FICD. *Nat. Commun.*, **12**, 2426.
42. Lauber,C., Goeman,J.J., Parquet Mdel,C., Nga,P.T., Snijder,E.J., Morita,K. and Gorbalenya,A.E. (2013) The footprint of genome architecture in the largest genome expansion in RNA viruses. *PLoS Pathog.*, **9**, e1003500.
43. Dangerfield,T.L., Huang,N.Z. and Johnson,K.A. (2020) Remdesivir is effective in combating COVID-19 because it is a better substrate than ATP for the viral RNA-Dependent RNA polymerase. *iScience*, **23**, 101849.
44. Wang,Q., Wu,J., Wang,H., Gao,Y., Liu,Q., Mu,A., Ji,W., Yan,L., Zhu,Y., Zhu,C. *et al.* (2020) Structural basis for RNA replication by the SARS-CoV-2 polymerase. *Cell*, **182**, 417–428.
45. Hein,P.P., Palangat,M. and Landick,R. (2011) RNA transcript 3′-proximal sequence affects translocation bias of RNA polymerase. *Biochemistry*, **50**, 7002–7014.
46. Jin,Z., Leveque,V., Ma,H., Johnson,K.A. and Klumpp,K. (2013) NTP-mediated nucleotide excision activity of hepatitis C virus RNA-dependent RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, E348–E357.
47. Deval,J., Powdrill,M.H., D'Abramo,C.M., Cellai,L. and Gotte,M. (2007) Pyrophosphorolytic excision of nonobligate chain terminators by hepatitis C virus NS5B polymerase. *Antimicrob. Agents Chemother.*, **51**, 2920–2928.
48. Belogurov,G.A. and Artsimovitch,I. (2019) The mechanisms of substrate selection, catalysis, and translocation by the elongating RNA polymerase. *J. Mol. Biol.*, **431**, 3975–4006.
49. Rozovskaya,T.A., Chenchik,A.A. and Beabealashvilli,R. (1982) Processive pyrophosphorolysis of RNA by *Escherichia coli* RNA polymerase. *FEBS Lett.*, **137**, 100–104.
50. Preissler,S., Rato,C., Perera,L., Saudek,V. and Ron,D. (2017) FICD acts bifunctionally to AMPylate and de-AMPylate the endoplasmic reticulum chaperone BiP. *Nat. Struct. Mol. Biol.*, **24**, 23–29.
51. Cai,H., Kapoor,A., He,R., Venkatadri,R., Forman,M., Posner,G.H. and Arav-Boger,R. (2014) In vitro combination of anti-cytomegalovirus compounds acting through different targets: role of the slope parameter and insights into mechanisms of Action. *Antimicrob. Agents Chemother.*, **58**, 986–994.
52. Das,K., Balzarini,J., Miller,M.T., Maguire,A.R., DeStefano,J.J. and Arnold,E. (2016) Conformational states of HIV-1 reverse transcriptase for nucleotide incorporation vs pyrophosphorolysis-binding of foscarnet. *ACS Chem. Biol.*, **11**, 2158–2164.
53. Drake,M.T., Clarke,B.L. and Khosla,S. (2008) Bisphosphonates: mechanism of action and role in clinical practice. *Mayo Clin. Proc.*, **83**, 1032–1045.
54. Ogino,T. and Green,T.J. (2019) RNA synthesis and capping by non-segmented negative strand RNA viral polymerases: lessons from a prototypic virus. *Front Microbiol*, **10**, 1490.
55. Shuman,S. and Lima,C.D. (2004) The polynucleotide ligase and RNA capping enzyme superfamily of covalent nucleotidyltransferases. *Curr. Opin. Struct. Biol.*, **14**, 757–764.
56. Brenner,C. (2002) Hint, Fhit, and GalT: function, structure, evolution, and mechanism of three branches of the histidine triad superfamily of nucleotide hydrolases and transferases. *Biochemistry*, **41**, 9003–9014.
57. Traut,T.W. (1994) Physiological concentrations of purines and pyrimidines. *Mol. Cell. Biochem.*, **140**, 1–22.
58. Castro-Roa,D., Garcia-Pino,A., De Gieter,S., van Nuland,N.A.J., Loris,R. and Zenkin,N. (2013) The Fic protein Doc uses an inverted

substrate to phosphorylate and inactivate EF-Tu. *Nat. Chem. Biol.*, **9**, 811–817.
59. Mukherjee,S., Liu,X., Arasaki,K., McDonough,J., Galan,J.E. and Roy,C.R. (2011) Modulation of Rab GTPase function by a protein phosphocholine transferase. *Nature*, **477**, 103–106.
60. Maio,N., Lafont,B.A.P., Sil,D., Li,Y., Bollinger,J.M., Krebs,C., Pierson,T.C., Linehan,W.M. and Rouault,T.A. (2021) Fe-S cofactors in the SARS-CoV-2 RNA-dependent RNA polymerase are potential antiviral targets. *Science*, **373**, 236–241.
61. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 486–501.