

# Characterization of SARS2 Nsp15 nuclease activity reveals it's mad about U

Meredith N. Frazier<sup>1</sup>, Lucas B. Dillard<sup>2</sup>, Juno M. Krahn<sup>2</sup>, Lalith Perera<sup>2</sup>, Jason G. Williams<sup>3</sup>, Isha M. Wilson<sup>1</sup>, Zachary D. Stewart<sup>1</sup>, Monica C. Pillon<sup>1</sup>, Leesa J. Deterding<sup>3</sup>, Mario J. Borgnia<sup>2</sup> and Robin E. Stanley<sup>1,\*</sup>

<sup>1</sup>Signal Transduction Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, 111 T. W. Alexander Drive, Research Triangle Park, NC 27709, USA,

<sup>2</sup>Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, 111 T. W. Alexander Drive, Research Triangle Park, NC 27709, USA and <sup>3</sup>Epigenetics and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, 111 T. W. Alexander Drive, Research Triangle Park, NC 27709, USA

Received June 02, 2021; Revised August 04, 2021; Editorial Decision August 05, 2021; Accepted August 09, 2021

## ABSTRACT

**Nsp15 is a uridine specific endoribonuclease that coronaviruses employ to cleave viral RNA and evade host immune defense systems. Previous structures of Nsp15 from across *Coronaviridae* revealed that Nsp15 assembles into a homo-hexamers and has a conserved active site similar to RNase A. Beyond a preference for cleaving RNA 3' of uridines, it is unknown if Nsp15 has any additional substrate preferences. Here, we used cryo-EM to capture structures of Nsp15 bound to RNA in pre- and post-cleavage states. The structures along with molecular dynamics and biochemical assays revealed critical residues involved in substrate specificity, nuclease activity, and oligomerization. Moreover, we determined how the sequence of the RNA substrate dictates cleavage and found that outside of polyU tracts, Nsp15 has a strong preference for purines 3' of the cleaved uridine. This work advances our understanding of how Nsp15 recognizes and processes viral RNA, and will aid in the development of new anti-viral therapeutics.**

## INTRODUCTION

The novel SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) emerged in late 2019 and became a worldwide pandemic that is still ongoing and has infected millions worldwide (1). Coronaviruses are members of the Nidovirus order, which encompasses large, positive-strand RNA viruses with genomes that range in size from 12 to 41 kb (2). The 30 kb SARS-CoV-2 genome encodes for 4 struc-

tural proteins that are part of the mature viral particle, 8 accessory proteins and 15 non-structural proteins (Nsps) (3). The Nsps are encoded in two open reading frames found in the first two-thirds of the viral genome. These proteins are translated by host ribosomes as two long polyproteins and are cleaved into functional proteins by the viral proteases (4). The Nsps play important roles in viral replication and pathogenicity and many of them are promising drug targets (3,4).

Nsp15 is a uridine specific endoribonuclease conserved across the *Coronaviridae* family (5). Enzymatic activity occurs in the C-terminal EndoU domain, which is more broadly conserved across nidoviruses, suggesting that this endoribonuclease activity is critically important for large, positive-strand RNA viruses (5,6). Work in animals and cell culture has shown that Nsp15 function is not necessary for viral replication, however Nsp15 nuclease activity is critically important for evasion of the host immune response to the virus, specifically by preventing the activation of dsRNA sensors (7–11). For example, in studies of porcine endemic diarrhea coronavirus (PEDV), Nsp15-deficient virus resulted in higher levels of type I and III interferon responses in cells, and piglets infected with the mutant virus had much higher survival rates than those infected with WT PEDV (8). A similar effect was also seen in studies of mouse hepatitis coronavirus (MHV); mice immunized with Nsp15 nuclease deficient virus were able to successfully clear WT virus, with commonly affected organs showing no pathology (7). Recent work also revealed a similar trend in the chicken infection bronchitis coronavirus (IBV), where animals infected with nuclease-deficient virus had reduced mortality and viral shedding (12). Therefore, Nsp15 is a promising therapeutic target for coronaviruses.

\*To whom correspondence should be addressed. Tel: +1 984 287 3568; Email: robin.stanley@nih.gov

Present address: Monica C. Pillon, Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030, USA.

One of the major outstanding questions about the function of Nsp15 is what is its RNA target for cleavage (5). Recent studies have begun to shed light on this important question (9,13). One study found a link between Nsp15 activity and the length of the polyuridine sequence at the 5' end of the template negative strand. When Nsp15-mutant MHV infected cells, there was a greater amount of polyuridine (PUN) RNA compared to cells infected with WT virus, suggesting Nsp15 cleaves the PUN RNA produced in the negative strand intermediate state (9). The PUN RNA was found to trigger the pathogen associated molecular pattern (PAMP) receptor MDA5, which mediates interferon response (9). Another study used cyclic phosphate RNA sequencing to identify Nsp15 cleavage products within MHV infected bone marrow-derived macrophages (13). This analysis revealed that Nsp15 cleaves numerous targets throughout the positive strand with a preference for cleaving between U<sup>^</sup>A and C<sup>^</sup>A sequences (13). More recent work with IBV demonstrated that Nsp15 nuclease activity prevents the accumulation of both dsRNA and cytoplasmic stress granules which have established anti-viral properties (14). Collectively these studies confirm that Nsp15 nuclease activity is critical to prevent the accumulation of viral dsRNA and activation of the immune response.

Numerous structures of Nsp15 have been determined from several *Coronaviridae* family members, however there are no structures of Nsp15 with more than a di-nucleotide bound, which has hindered our understanding of how Nsp15 recognizes its RNA targets. Crystal structures of Nsp15 revealed that Nsp15 assembles into a hexameric complex, formed from back-to-back trimers with the EndoU domains facing outward (15–19). The active site of Nsp15 shares considerable similarity to the well-studied endoribonuclease RNase A, and is composed of a catalytic triad including two histidines and lysine (18,20,21). These residues support a two-step reaction of transesterification and hydrolysis, however due to an altered position of one active site histidine, Nsp15 accumulates products from the transesterification reaction which contain a cyclic phosphate (20). Recent structures determined of Nsp15 bound to uridine nucleotides uncovered the molecular basis for uridine specificity, which is driven by a well-conserved serine residue within the uridine binding pocket (20,21). However, beyond the preference for uridines, it is unclear if Nsp15 has any additional specificity requirements. RNA sequencing suggests there is a preference for adenine 3' to the uridine, but the structural basis for this is unknown (13). In contrast, RNase A is known to have additional non-catalytic sites that affect substrate preference (22,23), prompting further characterization of how Nsp15 engages RNA.

Here, we used cryo-EM, molecular dynamics simulations, and in vitro RNA cleavage assays to probe the substrate specificity of SARS-CoV-2 Nsp15. We determined cryo-EM reconstructions of Nsp15 with RNA bound in the pre- and post-cleavage states. The structures revealed that, in contrast to RNase A, Nsp15 does not contain any additional well-ordered sites for RNA binding and recognition. This observation was further supported by molecular dynamics simulations with tri-nucleotide substrates. We probed RNA

specificity by determining how the nucleotide 5' and 3' of the uridine affects cleavage and found that Nsp15 has a preference for purines 3' of the cleaved pyrimidine. Finally, we looked at Nsp15's ability to cleave SARS-CoV-2 viral RNA substrates, such as the PUN and the transcriptional regulatory sequence (TRS). Collectively our work suggests that SARS CoV-2 Nsp15 is able to cleave a broad spectrum of RNA substrates and that this activity is driven by recognition of uridine within the active site.

## MATERIALS AND METHODS

### Protein expression and purification

Wild type (WT) and mutant Nsp15 constructs were created as described previously (20). Nsp15 was overexpressed in *Escherichia coli* C41 (DE3) competent cells in Terrific Broth with 100 mg/L ampicillin. At an optical density (600 nm) between 0.8 and 1.0, cultures were cooled at 4°C for 1 h prior to induction with 0.2 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were harvested after overnight expression at 16°C and stored at –80°C until use. Nsp15 purification was done as described previously (20). Briefly, cells were resuspended in Lysis Buffer (50 mM Tris pH 8.0, 500 mM NaCl, 5% glycerol, 5 mM β-ME, 5 mM imidazole) supplemented with cOmplete EDTA-free protease inhibitor tablets (Roche) and disrupted by sonication. The lysate was clarified at 26 915 × g for 50 min at 4°C and then incubated with TALON metal affinity resin (Clontech). His-Nsp15 was eluted from the resin with 250 mM imidazole, and buffer exchanged into Thrombin Cleavage Buffer (50 mM Tris pH 8.0, 150 mM NaCl, 5% glycerol, 2 mM β-ME, 2 mM CaCl<sub>2</sub>) for cleavage at room temperature for 3 h. The cleavage reaction was repassed over TALON resin and quenched with 1 mM phenylmethylsulfonyl fluoride (PMSF) prior to gel filtration using a Superdex-200 column equilibrated in SEC buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM MnCl<sub>2</sub>, 5 mM β-ME).

### Cryo-EM sample preparation

Purified Nsp15 was diluted in a low-salt buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 5 mM MnCl<sub>2</sub>, 5 mM β-ME) to 0.75 μM and incubated with excess RNA substrates (1 mM AU<sup>f</sup>A or AUA, see Supplementary Table S1) for 1 h at 4°C. UltrAuFoil R1.2/1.3 300 mesh gold grids (Quantifoil) were plasma cleaned (Pie Scientific) before use. The Nsp15/RNA mixture (3 μl) was deposited onto the grids, back-blotted for 3 seconds, and vitrified using an Automatic Plunge Freezer (Leica).

### Data collection and processing

Nsp15 images were collected using a Krios electron microscope at 300 keV with a Gatan K2 detector in super-resolution mode. Beam-induced motion and drift were corrected using MotionCor2 (24) and aligned dose-weighted images were used to calculate CTF parameters using CTFFIND4 (25). CryoSPARC v2 (26) was used in all subsequent image processing. Particles were selected by template-based particle picking, downsampled by a factor of 4, extracted with a box size of 64 and subjected to an initial

round of 2D classification. Full resolution particle projections from good classes were re-extracted using a box size of 256. *Ab initio* reconstruction was used to generate initial models; upon further inspection, there were no significant differences between the classes, which had roughly equal numbers of particles. Thus, the classes were grouped for further refinement in favor of better resolution. Three independent 3D refinement cycles were performed while applying C1, C3 and D3 symmetry respectively. Although previous apo- and UTP-bound datasets had D3 symmetry, the longer RNA bound in both datasets here resulted in particles that no longer had D3 symmetry, perhaps due to incomplete or mixed occupancy. Inspection of the C1 map did not reveal any asymmetric differences, although active site density was difficult to interpret for one half of the pre-cleavage state map. Therefore, C3 symmetry was used for model building and analysis for both datasets. Maps were re-scaled to optimize RMS fit to core domain residues of reference structure PDBID 6WLC (21).

### Model building

A SARS-CoV-2 Nsp15 crystal structure (PDBID 6WLC) was used as a starting model and fit into the cryo-EM maps using rigid body docking in Phenix (27). For the pre-cleavage state, which was captured with an AU<sup>f</sup>A trinucleotide, the density for the 5' A was weaker than the density for the U, so only the C5' group was modeled; no density was observed for the 3' A. For the post-cleavage state, the 5' A could be fit in the density along with the U. A combination of rigid body and real-space refinement in Phenix as well as iterative rounds of building in COOT (28) were used to improve the fit of the model. Molprobity (29) was used to evaluate the model (Table 1). Figures were prepared using Chimera (30) and Chimera X (31).

### FRET endoribonuclease assay

Nsp15 cleavage was monitored in real-time as described previously (19,20). Briefly, 6-mer substrates were labeled with 5'-fluorescein (FI) and 3'-TAMRA, where TAMRA quenches FI and cleavage is measured by increasing FI fluorescence (5'-FI-AAxxxA-TAMRA-3'; x nucleotides varied among substrates) (see Supplementary Table S1). The substrate (0.8 μM) was incubated with Nsp15 (2.5 nM) in RNA cleavage buffer (20 mM HEPES pH 7.5, 75 mM NaCl, 5 mM MnCl<sub>2</sub>, 5 mM DTT) at 25°C for 60 min. Fluorescence was measured every 2.5 min using a POLARstar Omega plate reader (BMG Labtech) set to excitation and emission wavelengths of 485 ± 12 and 520 nm, respectively. Three technical replicates were performed for each condition, and the assay was repeated with at least two independent protein preparations. Prism (Graphpad) was used to calculate significant differences using Dunnett's T3 multiple corrections test.

### Urea-PAGE endoribonuclease assay

Double fluorescently-labeled RNA substrates (5'-FI and 3'-Cy5, 500 nM) were incubated with Nsp15 (50 nM) in RNA cleavage buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 5

**Table 1.** Cryo-EM collection and processing statistics for the pre-cleavage and post-cleavage structures

Data collection and processing	Pre-cleavage state (EMBD-24137, PDBID: 7N33)	Post-cleavage state (EMBD-24101 PDBID: 7N06)
Microscope	Titan Krios	Titan Krios
Detector	Gatan K2	Gatan K2
Nominal magnification	165000x	165000x
Voltage (kV)	300	300
Electron exposure (e <sup>-</sup> /Å <sup>2</sup> )	54	54
Defocus range (μm)	-0.8 to -1.8	-0.8 to -1.8
Pixel size (Å)	0.4125	0.4125
Symmetry imposed	C3	C3
Number of micrographs	2563	3454
Initial particle images	944 302	1 592 506
Final particle images	383 275	1 058 228
Map resolution (Å)/FSC threshold	2.5/0.143	2.2/0.143
<b>Refinement</b>		
Resolution (Å)	2.5	2.2
B-factor used for map sharpening (Å <sup>2</sup> )	115.5	94.5
<b>Map to model CC</b>		
CC (mask)	0.87	0.88
CC (volume)	0.85	0.86
CC (peaks)	0.77	0.82
CC (box)	0.82	0.83
<b>Model composition</b>		
Non-hydrogen atoms	16 524	17 223
Protein residues	2076	2082
Nucleic acid	12	18
<b>Mean B factors (Å<sup>2</sup>)</b>		
Protein	53.45	33.82
Nucleic acid	74.64	49.73
<b>R.m.s. deviations</b>		
Bond lengths (Å)	0.005	0.007
Bond angles (°)	0.574	0.629
<b>Validation</b>		
Molprobity score	1.66	1.48
Clashscore	3.98	3.28
Poor rotamers (%)	2.93	2.10
<b>Ramachandranplot</b>		
Favored (%)	97.38	97.39
Allowed (%)	2.62	2.46
Disallowed (%)	0	0.14

mM MnCl<sub>2</sub>, 5 mM DTT, 1 u/μl RNasin ribonuclease inhibitor) at room temperature for 30 min, with samples collected at 0, 1, 5, 10 and 30 min. The reaction was quenched with 2× urea loading buffer (8 M urea, 20 mM Tris pH 8.0, 1 mM EDTA). Due to the expected size of cleavage products and the size of bromophenol blue, loading buffer without dye was used. To monitor the gel front, a control lane of protein only with bromophenol blue was run. To generate a ladder, alkaline hydrolysis of the RNA was carried out for 15 min at 90°C using 1 μM RNA in alkaline hydrolysis buffer (50 mM sodium carbonate pH 9.2, 1 mM EDTA) and quenched with 2× urea loading buffer. The cleavage reactions were separated using 15–20% TBE-urea PAGE gels and visualized with a Typhoon RGB imager (Amersham) using Cy2 (λ<sub>ex</sub> = 488 nm, λ<sub>em</sub> = 515–535 nm) and Cy5 (λ<sub>ex</sub> = 635 nm, λ<sub>em</sub> = 655–685 nm) channels.

### Mass spectrometry of RNA cleavage products

Mass spectrometry was performed as previously described (20). Briefly, the FRET RNA substrate of interest (0.8  $\mu$ M) was incubated  $\pm$  Nsp15 (2.5 nM) in RNA cleavage buffer for 30 min at RT. For mass spectrometry analysis, the reaction was chromatographically separated with a gradient of buffer A (400 mM hexafluoro-2-propanol, 3 mM triethylamine, pH 7.0) and buffer B (methanol). Parallel reaction monitoring (PRM) analyses were included in the MS analyses with included masses of  $m/z$  914.14; 923.14; 1463.42.

### Molecular dynamics simulations

Based on the RNA bound cryo-EM hexamer structure of Nsp15, the initial structure of Nsp15-AUA hexamer complex was prepared by manually introducing an adenine nucleotide at the B<sub>-2</sub> position. Except for H250, all histidine residues were selected to be N<sub>ε</sub> protonated. Since the ring nitrogen atoms on H250 were found to make two strong hydrogen bonds with the phosphate backbone and the carbonyl oxygen of S294, H250 was assigned the positively charged doubly-protonated form. After introducing all protons using the TLeap module of Amber.18 (32), the Nsp15-AUA hexamer system was solvated in 68 849 water molecules, while 203 sodium ions and 125 chloride ions provided the 100 mM salt concentration and the charge neutralization. A separate Nsp15-AUA monomer system, isolated from the hexamer, was also subjected to molecular dynamics. The monomer assembly was solvated with 24 545 water molecules. There were 57 sodium ions and 44 chloride ions also in the monomer system. The hexamer system consisted of 240 229 atoms while the monomer system had 79 295 atoms. The boundaries of the water boxes were at least 15 Å away from any protein or RNA atoms.

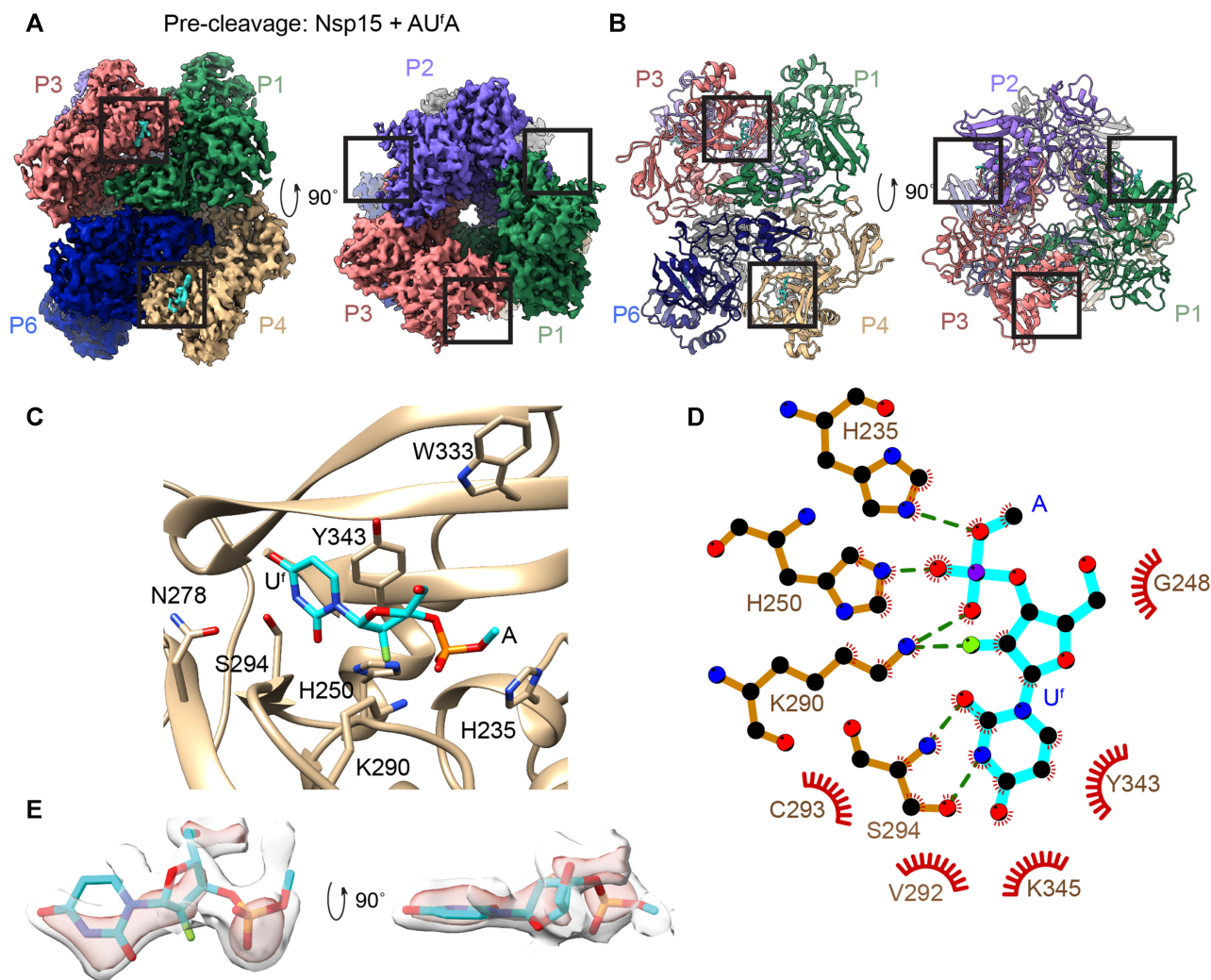
After proper equilibration of each system over 30 ns under various conditions, the CUDA implementation of the PMEMD module of Amber.18 with the amino acid represented by the FF14SB force field was used to simulate unconstrained dynamics for 500 ns for hexamer and monomer systems at 2 fs time step and 300 K under constant pressure. The Amber FF14 RNA force field was used for the ribonucleotide trimer. The particle mesh Ewald method was used in dealing with long range Coulomb and van der Waals interactions. For each system, two additional 500 ns simulations were performed. The starting structures of the additional runs were selected from the 30 and 40 ns conformations of the primary simulation with the randomized initial velocities to simulate alternate trajectories. The MMGBSA module of Amber.18 was implemented in free energy estimations with the selection of 0.15M salt concentration and the default parameters (IGB = 5) in the Amber module. Since the trinucleotide was not bound to the binding site residues during the entire half a microsecond production runs in most systems, the energy calculations were performed for each 50 ns segments (with 50 samples selected at each nanosecond) separately for each trajectory. When calculating the residues interaction energies, only the values from 50 ns segments with bound trinucleotides were selected.

## RESULTS

### Cryo-EM reconstructions of Nsp15 bound to RNA in pre- and post-cleavage states reveal substrate binding interactions

To gain insight into RNA cleavage by Nsp15, we determined cryo-EM structures of Nsp15 with RNA bound in pre- and post-cleavage states (Table 1, Figures 1 and 2). Given the similar active site arrangement and chemistry between Nsp15 and RNase A, we hypothesized that analogous to RNase A, there may be additional base specific binding pockets in Nsp15 (23,33). RNase A has multiple phosphate and base binding pockets that mediate the position of the base 5' (herein referred to as B<sub>-1</sub>) and 3' (herein referred to as B<sub>+1</sub>) to the scissile phosphate. Similar to Nsp15, the base binding pocket at the B<sub>-1</sub> position in RNase A confers specificity to pyrimidines, through a well conserved threonine (S294 in SARS-CoV-2 Nsp15). RNase A prefers purines following the pyrimidine and this preference is mediated by interactions in the B<sub>+1</sub> binding pocket.

To determine if Nsp15 has a B<sub>-1</sub> binding pocket, we trapped Nsp15 in a pre-cleavage state by incubating Nsp15 with a modified RNA containing a 2'-fluorine substituting for the 2'-OH on the uridine ribose that prevents catalysis (AU<sup>f</sup>A; Supplementary Table S1) prior to vitrification. 2'-fluorine modifications have been well-characterized and exploited for generating ribonuclease resistant oligos (34,35); the small size of fluorine in particular makes it an ideal modification unlikely to disturb the structure. Cryo-EM data were collected using a Titan Krios microscope (Supplementary Figure S1 and Table 1). The pre-cleavage state map was determined to 2.5 Å resolution with C3 symmetry applied. RNA density within the EndoU active site was observed in three of the six active sites, corresponding to one of the two trimers in the C1 map; poor resolution of the EndoU domains of the other trimer hindered conclusive identification of RNA within that trimer. We did not observe any asymmetric features within the trimers so C3 symmetry was subsequently applied due to the resolution enhancement from effectively tripling the particle number. Overall, the structure of the pre-cleavage state is very similar to previously determined SARS-CoV-2 Nsp15 structures bound to mono- or di-nucleotides, with RMSD values of <1 Å (Figure 1 and Supplementary Figure S2). We were able to visualize the complete, uncut scissile phosphate of the trinucleotide substrate (AU<sup>f</sup>A), although only the 2'-fluorouridine and the ribose C5' of the B<sub>+1</sub> A could be modeled in the density (Figure 1C and 1E). The fluorine was properly oriented within the catalytic triad (H235/H250/K290) even though no chemistry could take place; this also allowed for the U 3'-PO<sub>4</sub> to be positioned for hydrogen bonding with the catalytic triad. Consistent with UMP- and UTP-bound structures, the uracil base is poised to form hydrogen bonds with S294, such that Nsp15 can discriminate between U and other bases, while the uridine ribose group is oriented with Y343 to form van der Waals interactions (Figure 1D and Supplementary Figure S2) (20,21). Thus, this positioning represents a uridine poised for cleavage. The lack of well resolved density for either adenine base suggests that in contrast to RNase A, beyond the uridine recognition site Nsp15 does not have strong secondary base binding sites.

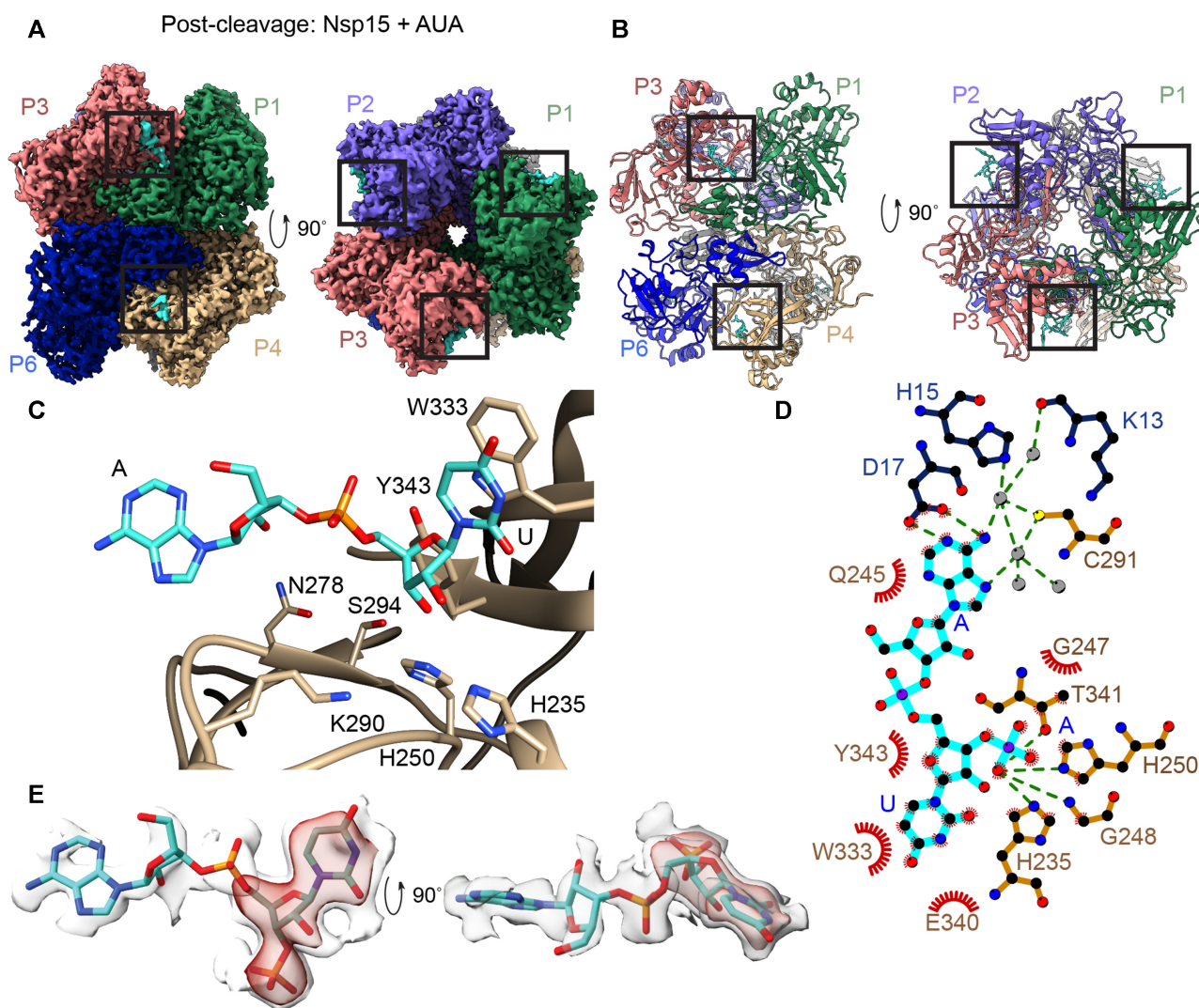


**Figure 1.** Pre-cleavage cryo-EM structure (Nsp15 + AU<sup>f</sup>A) shows the complete, uncut scissile bond. Side and top views of the cryo-EM map (A) and model (B). Protomers are colored and labeled. The active site is boxed in black and the RNA is colored teal. (C) Zoomed in active site with RNA bound. (D) 2D Ligplot (58) representation of interactions between the RNA and Nsp15. Green dashed lines represent hydrogen bonds (bond lengths < 3.0 Å). (E) The density for the AU<sup>f</sup>A ligand is shown at two different contour levels in grey and red. The density was best modeled with a complete uridine nucleotide; only a small part of the B<sub>+1</sub> adenine fit in the observed density.

To further probe RNA recognition by Nsp15 we determined the cryo-EM structure of Nsp15 in the post-cleavage state. We captured the post-cleavage state by incubating Nsp15 with excess unmodified RNA (AUA) prior to vitrification and cryo-EM data collection. Use of an unmodified RNA had the potential to lead to the capture of multiple states of Nsp15 (pre-cleavage, cyclic phosphate intermediate, and mono-phosphate product), however during classification and refinement we were only able to identify a single state containing the 5' final product (AU-3'P). Inspection of the C1 map revealed unambiguous density for RNA in all six active sites, however the RNA density was better resolved in one of the two trimers, so C3 symmetry was applied. The final cryo-EM reconstruction of the post-cleavage state went to 2.2 Å resolution and the map contains well resolved side chain density throughout the entire molecule. Analogous to the pre-cleavage state the overall structure of the post-cleavage state is similar to previ-

ous Nsp15 structures, with RMSD values of <1 Å (Figure 2 and Supplementary Figure S3), suggesting that there are no large conformational changes following transesterification and hydrolysis.

The RNA density within the active site was modeled as AU-3'P RNA, which represents the final 5' product following transesterification and hydrolysis (Figure 2C and E). The high-resolution density for the uracil base is unambiguous (Figure 2E). Notably, the position of the uracil base and ribose differs from the pre-cleavage state and the majority of previously determined nucleotide or nucleotide analogue structures of Nsp15 (Figure 3 and Supplementary Movie 1). While the 3'-PO<sub>4</sub> remains in the same place as the pre-cleavage state, the uracil has moved to pi-stack with W333 instead of interacting with Y343 and S294, a >10 Å movement of the base (Figure 2D). The observed movement of the uracil in the post-cleavage state aligns with both a recently published crystal structure of Nsp15 bound to 3'

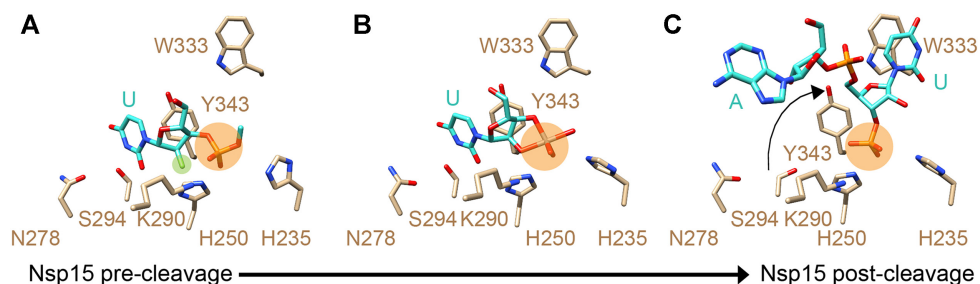


**Figure 2.** Post-cleavage cryo-EM structure (Nsp15 + AUA) reveals a change in U positioning. Side and top views of the cryo-EM map (A) and model (B). Protomers are colored and labeled. The active site is boxed in black and the RNA is colored teal. (C) Zoomed in active site with RNA bound. (D) 2D Ligplot (58) representation of interactions between the RNA and Nsp15. Green dashed lines represent hydrogen bonds (bond lengths of  $<3.25$  Å). (E) The density for the AUA ligand is shown at two different contour levels in grey and red. The high contour shows the exceptional density we saw for the uridine nucleotide, while the B<sub>-2</sub> adenine density is weaker.

UMP (PDB ID 6X4I, (21)) and a deposited crystal structure of Nsp15 bound to uridine 3',5'-diphosphate (PDB ID 7K1O), which represent a minimal product nucleotide (Supplementary Figure S2D). Those structures, along with our post-cleavage state, are the only Nsp15-substrate structures where the nucleotide substrate contains a free 3' PO<sub>4</sub>. The alignment seen between these structures provides further support that the repositioning of the uracil base following cleavage is a relevant post-cleavage state, though it may be one of several states sampled post-cleavage (Figure 3 and Supplementary Movie 1). While it is unclear why the uracil would leave the uracil binding pocket this structure suggests that W333 could be another critical residue for RNA binding within Nsp15. The role of W333 in RNA binding is further supported by a recent crystal structure of Nsp15 bound to a 5' GpU dinucleotide in which the guanine base pi-stacks with W333 (21). We also noticed the position of

S294 flips from the position it uses to form a hydrogen bond with the uracil base when it sits in that pocket. The result of this movement is that S294 moves out of hydrogen bonding distance from N278.

In contrast to the pre-cleavage state, we also observed weak density corresponding to the adenine base from the AU-3'P product. The density for the adenine is less well defined than the uridine but lies adjacent to the N-terminal domain (NTD) from a neighboring Nsp15 protomer (Figure 2C and D). The majority of these interactions are mediated through water molecules, although D17 does make base specific interactions with the adenine. Intriguingly, RNA crosslinking and mutational analysis with SARS-CoV-1, identified this same region of the NTD as being important for RNA binding (36). While the adenine density is poor the Nsp15 side chain density was well-resolved, and we observed unexpected density near H15 and C291



**Figure 3.** Nsp15 structures across the complete reaction mechanism. In all panels, active site residues of interest are shown as tan sticks, the RNA as turquoise sticks, and the scissile phosphate highlighted in orange. (A) The pre-cleavage structure with U<sup>f\*</sup> (highlighted in green) poised for cleavage (this work). (B) The cyclic intermediate state determined using the mimic uridine-2',3'-vanadate (PDB: 7K11 (21)). (C) The post-cleavage structure with the U pivoting to pi-stack with W333 (this work).

in close proximity to the active site (Supplementary Figure S6A and B). The density is continuous with the cysteine side chain suggesting a putative cysteine modification. Analysis of our recombinant Nsp15 by mass spectrometry suggests this could be a  $\beta$ ME adduct, although a full adduct does not fit into the observed density (data not shown). However, we cannot rule out the possibility that the  $\beta$ ME modification occurred during sample preparation for mass spectrometry as mass spectrometry also identified a  $\beta$ ME modification on C293, an interior facing residue with no extra density in our map. Due to this ambiguity, the additional density was modeled as water molecules but this observation suggests that C291 could be a reactive cysteine. Collectively the structures of Nsp15 in pre- and post-cleavage states revealed new insight into RNA recognition that prompted us to examine the function of residues surrounding the active site.

#### Molecular dynamics simulations and energy calculations support cryo-EM structural observations

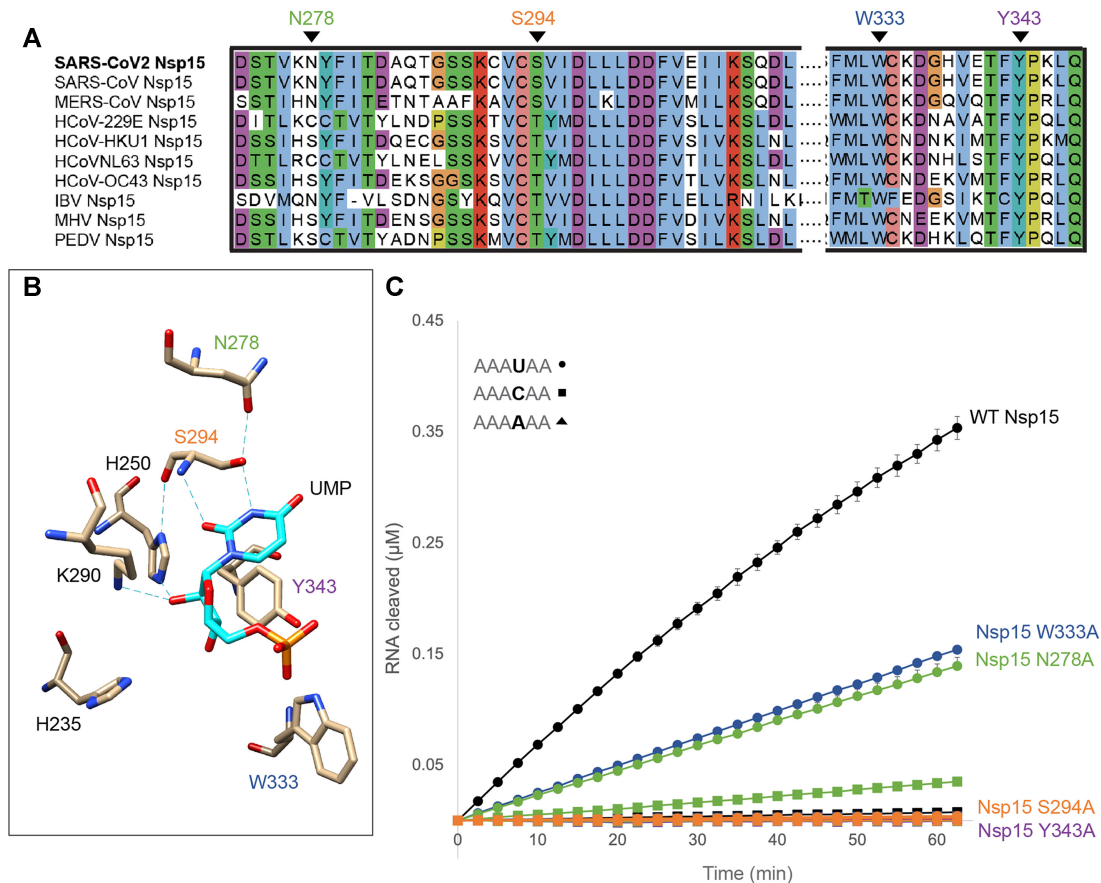
Our cryo-EM structure provided partial density for the uncleaved trinucleotide RNA substrate, with the nucleotides adjacent to the uridine seemingly highly dynamic. Therefore, we turned to molecular dynamics simulations to further characterize the behavior of the nucleotides near the active site. Root mean square deviations (RMSDs) were used to establish the stability of the simulated systems (Supplementary Figure S4A) in which the isolated monomer systems displayed elevated dynamics (as assessed by RMSDs) compared to the protomers assembled into the hexamer. The same trends were seen in both B-factors and the dynamic cross correlations (Supplementary Figure S4C). This is consistent with our previous molecular dynamics simulations revealing that the hexamer is important for protein stability (20). To assess RNA trinucleotide binding to Nsp15, we monitored multiple distances during the dynamic simulations, including three distances from the catalytic triad (K290, H235, H250) and two distances from the uridine discriminator (S294): NZ(K290)-P(PO<sub>4</sub>); NE2(H235)-O5'(B<sub>+1</sub>); NE2(H250)-O(PO<sub>4</sub>); N3(S294)-O2(U); OG(S294)-N3(U). These distances (Supplementary Figure S4B) guided us in deciding whether or not the AUA trinucleotide was bound, leading to subsequent analyses of dynamics behavior and related interaction strengths at the residue level. The ther-

mal fluctuations were evaluated through B-factors calculated from the simulations (Supplementary Figure S5); much larger fluctuations were displayed for the bases of both neighboring nucleotides compared to the uridine.

We then used MM/GBSA (molecular mechanics/generalized Born surface area) analysis to evaluate the contributions of Nsp15 active site residues to binding free energies. Residues providing binding energies with magnitudes above a 0.5 kcal/mol threshold are listed in Supplementary Tables S2 and S3. K290 and H250 provide the biggest contribution for the binding of the RNA trinucleotide, further highlighting the importance of the B<sub>-1</sub> phosphate for driving nucleotide binding. Similar characteristics were observed in the structurally similar RNase A system (22). Aside from the phosphate in the active site, the residues with the strongest energetic contributions mainly do so through interactions with the base and ribose of the U. Residues interacting with the B<sub>+1</sub> base overall feature larger energy values than the B<sub>-2</sub> base (with the exception of W333), suggesting that while both bases are not as fixed as U, the B<sub>+1</sub> base has more favorable interactions with Nsp15 active site residues (Supplementary Tables S2 and S3).

#### Nsp15 active site mutants exhibit reduced or abrogated RNA cleavage

To determine the significance of SARS-CoV-2 Nsp15 active residues in mediating RNA interactions and supporting cleavage we made a series of single point mutations. We expressed and purified the following Nsp15 active site variants: N278A, S294A, W333A and Y343A, which are largely well-conserved across coronaviruses (Figure 4A, Supplementary Figures S6–S8). All four active site variants were purified as stable hexamers, indicating that they did not disrupt the oligomerization of Nsp15 (Supplementary Figure S6). We measured RNA cleavage with a FRET-based assay (19,20) using 6-mer RNA substrates with 5' fluorescein and 3' TAMRA labels. TAMRA quenches the fluorescein signal of the uncleaved RNA, therefore increasing fluorescence over time serves as a readout for cleavage. We used three different RNA substrates for this assay to address Nsp15's ability to cleave RNA following either a U or a C, with a 6-mer polyA RNA serving as the negative control. The S294A and Y343A active site variants lack ap-



**Figure 4.** Nsp15 active site mutants reveal roles of critical residues. (A) Sequence alignment for selected regions of the Nsp15 EndoU domain across human coronaviruses and several major animal coronaviruses. Mutated residues are labeled. Residues that were mutated are colored to match the corresponding reaction curves. (B) Active site model of a pre-cleavage state (mimicked with 5' UMP, PDB: 6WLC) depicting the catalytic triad and other important active site residues. (C) FRET time course data for Nsp15 active site mutants. Nsp15 WT and mutants (2.5 nM) were incubated with RNA (0.8  $\mu$ M) at room temperature and fluorescence was monitored every 2.5 min for an hour. The average of a representative technical triplicate is plotted with standard deviation error bars. At least two biological replicates were performed for each mutant. Three different substrates were tested for each mutant: AAAUAA (circles), AAAAAA (squares), and AAACAA (triangles). Each mutant is represented by a different color: WT Nsp15 (black), Nsp15 W333A (blue), Nsp15 N278A (green), Nsp15 S294A (orange) and Nsp15 Y343A (purple).

preciable nuclease activity (Figure 4C). This confirms the critical importance of Y343 in orienting the uridine ribose and S294 in hydrogen bonding with the uracil base. S294 and Y343 are conserved in SARS-CoV-1 and the results of our mutational analysis are in good agreement with earlier studies with SARS-CoV-1 Nsp15 (36). Mutating N278A, the residue that interacts with S294 to position it correctly, resulted in hexameric enzyme that was  $\sim$ 2-fold less active and did not discriminate between U and C as strongly as WT (Figure 4C). At the endpoint of this assay, WT Nsp15 cleaved 40 $\times$  more U/C, while Nsp15 N278A only cleaved 3 $\times$  more U/C. Thus, N278 is important for maintaining uridine preference in SARS-CoV-2 Nsp15. Similar to N278A, the W333A variant was  $\sim$ 2-fold less active than WT Nsp15. This suggests that while W333 can provide a platform for pi-stacking with RNA bases, this interaction is not absolutely critical for RNA cleavage.

We also determined the importance of several residues near the weak adenine (B<sub>-2</sub>) base interaction site observed in the post-cleavage state, which lies at an interface with the EndoU domain of one protomer and the

NTD of another protomer within the same trimer. Given that we noticed extra density extending from the C291 side chain, we tested the importance of this residue in mediating cleavage (Supplementary Figure S6). Mutating C291A did not significantly affect the oligomerization or activity of Nsp15, which is not surprising given that it is not well conserved (only SARS-CoV-1 also has a cysteine at that position; Supplementary Figure S7). However, mutating the nearby H15A (from a neighboring protomer) disrupted the oligomerization of Nsp15 and resulted in predominantly inactive, monomeric enzyme (Supplementary Figure S6D, E and F). Previous work with SARS-CoV-1 Nsp15 revealed that the first 25 residues of the NTD were important for oligomerization, as a truncation produced monomeric Nsp15 (36,37). Thus, mutation of H15 which lies in a hinge region prior to a helix that interacts with the neighboring protomer, may be disrupting the folding or packing of the NTD.

We mutated two additional residues near the B<sub>-2</sub> base from the NTD including K13A and D17S, (Figure 2D) and saw no significant change in oligomerization compared to



WT Nsp15 (Supplementary Figure S6D and E). However, activity of K13A was  $\sim 2$ -fold lower, while activity of D17S was  $\sim 2$ -fold higher suggesting the neighboring NTD does affect EndoU function (Supplementary Figure S6F). One possible explanation for these results is charge stabilization of the RNA, since overall the EndoU surface is very negative. The robust extra density near C291 and H15 includes a shorter distance between atoms reminiscent of a metal ion interaction. Manganese ions are important for Nsp15 activity, but have never been localized in the active site, suggesting they are not structural metals. One hypothesis is that  $Mn^{2+}$  ions provide charge shielding to increase the favorability of the RNA-Nsp15 interaction, and thus are more transient. Given that loss of the K13 positive charge decreased activity, and the loss of the D17 negative charge increased activity, our data may point to a role in charge stabilization for this region of the active site. Overall, the identification of these new active site residues from a neighboring protomer important for Nsp15 activity provide more evidence for allosteric regulation and suggest a new way to inhibit this enzyme by targeting the NTD/EndoU interface. All together our structural and functional analysis highlights the significance of several Nsp15 residues in mediating RNA cleavage, specificity, and oligomerization.

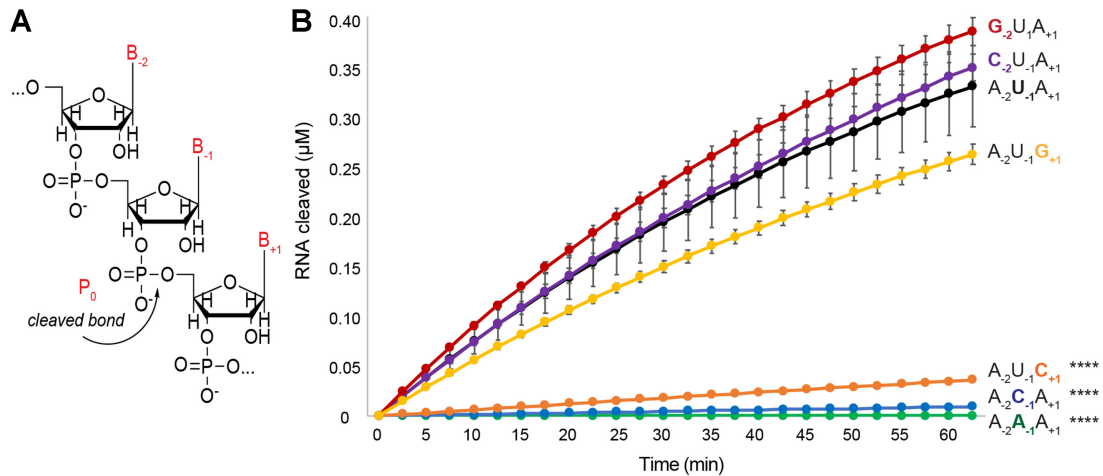
#### ***In vitro* FRET endoribonuclease assay reveals Nsp15 prefers a purine in the position following the uridine**

While the structures of Nsp15 in the pre- and post-cleavage states did not reveal strong secondary base binding sites, RNA cleavage assays revealed that Nsp15 has a preference for a purine 3' to the uridine in the cleavage site. Using the FRET endoribonuclease assay described above, we studied the importance of the bases in the  $-2$  and  $+1$  position ( $B_{-2}$ ,  $B_{+1}$ ) relative to the uridine being cleaved ( $B_{-1}$ ) using a series of oligomers with a single nucleotide change (Supplementary Table S1 and Figure 5A). Substituting an A or C for U in the  $B_{-1}$  position resulted in minimal cleavage, confirming Nsp15's specificity for cleavage following a U. In contrast, substitutions in the  $B_{-2}$  position did not significantly affect Nsp15 cleavage. However, substituting a C in the  $B_{+1}$  position resulted in significantly less cleavage compared to an A or G (Figure 5B). We further looked at the effect of substituting a U in either the  $B_{-2}$  or  $B_{+1}$  position. Because the presence of two U's introduces a second cleavage site, we measured time courses with either U or a modified un-cleavable U in the  $B_{-2}$  or  $B_{+1}$  position. Addition of a U in the  $B_{-2}$  position does not impact cleavage significantly but addition of a U in the  $B_{+1}$  position led to a reduction in cleavage, albeit not as significant of a reduction as with a C (Supplementary Figure S9A). We further analyzed the cleavage products of the unmodified double U RNAs by mass spectrometry which showed the accumulation of cleavage products following the uridine nucleotides at all three positions (Supplementary Figure S9B and C). Therefore, the differences observed in cleavage with the double U constructs suggests that the position of the U relative to the 5' or 3' end may also impact cleavage. Based on these endoribonuclease assay results, we define the consensus motif for *in vitro* Nsp15 cleavage on ssRNA as  $(N)(U)^{(R > U >> C)}$  (where N is any nucleotide and R is a purine). Together

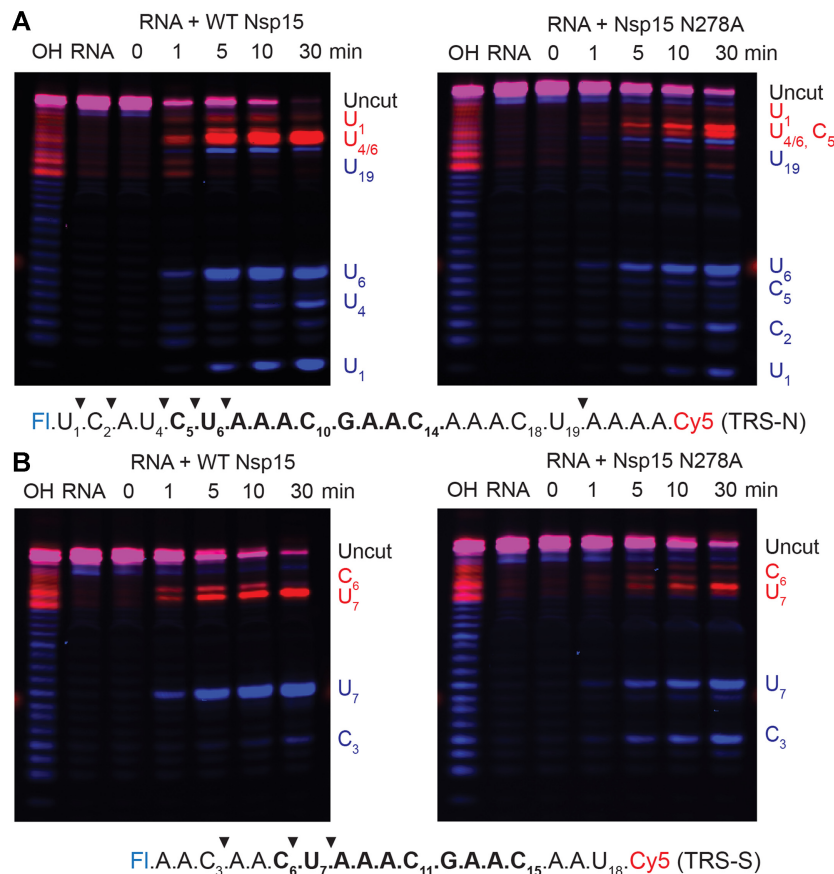
these results reveal that Nsp15 has a preference for purines over pyrimidines in the  $B_{+1}$  position, however Nsp15 can efficiently process substrates with tandem U's.

#### **Cleavage of biologically relevant substrates shows Nsp15 has a clear preference for U<sup>^</sup>A versus U<sup>^</sup>C**

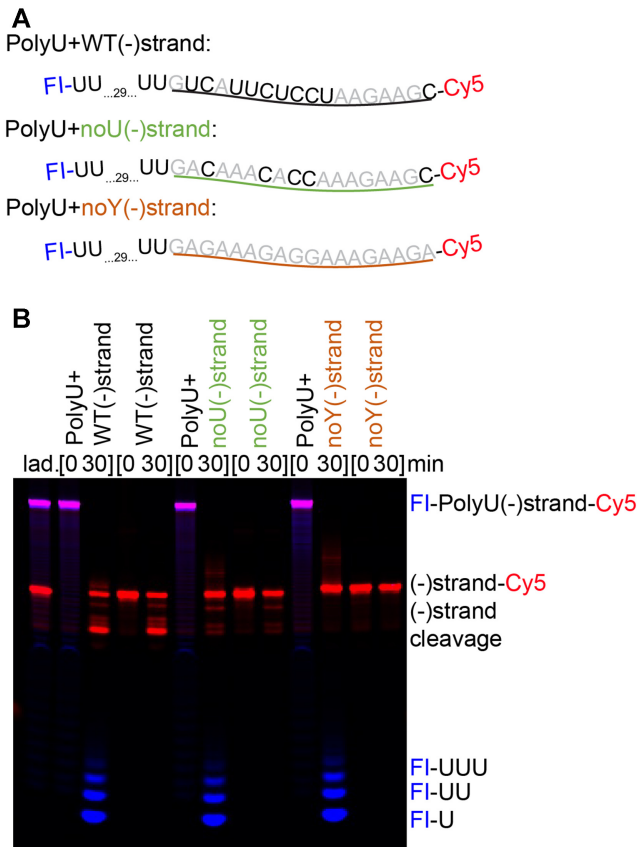
Our FRET-based cleavage assays revealed that Nsp15 demonstrates selectivity beyond uridine recognition in small 6-mer substrates leading us to ask whether this specificity is conserved in a) longer and b) more biologically relevant RNA substrates. Cyclic phosphate sequencing of MHV (a mouse model coronavirus) infected macrophages revealed MHV Nsp15 cleavage products throughout the positive strand of the MHV genome, including in conserved transcriptional regulatory sequences (TRS) (13). TRS sequences are short conserved motifs that lead to discontinuous viral RNA synthesis to produce sub-genomic RNAs (38). The consensus body TRS body (TRS-B) sequence from SARS-CoV-2 precedes each gene and promotes template switching by hybridizing to the TRS leader (TRS-L). The core TRS-B sequence includes one uridine and several cytidines, making it a putative substrate for Nsp15. We synthesized  $\sim 20$ -mer oligos containing the consensus TRS-B and flanking sequence for the nucleoprotein (N) as well as the spike (S) protein sub-genomic RNAs with labels on both the 5' and 3' ends (Supplementary Table S1). We analyzed these oligos with the RNAfold server to confirm that the short constructs used in our assays are unlikely to form secondary structures (39). We incubated WT Nsp15 with the TRS-N and TRS-S containing RNA substrates and then resolved the cleavage products on denaturing urea gels. We observed over time the accumulation of cleavage products following the uridines in both the TRS-N and TRS-S substrates (Figure 6). The time courses are consistent with the FRET-based cleavage data, although we had to use a higher ratio of protein to RNA to be able to visualize the products well. These higher protein concentrations are in-line with other Nsp15 activity assays (40). This data demonstrates that WT Nsp15 prefers purines in the  $B_{+1}$  position in longer substrates as we observed for the six nucleotide substrates. There are four U's within the TRS-N sequence with differing  $B_{+1}$  nucleotides:  $U_1^{\wedge}C$ ,  $U_4^{\wedge}C$ ,  $U_6^{\wedge}A$  and  $U_{19}^{\wedge}A$ . We observed faster accumulation of the  $U_6^{\wedge}A$  cleavage product over the  $U_1^{\wedge}C$  and  $U_4^{\wedge}C$  products. Cleavage at  $U_{19}^{\wedge}A$  was minimal, an effect that could be due to a bias for shorter products in this assay and/or issues detecting smaller Cy5-products (41,42). The TRS-S sequence contains a single  $U_7^{\wedge}A$  cleavage site. In addition to the  $U_7$  cleavage product observed with the TRS-S substrate, we also detect a small amount of cleavage following  $C_3^{\wedge}A$  and  $C_6^{\wedge}U$ . We repeated these cleavage assays with the Nsp15 N278A variant that demonstrated change in U/C specificity in our FRET cleavage assay (Figure 4). Consistent with our assay results with a six nucleotide substrate, the N278A Nsp15 variant cleaved more slowly and produced more C cleavage products than WT Nsp15 (Figure 6). We further tested the  $B_{+1}$  purine preference using another 20-mer oligomer (Supplementary Table S1) that was previously used to measure cleavage in SARS-CoV-2 Nsp15 (21) featuring multiple uridines with different  $B_{+1}$  bases. The time course again shows that the



**Figure 5.** Nsp15 RNA specificity identified by a FRET endoribonuclease assay reveals the importance of the B<sub>+1</sub> position. (A) Illustrative RNA diagram highlighting the scissile phosphate bond (P<sub>0</sub>), and defining the nomenclature of the bases of interest (B<sub>-2</sub>, B<sub>-1</sub>, B<sub>+1</sub>). (B) Reaction curves for comparing cleavage of different RNA sequences. WT Nsp15 (2.5 nM) was incubated with RNA (0.8 µM) at room temperature and fluorescence was monitored every 2.5 min for an hour. The average of a representative technical triplicate is plotted with standard deviation error bars. At least two biological replicates were performed for each mutant. For simplicity, only the three bases that were changed in this experiment are listed by the respective curves. U in the B<sub>-1</sub> position with A in the B<sub>-2</sub> and B<sub>+1</sub>, is shown in black. B<sub>-1</sub> was changed to an A (green) and C (blue). B<sub>-2</sub> was changed to G (red) and C (purple). B<sub>+1</sub> was changed to G (yellow) and C (orange). Dunnett's T3 multiple corrections test was performed using Prism for pairwise comparisons. The *P* < 0.0001 for the indicated (\*\*\*\*) curves.



**Figure 6.** Cleavage gels of biologically relevant TRS sequences show B<sub>+1</sub> sequence preference. The sequence is shown below the gels with the conserved consensus sequence in bold and the labels colored to match the overlays. Nsp15 (50 nM) was incubated with 5'-FI-RNA-Cy5-3' (500 nM) for 30 min at room temperature. Time course samples were taken at the indicated times and quenched with loading buffer. Alkaline hydrolysis of the RNA was performed to produce a ladder (lane OH) Representative gels are shown from three technical replicates. (A) Cleavage gels for WT Nsp15 (left) and Nsp15 N278A (right) with TRS-N. Images for the 5' (blue) and 3' (red) products were overlaid. (B) Cleavage gels for WT Nsp15 (left) and Nsp15 N278A (right) with TRS-S. Images for the 5' (blue) and 3' (red) products were overlaid.



**Figure 7.** Cleavage gels of negative strand sequences with both the PUN and internal Us shows Nsp15 acts across the sequence. Nsp15 (50 nM) was incubated with RNA (500 nM) for 30 min at room temperature. The polyU-NS substrates had 5'-FI and 3'-Cy5 labels, while the NS substrates only had 3'-Cy5 labels. (A) Cartoon depiction of the substrates tested. (B) Summary cleavage gel for the time course reactions showing only the 0- and 30-min samples for each substrate.

products accumulate preferentially, with the product resulting from a purine in the  $B_{+1}$  position ( $U_8^A G$ ) predominating (Supplementary Figure S10). Collectively the results from our gel-based cleavage assays support the results from our FRET cleavage experiments that Nsp15 prefers to cleave U's following purines.

Finally, we also looked at Nsp15's ability to degrade polyU sequences and found that Nsp15 efficiently degrades polyU containing RNAs *in vitro*. In addition to having numerous targets across the positive strand, previous work with MHV Nsp15 demonstrated activity in the PUN tract of the negative strand RNA *in vitro* and *in vivo* (9). We assessed the ability of SARS-CoV-2 Nsp15 to cleave polyU sequences in two ways. First, we performed a cleavage assay with a 5' fluorescein labeled U17 oligomer. Nsp15 acted efficiently on this substrate, leading to an accumulation of the single nucleotide product over time (Supplementary Figure S11). Similar to a previous study with MHV-CoV-1 Nsp15 (9), we also tested substrates with the SARS-CoV-2 wild-type negative strand sequence with and without the 5'-polyU sequence, as well as oligomers with the negative strand changed to have no uridines or no pyrimidines, with 5'-fluorescein and 3'-Cy5 labels (Figure 7, Sup-

plementary Figure S12 and Table S1). The dual labels allowed us to track formation of both 5'-polyU products and the 3'-negative strand products. Similar to the polyU alone cleavage, Nsp15 efficiently degraded the entire PUN RNA sequence at the 5' end. In addition to cleaving the PUN, Nsp15 cleaved uridines within the negative strand. When we removed the U's from the negative strand, a small amount of C cleavage products were observed, whereas removal of all pyrimidines resulted in an RNA that was not cleaved by Nsp15. These results are in excellent agreement with the earlier work showing that Nsp15 targets the PUN RNA sequence, however Nsp15 activity is not restricted to the PUN as it also cleaves additional uridines within the negative strand. Our structural, molecular dynamics, and biochemical data combined demonstrate that Nsp15 activity is driven by interactions with uridines in viral RNA. Thus, our current model features the six active sites cleaving U's efficiently, in both the 5' polyU sequence of the negative strand and throughout the rest of the genome. This occurs regardless of the surrounding sequence, although our data show Nsp15 does prefer a purine in the  $B_{+1}$  position outside of the negative strand polyU sequence.

## DISCUSSION

Despite its critical role in coronavirus pathogenesis, how Nsp15 recognizes its RNA targets was poorly understood. Here we characterized the cleavage motif of SARS-CoV-2 Nsp15 to provide insight into how Nsp15 processes viral RNA. We determined cryo-EM structures of Nsp15 in the pre- and post-cleavage states. While we hypothesized that Nsp15 would contain additional base binding sites analogous to RNase A, we did not observe any strong secondary base binding sites in our cryo-EM reconstructions, which suggests that beyond the uridine binding pocket the rest of the RNA is not fixed in place by strong interactions, and therefore averaged out during data processing. We did observe some density for the  $B_{-2}$  base preceding the uridine in the post-cleavage reconstruction towards the edge of the active site and adjacent to the NTD of a neighboring protomer. We identified several NTD residues from the adjacent protomer that interact with the  $B_{-2}$  adenine in our model and are important for oligomerization and nuclease activity. Therefore, in addition to being a more stable assembly (20,21), the hexamer may also be necessary for the NTD to support engagement of RNA in the active site. Thus, our structure-based point mutations suggest Nsp15 could be inhibited by disrupting the EndoU/NTD interface at the edge of the active site, which should destabilize the hexamer and lead to inactive monomeric enzyme.

While previous work established the role of S294 in conferring uridine specificity (20,21,36), here we identified N278 as a key residue in maintaining uridine preference in SARS-CoV-2 Nsp15. This residue is identical in SARS-CoV-1 and MERS Nsp15, but is only somewhat conserved (i.e. polar) across other coronaviruses that cause human disease (Supplementary Figure S7). Our structures show that N278 interacts with S294 to orient the hydrogen bond network to select uridines. The N278A variant showed reduced activity on a single U-containing substrate, but enhanced activity on a single C-containing substrate in com-

parison with WT Nsp15. Our SARS-CoV-2 S294A mutant appeared to have less activity than the equivalent SARS-CoV-1 and MERS mutants, which showed a ~2–5 fold reduction in activity (19,36). However, this may be due to differences in the concentrations and substrate used in our assays, and may further suggest that the weaker binding affinity of the S294A mutant can be overcome with an increase in enzyme. The conservation of S294 shows a similar pattern to N278; the equivalent residue is a serine in SARS and MERS Nsp15, but is a threonine in the other human coronaviruses. Similarly, MHV Nsp15 has a serine at the equivalent position to N278, and a threonine equivalent to S294 (Supplementary Figure S7). Cyclic phosphate sequencing of MHV-infected cells revealed that MHV Nsp15 cleaves both U<sup>A</sup> and C<sup>A</sup> sequences (13), which differs from our *in vitro* activity assays with SARS-CoV-2 Nsp15 that reveal a clear preference for cleaving 3' of U over C. While we cannot exclude the possibility that other viral or host factors may influence the sequence specificity of Nsp15 which could give rise to non-uridine cleavage products, our data suggests that the N278/S294 pair in SARS-CoV-2, SARS-CoV-1, and MERS may lead to a stronger preference for cleaving uridines.

We also probed substrate specificity via *in vitro* cleavage assays and found SARS-CoV-2 Nsp15 exhibited robust cleavage of PUN RNA, but appeared to prefer purines 3' of the cleaved uridine in sequences outside of the polyU tract. Thus, Nsp15 appears to have evolved to cleave uridines rather indiscriminately. Interestingly, the base composition of coronaviruses is heavily biased towards uridines (43), which suggests both a regulatory interplay and an additional need to target Nsp15 so it is not acting in an uncontrolled fashion. For example, the nascent viral RNA destined to be packaged must be protected from Nsp15 activity. Additionally, Nsp15 activity towards TRS sequences suggests it could play a role in regulating transcription, since TRS sequences play an important role in coronavirus transcription.

There are several putative mechanisms for regulation of Nsp15 nuclease activity. Modification and/or secondary structure of the viral RNA could regulate nuclease activity. For example, coronaviruses encode a 2'-O-methyltransferase (Nsp16); the modification produced by this enzyme would be expected to hinder Nsp15 endoribonuclease activity (38,44). SARS-CoV-1 Nsp15 does not cleave 2'-methylated RNA substrate efficiently and preferentially cleaves unpaired U's within a structured RNA substrate (36). Recent papers have mapped the secondary structure of the SARS-CoV-2 genome and found evidence for highly stable RNA folding throughout the genome (45–49). Beyond RNA modification and secondary structure, another potential mechanism of Nsp15 nuclease regulation is through compartmentalization. Extensive membrane rearrangement occurs within cells during coronavirus infection, leading to the formation of double membrane vesicles and other convoluted membrane structures (50–52). These membrane rearrangements are thought to help hide the transcription intermediate, double stranded RNA from host cell anti-viral sensors. This could lead to a local concentration of viral enzymes and viral RNA so that viral RNA editing is favored over host RNA editing. Recent

work demonstrated Nsp15 endoribonuclease activity also hindered the formation of anti-viral stress granules by regulating the formation of viral dsRNA (14). There is also evidence that Nsp15 may inhibit autophagy (53). Thus, the cellular localization of Nsp15 could also affect its targeting. Finally, other viral proteins may influence Nsp15 RNA targets and regulation in host cells, as Nsp15 is believed to localize within the replication-transcription complex of Nsps, including the RdRp complex (54,55). Stable or transient interactions with other RTC proteins could thus direct Nsp15 specificity and activity.

Overall, this work establishes that on its own SARS-CoV-2 Nsp15 is a largely non-specific endoribonuclease with recognition for a minimal consensus motif (N)(U)<sup>(R > U >> C)</sup> (where N is any base and R is a purine). However, other viral and host factors may influence Nsp15's cleavage preferences therefore additional studies are needed to determine SARS2 Nsp15's cleavage preference inside the host. Our data show that Nsp15 acts in a distributive fashion to catalyze cleavage following uridines. Nsp15 is a key player in blocking activation of host dsRNA sensors by preventing the accumulation of viral RNA and is therefore a promising therapeutic target (9,13,14). Prokaryotic and eukaryotic RNase A and EndoU family proteins have been characterized as distant homologues, with extensive bioinformatics revealing new members of the families throughout the kingdoms of life (56). These two families share a catalytic triad that represents an ancient RNA processing mechanism, which has been adapted to function in host defense and innate immunity in many instances (57). This work reveals that similar to RNase A, Nsp15 is a broad-spectrum endoribonuclease primarily guided to its cleavage targets by recognition of a single uridine. Our integrated structural and functional characterization of Nsp15 provides a platform for the development of innovative therapeutic strategies against coronaviruses.

## DATA AVAILABILITY

Atomic coordinates and structure factors for the reported cryo-EM structures have been deposited with the Protein Data Bank (PDB) and the Electron Microscopy Data Bank (EMDB) under accession numbers EMDB-24101, EMDB-24137, PDB ID: 7N06, and PDB ID: 7N33.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We would like to thank Dr. Rick Huang and Allison Zeher for help with cryoEM data collection. This work utilized the NCI/NICE Cryo-EM Facility. We would like to thank all members of the Stanley Lab and NIEHS cryo-EM core facility for helpful discussions on this project. Finally, we would like to thank Drs. Traci Hall and Oswaldo Lozoya for their critical reading of this manuscript.

## FUNDING

US National Institutes of Health Intramural Research Program; US National Institute of Environmental Health Sci-

ences (NIEHS) [ZIA ES103247 to R.E.S., Z01 ES043010 to L.P., 1ZI CES102488 to J.G.W., 1ZI CES103206 to L.J.D., ZIC ES103326 to M.J.B.]; NIH Intramural Targeted Anti-COVID-19 (ITAC) Program funded by the National Institute of Allergy and Infectious Diseases [1ZIAES103340]. Funding for open access charge: NIEHS.  
*Conflict of interest.* None declared.

## REFERENCES

- Morens,D.M. and Fauci,A.S. (2020) Emerging pandemic diseases: how we got to COVID-19. *Cell*, **182**, 1077–1092.
- Gulyaeva,A.A. and Gorbalenya,A.E. (2021) A nidovirus perspective on SARS-CoV-2. *Biochem. Biophys. Res. Commun.*, **538**, 24–34.
- Yoshimoto,F.K. (2020) The proteins of severe acute respiratory syndrome coronavirus-2 (SARS CoV-2 or n-COV19), the cause of COVID-19. *Protein J.*, **39**, 198–216.
- Hartenian,E., Nandakumar,D., Lari,A., Ly,M., Tucker,J.M. and Glaunsinger,B.A. (2020) The molecular virology of coronaviruses. *J. Biol. Chem.*, **295**, 12910–12934.
- Deng,X. and Baker,S.C. (2018) An ‘Old’ protein with a new story: Coronavirus endoribonuclease is important for evading host antiviral defenses. *Virology*, **517**, 157–163.
- Zheng,A., Shi,Y., Shen,Z., Wang,G., Shi,J., Xiong,Q., Fang,L., Xiao,S., Fu,Z.F. and Peng,G. (2018) Insight into the evolution of nidovirus endoribonuclease based on the finding that nsp15 from porcine Deltacoronavirus functions as a dimer. *J. Biol. Chem.*, **293**, 12054–12067.
- Deng,X., Hackbart,M., Mettelman,R.C., O’Brien,A., Mielech,A.M., Yi,G., Kao,C.C. and Baker,S.C. (2017) Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, E4251–E4260.
- Deng,X., van Geelen,A., Buckley,A.C., O’Brien,A., Pillatzki,A., Lager,K.M., Faaberg,K.S. and Baker,S.C. (2019) Coronavirus endoribonuclease activity in porcine epidemic diarrhea virus suppresses type I and type III interferon responses. *J. Virol.*, **93**, e02000-18.
- Hackbart,M., Deng,X. and Baker,S.C. (2020) Coronavirus endoribonuclease targets viral polyuridine sequences to evade activating host sensors. *Proc. Natl. Acad. Sci. U. S. A.*, **117**, 8094–8103.
- Wu,Y., Zhang,H., Shi,Z., Chen,J., Li,M., Shi,H., Shi,D., Guo,L. and Feng,L. (2020) Porcine epidemic diarrhea virus nsp15 antagonizes interferon signaling by RNA degradation of TBK1 and IRF3. *Viruses*, **12**, 599.
- Kindler,E., Gil-Cruz,C., Spanier,J., Li,Y., Wilhelm,J., Rabouw,H.H., Zust,R., Hwang,M., V’Kovski,P., Stalder,H. *et al.* (2017) Early endonuclease-mediated evasion of RNA sensing ensures efficient coronavirus replication. *PLoS Pathog.*, **13**, e1006195.
- Zhao,J., Sun,L., Zhao,Y., Feng,D., Cheng,J. and Zhang,G. (2020) Coronavirus endoribonuclease ensures efficient viral replication and prevents protein kinase R activation. *J. Virol.*, **95**, e02103-20.
- Anar,R., Li,Y., Kindler,E., Cooper,D.A., Ransom,M., Thiel,V., Weiss,S.R., Hesselberth,J.R. and Barton,D.J. (2020) Physiologic RNA targets and refined sequence specificity of coronavirus EndoU. *RNA*, **26**, 1976–1999.
- Gao,B., Gong,X., Fang,S., Weng,W., Wang,H., Chu,H., Sun,Y., Meng,C., Tan,L., Song,C. *et al.* (2021) Inhibition of anti-viral stress granule formation by coronavirus endoribonuclease nsp15 ensures efficient virus replication. *PLoS Pathog.*, **17**, e1008690.
- Ricagno,S., Egloff,M.P., Ulferts,R., Coutard,B., Nurizzo,D., Campanacci,V., Cambillau,C., Ziebuhr,J. and Canard,B. (2006) Crystal structure and mechanistic determinants of SARS coronavirus nonstructural protein 15 define an endoribonuclease family. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 11892–11897.
- Xu,X., Zhai,Y., Sun,F., Lou,Z., Su,D., Xu,Y., Zhang,R., Joachimiak,A., Zhang,X.C., Bartlam,M. *et al.* (2006) New antiviral target revealed by the hexameric structure of mouse hepatitis virus nonstructural protein nsp15. *J. Virol.*, **80**, 7909–7917.
- Cao,J., Wu,C.C. and Lin,T.L. (2008) Turkey coronavirus non-structure protein NSP15—an endoribonuclease. *Intervirology*, **51**, 342–351.
- Kim,Y., Jedrzejczak,R., Maltseva,N.I., Wilamowski,M., Endres,M., Godzik,A., Michalska,K. and Joachimiak,A. (2020) Crystal structure of Nsp15 endoribonuclease NendoU from SARS-CoV-2. *Protein Sci.*, **29**, 1596–1605.
- Zhang,L., Li,L., Yan,L., Ming,Z., Jia,Z., Lou,Z. and Rao,Z. (2018) Structural and biochemical characterization of endoribonuclease Nsp15 encoded by Middle East respiratory syndrome coronavirus. *J. Virol.*, **92**, e00893-18.
- Pillon,M.C., Frazier,M.N., Dillard,L.B., Williams,J.G., Kocaman,S., Krahn,J.M., Perera,L., Hayne,C.K., Gordon,J., Stewart,Z.D. *et al.* (2021) Cryo-EM structures of the SARS-CoV-2 endoribonuclease Nsp15 reveal insight into nuclease specificity and dynamics. *Nat. Commun.*, **12**, 636.
- Kim,Y., Wower,J., Maltseva,N., Chang,C., Jedrzejczak,R., Wilamowski,M., Kang,S., Nicolaescu,V., Randall,G., Michalska,K. *et al.* (2021) Tipiracil binds to uridine site and inhibits Nsp15 endoribonuclease NendoU from SARS-CoV-2. *Commun Biol*, **4**, 193.
- Prats-Ejarque,G., Lu,L., Salazar,V.A., Moussaoui,M. and Boix,E. (2019) Evolutionary trends in RNA base selectivity within the RNase a superfamily. *Front Pharmacol*, **10**, 1170.
- Fontecilla-Camps,J.C., de Llorens,R., le Du,M.H. and Cuchillo,C.M. (1994) Crystal structure of ribonuclease A.d(ApTpApApG) complex. Direct evidence for extended substrate recognition. *J. Biol. Chem.*, **269**, 21526–21531.
- Zheng,S.Q., Palovcak,E., Armache,J.P., Verba,K.A., Cheng,Y. and Agard,D.A. (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods*, **14**, 331–332.
- Rohou,A. and Grigorieff,N. (2015) CTFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.*, **192**, 216–221.
- Punjani,A., Rubinstein,J.L., Fleet,D.J. and Brubaker,M.A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods*, **14**, 290–296.
- Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 213–221.
- Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 486–501.
- Williams,C.J., Headd,J.J., Moriarty,N.W., Prisant,M.G., Videau,L.L., Deis,L.N., Verma,V., Keedy,D.A., Hintze,B.J., Chen,V.B. *et al.* (2018) MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.*, **27**, 293–315.
- Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF Chimera – a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
- Goddard,T.D., Huang,C.C., Meng,E.C., Pettersen,E.F., Couch,G.S., Morris,J.H. and Ferrin,T.E. (2018) UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.*, **27**, 14–25.
- Case,D.A., Brozell,S.R., Cerutti,D.S., Cheatham,T.E. III, Cruzeiro,V.W.D., Darden,T.A., Duke,R.E., Ghoreishi,D., Gilson,M.K., Gohlke,H. *et al.* (2018) In: *AMBER 2018*. University of California, San Francisco.
- Moussaoui,M., Guasch,A., Boix,E., Cuchillo,C. and Nogues,M. (1996) The role of non-catalytic binding subsites in the endonuclease activity of bovine pancreatic ribonuclease A. *J. Biol. Chem.*, **271**, 4687–4692.
- Guo,F., Li,Q. and Zhou,C. (2017) Synthesis and biological applications of fluoro-modified nucleic acids. *Org. Biomol. Chem.*, **15**, 9552–9565.
- Heidenreich,O., Pieken,W. and Eckstein,F. (1993) Chemically modified RNA: approaches and applications. *FASEB J.*, **7**, 90–96.
- Bhardwaj,K., Palaninathan,S., Alcantara,J.M., Yi,L.L., Guarino,L., Sacchettini,J.C. and Kao,C.C. (2008) Structural and functional analyses of the severe acute respiratory syndrome coronavirus endoribonuclease Nsp15. *J. Biol. Chem.*, **283**, 3655–3664.
- Joseph,J.S., Saikatendu,K.S., Subramanian,V., Neuman,B.W., Buchmeier,M.J., Stevens,R.C. and Kuhn,P. (2007) Crystal structure of a nonomeric form of severe acute respiratory syndrome

- coronavirus endonuclease nsp15 suggests a role for hexamerization as an allosteric switch. *J. Virol.*, **81**, 6700–6708.
38. V'Kovski,P., Kratzel,A., Steiner,S., Stalder,H. and Thiel,V. (2021) Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.*, **19**, 155–170.
  39. Gruber,A.R., Bernhart,S.H. and Lorenz,R. (2015) The ViennaRNA web services. *Methods Mol. Biol.*, **1269**, 307–326.
  40. Canal,B., Fujisawa,R., McClure,A.W., Deegan,T.D., Wu,M., Ulferts,R., Weissmann,F., Drury,L.S., Bertolin,A.P., Zeng,J. *et al.* (2021) Identifying SARS-CoV-2 antiviral compounds by screening for small molecule inhibitors of nsp15 endoribonuclease. *Biochem. J.*, **478**, 2465–2479.
  41. Killelea,T., Saint-Pierre,C., Ralec,C., Gasparutto,D. and Henneke,G. (2014) Anomalous electrophoretic migration of short oligodeoxynucleotides labelled with 5'-terminal Cy5 dyes. *Electrophoresis*, **35**, 1938–1946.
  42. Niewoehner,O. and Jinek,M. (2016) Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. *RNA*, **22**, 318–329.
  43. Berkhout,B. and van Hemert,F. (2015) On the biased nucleotide composition of the human coronavirus RNA genome. *Virus Res.*, **202**, 41–47.
  44. Chen,Y., Su,C., Ke,M., Jin,X., Xu,L., Zhang,Z., Wu,A., Sun,Y., Yang,Z., Tien,P. *et al.* (2011) Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. *PLoS Pathog.*, **7**, e1002294.
  45. Rangan,R., Watkins,A.M., Chacon,J., Kretsche,R., Kladwang,W., Zheludev,I.N., Townley,J., Rynge,M., Thain,G. and Das,R. (2021) De novo 3D models of SARS-CoV-2 RNA elements from consensus experimental secondary structures. *Nucleic Acids Res.*, **49**, 3092–3108.
  46. Huston,N.C., Wan,H., Strine,M.S., de Cesaris Araujo Tavares,R., Wilen,C.B. and Pyle,A.M. (2021) Comprehensive in vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. *Mol. Cell*, **81**, 584–598.
  47. Tavares,R.C.A., Mahadeshwar,G., Wan,H., Huston,N.C. and Pyle,A.M. (2020) The global and local distribution of RNA structure throughout the SARS-CoV-2 genome. *J. Virol.*, **95**, e02190-20.
  48. Rangan,R., Zheludev,I.N., Hagey,R.J., Pham,E.A., Wayment-Steele,H.K., Glenn,J.S. and Das,R. (2020) RNA genome conservation and secondary structure in SARS-CoV-2 and SARS-related viruses: a first look. *RNA*, **26**, 937–959.
  49. Manfredonia,I., Nithin,C., Ponce-Salvatierra,A., Ghosh,P., Wirecki,T.K., Marinus,T., Ogando,N.S., Snijder,E.J., van Hemert,M.J., Bujnicki,J.M. *et al.* (2020) Genome-wide mapping of SARS-CoV-2 RNA structures identifies therapeutically-relevant elements. *Nucleic Acids Res.*, **48**, 12436–12452.
  50. Wolff,G., Melia,C.E., Snijder,E.J. and Barcena,M. (2020) Double-membrane vesicles as platforms for viral replication. *Trends Microbiol.*, **28**, 1022–1033.
  51. Knoops,K., Kikkert,M., Worm,S.H., Zevenhoven-Dobbe,J.C., van der Meer,Y., Koster,A.J., Mommaas,A.M. and Snijder,E.J. (2008) SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biol.*, **6**, e226.
  52. Gosert,R., Kanjanahaluethai,A., Egger,D., Bienz,K. and Baker,S.C. (2002) RNA replication of mouse hepatitis virus takes place at double-membrane vesicles. *J. Virol.*, **76**, 3697–3708.
  53. Hayn,M., Hirschenberger,M., Koepke,L., Nchioua,R., Straub,J.H., Klute,S., Hunszinger,V., Zech,F., Prelli Bozzo,C., Aftab,W. *et al.* (2021) Systematic functional analysis of SARS-CoV-2 proteins uncovers viral innate immune antagonists and remaining vulnerabilities. *Cell Rep*, 109126.
  54. Athmer,J., Fehr,A.R., Grunewald,M., Smith,E.C., Denison,M.R. and Perlman,S. (2017) In situ tagged nsp15 reveals interactions with coronavirus replication/transcription complex-associated proteins. *mBio*, **8**, e02320-16.
  55. Shi,S.T., Schiller,J.J., Kanjanahaluethai,A., Baker,S.C., Oh,J.W. and Lai,M.M. (1999) Colocalization and membrane association of murine hepatitis virus gene 1 products and De novo-synthesized viral RNA in infected cells. *J. Virol.*, **73**, 5957–5969.
  56. Mushegian,A., Sorokina,I., Eroshkin,A. and Dlakic,M. (2020) An ancient evolutionary connection between ribonuclease A and EndoU families. *RNA*, **26**, 803–813.
  57. Li,J. and Boix,E. (2021) Host defence RNases as antiviral agents against enveloped single stranded RNA viruses. *Virulence*, **12**, 444–469.
  58. Laskowski,R.A. and Swindells,M.B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.*, **51**, 2778–2786.