Regulation of coronavirus nsp15 cleavage specificity by RNA structure

Indraneel A. Salukhe

A dissertation
submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington
2023

Reading Committee
Jennifer Hyde, Chair
Michael Emerman
Jason Smith

Program authorized to offer degree:
Microbiology

University of Washington

Abstract

Regulation of coronavirus nsp15 cleavage specificity by RNA structure

Indraneel A. Salukhe

Chair of the Supervisory Committee:
Jennifer Hyde
Department of Microbiology

SARS-CoV-2, the etiologic agent of the COVID-19 pandemic, has had a significant impact on global public health. However, SARS-CoV-2 is only one of multiple pathogenic human coronaviruses (CoVs) to have emerged since the turn of the century. CoVs encode for several nonstructural proteins (nsps) that are essential for viral replication and pathogenesis. Among them is nsp15, a uridine-specific viral endonuclease that is important in evading the host immune response and promoting viral replication. Despite the established endonuclease function of nsp15, other determinants of its cleavage specificity have only recently begun to be investigated. In this study we investigate the role of RNA secondary structure in SARS-CoV-2 nsp15 endonuclease activity.

We identified regions of differing predicted RNA secondary structure across the SARS-CoV-2 genome. Using a series of *in vitro* endonuclease assays, we observed that thermodynamically stable RNA structures were protected from nsp15 cleavage relative to RNAs lacking stable structure. We leveraged the s2m RNA from the SARS-CoV-1 3'UTR as a model RNA structure for our studies as it adopts a well-defined structure with several uridines, two of which are unpaired and thus highly probable targets for nsp15 cleavage. We found that SARS-CoV-2 nsp15 specifically cleaves s2m at the unpaired uridine within the pentaloop of the RNA. Further investigation revealed that the position of this uridine also impacted nsp15 cleavage efficiency suggesting that positioning within the pentaloop is necessary for optimal presentation of the scissile uridine and alignment within the

nsp15 catalytic pocket. Our findings indicate that RNA secondary structure is an important determinant of nsp15 cleavage and provides insight into the molecular mechanisms of RNA recognition by nsp15. Understanding the broader implications of nsp15 activity will provide further insight not only into CoV biology but also into drug development against nsp15.

Table of Contents

List of Tables and Figures

For Comet

Acknowledgements

I would not be here without the support of several mentors and friends. Thank you to Jenny for allowing me to explore my interests with my first foray into virology research and drug discovery. Nina and Ajai, thank you for your mentorship during my rotations. I would have been fortunate to have done my thesis work in your labs as well. I would like to thank my committee for their guidance over the course of 6 years, a pandemic, a failed project, and, ultimately, a successful one. I would like to specifically recognize Michael Emerman and Michael Lagunoff. Thank you for helping me navigate the scientific side of grad school. More importantly to me, your constant support outside of my research is why I have been able to graduate.

Wes Van Voorhis and Ryan Choi have been instrumental collaborators in the success of my thesis work. Thank you for taking a chance on me and entrusting me to carry forward a part of your lab's work. I would not have been able to complete my project without your assistance and advice.

I was grateful not only to learn under Mira and Kendall, but also to have the opportunity teach with them as colleagues. Thank you, Amy and Andrea for your boundless compassion and support. I am tremendously fortunate to have known you not just as advisors, but also as friends.

Mayumi, Kaitlyn, Phil, Brittany, Renae, Lyndsey, Noldy, Nikki and so many more of the microbiology grad students, I don't have the space to give you your due here, but I feel incredibly lucky to have you as friends. Thank you for letting me lean on you at the drop of a hat. Lyndsey, if carmelladansen died with us at a beach house in Oregon, she had a good run.

Sarah, you kept me afloat. The first hour of my day for the past year with you was critical for me to make it out. Thank you for being the person I could ask all my stupid questions to and for commiserating with me for six years. Can you imagine if we didn't end up liking each other?

Thank you, Mitch, Nikhil, Akshay, Rahul, Anay, Yizy, Elliot, Haley, and Grant for being with me and supporting me through everything. Throughout writing my grad school applications, training grants, manuscripts, and my thesis, you were usually asleep, Conor, but you were there, and that mattered to me.

Coach Pete, the values that you instilled in me as a 19-year old kid have stood with me since college. They've carried me through grad school, and I'll hold onto them for the rest of my life. Thank you for being the mentor that I needed.

Rochelle, Shivam, and Monica, when I started grad school, you welcomed me with your friendship, kindness, and generosity. Despite all the tribulations of the past six years, I would choose to do it all over again if I got to do it with you. I wouldn't have been able to make it without you. Thank you for being with me through it all.

To my parents who have made all of this possible: thank you for the rides to school and practice. Thank you for bearing with me, for all the sacrifices you've made, and your unwavering support. This accolade is yours more than it is mine.

I have been at the University of Washington for more than a decade now. In the past 11 years, I made countless friends and even more memories. This place has been good to me, and I am eternally grateful for all of it. Even as I say goodbye to UW, I will always be a Washington Husky.

Go Dawgs

Chapter I: Introduction

Coronaviruses

Coronaviridae are a family of RNA viruses that infect a wide range of species including mice, cows, and humans[1-20]. The family *Coronaviridae* is within the order *Nidovirales* and can be further categorized into two subfamilies: *Letovirinae* and *Orthocoronaviridae*. Most of the viruses in the family *Coronaviridae* are found in the *Orthocoronaviridae* which is subdivided into four genera (Figure 1.1)[21]. The *Alphacoronaviruses* and *Betacoronaviruses* infect mammals while the *Gammacoronaviruses* and *Deltacoronaviruses* primarily infect birds. Among the *Betacoroanviruses* are the most well-known coronaviruses (CoVs) including the long-studied model coronavirus, mouse hepatitis virus (MHV) and the highly pathogenic human coronaviruses, Severe acute respiratory syndrome coronavirus (SARS-CoV-1), Middle East respiratory syndrome-related coronavirus (MERS-CoV), and SARS-CoV-2[22]. CoVs have been studied for many years with the prototypical virus in the family, MHV having first been isolated in 1949[2,3]. Other coronaviruses of veterinary concern include infectious bronchitis virus (IBV) and Porcine Epidemic Diarrhea Virus (PEDV)[5,10,19,20,23,24].

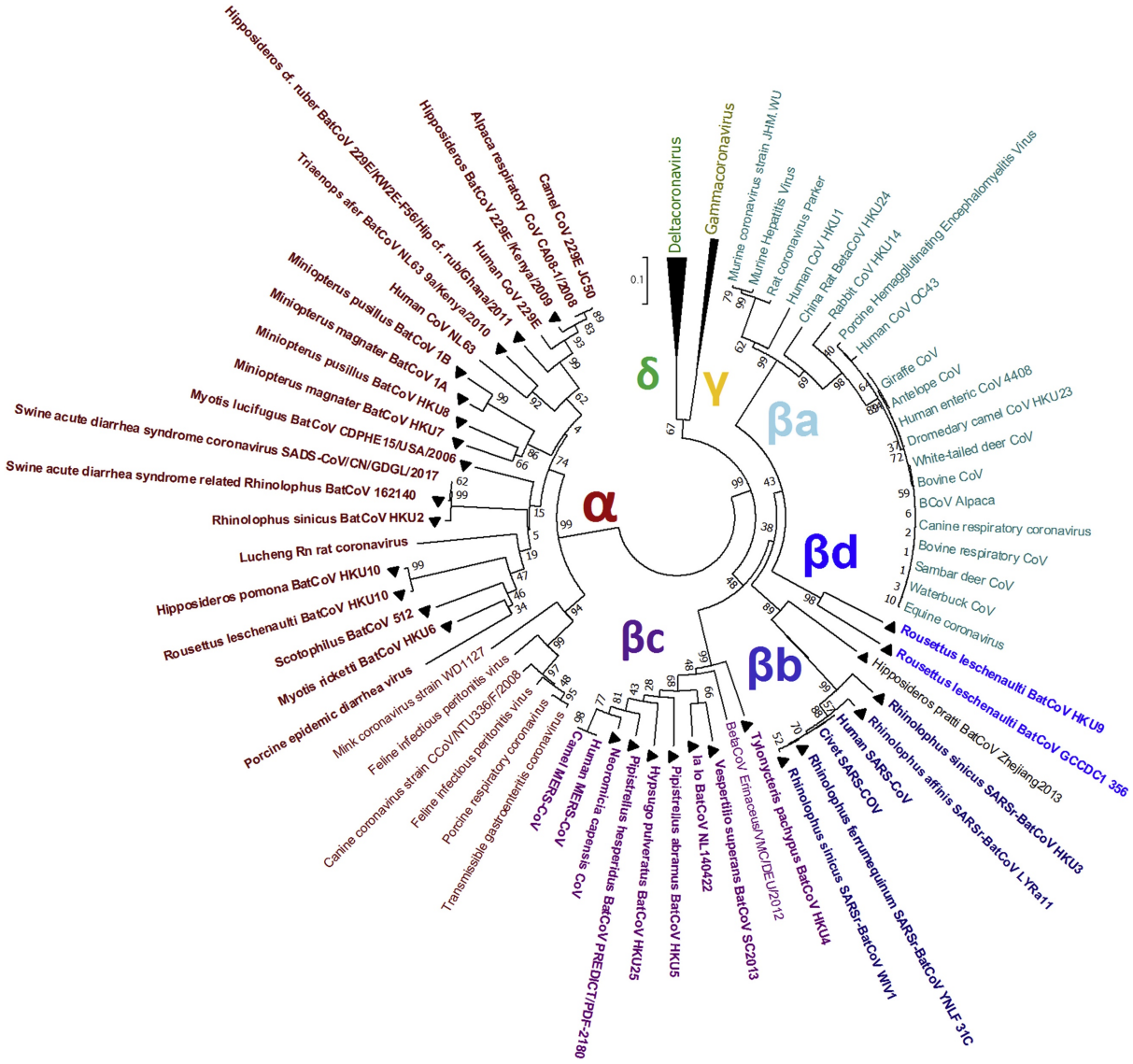


Figure 1.1 Phylogenetic tree depicting the evolutionary relationship of CoVs. This tree was generated based on the RNA-dependent-RNA-polymerase sequences of the queried viruses by Luk et al. *Infection, Genetics and Evolution* 2019. The four genera of *orthocoronaviridae* are depicted above with a focus on the alpha and betacoronaviruses. Notable pathogens in the betacoronaviruses include Severe acutre respiratory syndrome-CoV (SARS-CoV), Middle East respiratory syndrome-CoV (MERS-CoV), SARS-CoV-2.

Isolated in 1931[5,24,25], IBV was the first coronavirus to be discovered[22,26]. In the late 1920s and into the 1930s, reports of novel respiratory infections in baby chicks were noted. Soon after, IBV was found to be the causative agent of these. IBV impacted the poultry industry significantly as chicks would not feed very much leading to poor weight gain. Commercial broilers with IBV infections in their flocks would lose three to eight times as many chickens as uninfected flocks. Ultimately, broilers that contained IBV-infected chickens could lose 3% in income relative to broilers with healthy chickens[23].

Mortality of IBV in chickens varies with chicken breed and age[27]. The mortality rate of IBV in one week old pathogen-free chicks is 50% and drops further to 30% once the chicks reach two weeks old[28]. The mortality rate of chickens infected with IBV also increases with coinfections of *E. coli*[27,29]. To mitigate the impact of IBV on the poultry industry, both live and inactive virus vaccines have been developed. These vaccines have proven to be effective in preventing morbidity and mortality from IBV[30,31]. However, vaccines or antibodies from prior infections do not provide protection across strains of IBV[8,10]. As a result, the poultry industry is continuing to develop strategies to best combat IBV variants that are present in the local area. Although IBV does not infect humans, it was the first example of a CoV having a profound impact on human affairs and lifestyle.

MHV is a *Betacoronavirus* that has long been used as a model for human disease since it was first isolated from a paralyzed mouse in the 1940s[2]. This strain, termed JHM, not only caused paralysis, but also resulted in encephalitis and demyelination of the brain and spinal cord[3]. A second strain of MHV was isolated in 1961 and termed A59. In addition to encephalitis and demyelination, A59 caused hepatitis in mice[1,9]. However, A59 is generally considered to be a more attenuated strain of MHV than JHM as the latter is uniformly lethal in naïve mice whereas mortality in A59 infected mice is lower by comparison[32]. As a result, A59 has been the preferred virus to use as a model to study demyelination and multiple sclerosis in mice. As with IBV, immunocompromised animals infected with MHV developed more severe symptoms than healthy mice[33-35]. Additionally, the most severe lesions were found on 1-week old mice relative to their 3 and 12-week old counterparts indicating that age also factors into MHV pathogenesis[35]. Since the turn of the century, several groups have investigated the role of the immune response during

MHV infection. These studies have informed similar mechanisms of action in pathogenic human coronaviruses.

The first human CoVs (huCoVs), B814 and 229E were isolated in 1966[6,7]. Preliminary studies indicated that all demographics and ranges who were infected with these huCoVs presented with mild symptoms[36,37] while only in rare cases did infections cause pneumonia[38]. However, several studies following the initial characterization of B814 and 229E found that coronaviruses disproportionately cause more severe symptoms and lower respiratory tract infections in elderly and immunocompromised individuals[39-42]. As with other respiratory viruses, patients with asthma in particular encountered exacerbations of their asthmatic symptoms when infected by coronaviruses[43-45]. Ultimately, B814 was excluded from the CoV taxonomy due to difficult in culturing the virus[46]. Prior to 2002, human coronaviruses were restricted to causing the common cold and often did not cause severe symptoms[6,7]. Since then, further human coronaviruses have been isolated, but none of these viruses caused severe disease in humans until the emergence of SARS-CoV in 2002, MERS-CoV in 2012, and SARS-CoV-2 in 2019. To date, there are seven human CoVs including the three highly pathogenic CoVs mentioned above.

Coronavirus replication

CoVs have a positive sense single-stranded RNA genome (+ssRNA) the first two thirds of which is composed of two open reading frames (ORFs) that encode for the nonstructural genes (Figure 1.2). Like many RNA viruses, coronaviruses have a cytoplasmic replication cycle. The general replication strategy for CoVs is outlined in brief here and discussed in further detail below. Initially, CoVs bind to cellular receptors to promote viral uptake into the cytoplasm in an endosome. Following acidification of the endosome, the incoming viral genomic RNA is released into the cytoplasm. The genome is then directly translated from the two ORFs 1a and 1ab into polyproteins (pp1a and pp1ab). Viral proteases then cleave these polyproteins into single proteins that form the replication-transcription-complex (RTC)[47]. Genome replication takes place at double-membrane vesicles (DMVs) at the endoplasmic reticulum (ER)[48-51]. The viral genome is first replicated to negative sense, which is then used as a template for positive sense

genome synthesis. In addition to full-length negative sense RNA, the virus also generates subgenomic negative sense RNAs through transcription regulatory sequence (TRS) elements. These negative sense RNAs are used as templates to produce positive sense subgenomic RNAs which are translated into the structural and accessory genes. Following the production of the structural proteins, they are translocated to the endoplasmic reticulum and subsequently to the ER-to-Golgi intermediate compartment (ERGIC) through the secretory pathway. Here, the viral genomes are encapsidated and bud into the lumen of the ERGIC. Finally, the infectious virions are released from the cell through the secretory pathway and exocytosis.

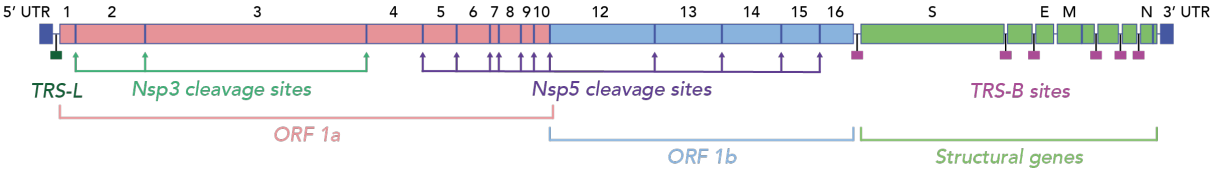


Figure 1.2 Organization of the coronavirus genome. Schematic of the coronavirus genome. The first two-thirds of the genome is composed of two open reading frames (ORF1a in light red, ORF1b in light blue) that encode the viral nsps. Identity of nsps are indicated above each gene. Cleavage sites of each viral protease are indicated below ORF1a and 1b. Structural genes (light green) are encoded in their own open reading frames at the 3' end of the genome. Notable structural genes are labeled above (S: spike, E: envelope, M: membrane, N: nucleocapsid). TRS-L site is labeled in dark green, and TRS-B sites are indicated in magenta.

To begin the infection process, the virus uses the spike (S) protein to bind to the cell surface receptors. SARS-CoV, and SARS-CoV-2 share a host cell receptor that permits infection in angiotensin-converting enzyme 2 (ACE2)[52-55]. As with SARS-CoV, the cell surface serine protease, TMPRSS2, was also identified as secondary surface protein that SARS-CoV-2 use for priming and entry[55]. Additionally, both SARS-CoVs can use the proteases cathepsin B and L (catB and catL, respectively) for priming[55,56]. Priming by these host proteins important for cleavage the viral S protein into a state that allows for fusion of the viral and cellular membranes[55].

Consequently, the expression and distribution of these surface proteins dictates tropism of these viruses. ACE2 is expressed in a wide variety of human tissues including the small intestine, testis, kidney, heart, thyroid, lung, and liver[57,58]. Interestingly, ACE2 expression in the respiratory tract is lower than many of the other organs,[57,58] and studies have shown that these viruses can infect other organs apart from the lungs[59,60]. The clinical presentation of both SARS-CoV and SARS-CoV-2 support this idea as symptoms of some infections include headache, vomiting, nausea, and diarrhea apart from the most widely observed symptoms of coughing and pneumonia[59,61]. Interestingly, a recent investigation showed that individuals infected with SARS-CoV-2 express higher levels of ACE2 in their respiratory tract relative to a healthy cohort[62]. Altogether, these studies indicate a critical role for ACE2 during SARS-CoV and SARS-CoV-2 infection. Following receptor binding, the virus is taken up in an endosome and the viral envelope fuses with the endosomal membrane thereby depositing the RNA into the host cytoplasm[63-65].

After entry, the viral genome can be directly translated due to its nature as a +ssRNA genome. The first two thirds of the genome is composed of two opening reading frames (ORFs): ORF1a and the full length ORF1a and ORF1b. The viral polyprotein is translated from these ORFs into polyprotein 1a and polyprotein 1ab. The production of the full-length polyprotein 1ab relies on a ribosomal frameshift at the end ORF of 1a which occurs at frequency of approximately 50%[66]. These two reading frames are translated into their respective polyproteins and make up the nonstructural proteins (nsps) encoded by the virus. The suboptimal efficiency of this frameshift generates a higher ratio of nsps found in ORF1a relative to those in ORF1b. The polyproteins are cleaved into individual nonstructural proteins by the two viral proteases, nsp3 and nsp5[67-70]. Nsp3 is specifically

responsible for releasing the three N-terminal proteins nsp1, nsp2, and nsp3 from the C-terminal end of polyprotein[71]. Nsp5 cleaves the polyprotein at 11 sites to produce the remaining mature viral proteins[68,72-75].

Upon proteolytic cleavage, nsp1 is the first viral protein released[76]. Nsp1 is particularly important in suppressing the host IFN response and degrades cellular RNAs thereby inhibiting host protein synthesis[77]. Nsp3, nsp4, and nsp6 have been implicated in the formation of DMVs[78-81]. DMVs are the site at which CoV replication takes place[48-51], however, the role of these organelles is not well understood beyond this. The formation of these replication organelles presumably benefits the virus by shielding the double stranded RNA (dsRNA) pathogen associated molecular patterns (PAMPs) from the host immune system[82,83]. Additionally, the presence of viral proteins and RNA to these sites may provide the requisite macromolecular concentration for viral RNA synthesis[83].

Nsp7 and 8 are presumed to be cofactors for the CoV RNA-dependent RNA polymerase (RdRP). The RdRP is more processive in the presence of nsp7 and 8[84]. These two proteins oligomerize to form a hexadecameric superstructure with an octamer of nsp8 held together by eight copies of nsp7. This complex forms an internal channel that is positively charged which suggests that it may play a role in nucleic acid binding[85]. While nsp12 is the primary RdRP used by CoVs for replication, nsp8 was found to encode a second RdRP. Though nsp8 does function as an RdRP, it can only generate short RNA products of fewer than 6 nucleotides similar to the function of primases[86,87]. Nsp12 polymerase activity is known to be primer-dependent *in vitro*[88]. While the biochemical activity of nsp8 has been studied extensively[85,89], its role during viral replication has not been established. One recent study attempted to investigate this by generating nsp8 mutant viruses in 229E. However, the authors were not able to recover viable mutant viruses[90]. This suggests a critical role for nsp8 during replication though the exact function of nsp8 remains unclear. However, because of its function as a secondary RdRP and primase nsp8 appears to be an obvious candidate to operate in conjunction with nsp12 during replication of the viral genome.

The products encoded by ORF1b contain several proteins with known enzymatic functions the most prominent of which is the RdRP encoded by nsp12. Nsp12 is made up of two domains: the canonical RdRP domain that occupies the C terminal half of the

protein and an N-terminal nidovirus RdRP-associated nucleotidyltransferase (NiRAN) domain[91]. Previous work has identified that activity of the NiRAN domain is essential for replication of SARS-CoV[91]. However, beyond this, the role of the nucleotidyltransferase in CoVs is unknown. CoV genome replication is carried out by the RdRP domain of nsp12 [92,93]. This domain contains canonical conserved sequence and structural subdomains of other viral RdRPs and reverse transcriptases and, expectedly, functions similarly[94,95]. Whether nsp12 can begin RNA synthesis *de novo* or requires a primer for initiation during infection is continues to be debated. As stated above, however, nsp12 requires a primer for initiation of RNA synthesis *in vitro*, and nsp8 remains the most likely candidate to generate primers for nsp12. Due to its critical role in the CoV replication cycle, nsp12 is a major target for antivirals and therapeutics including the well-described small molecules remdesivir and ribavirin[96-98].

Apart from using nsp7 and nsp8 as cofactors, nsp13 and nsp14 improve the functionality of nsp12 through their enzymatic activities. Nsp13 has both helicase and 5'-triphosphatase activity[99,100]. The helicase maps to the C-terminal end of the protein while the N-terminal end encodes for a zinc binding domain (ZBD) that is a unique genetic marker of the *Nidovirales*[101,102]. Based on work done in arteriviruses, another family within the *Nidovirales*[103], the ZBD of nsp13 in CoVs is believed to be necessary for helicase activity[104]. As with many other positive sense RNA viruses, the CoV helicase falls within the SF1 superfamily of helicases[105,106]. While several +RNA viruses encode helicases, the exact function of the enzyme during viral replication is not well understood. They are presumed to function in conjunction with the viral polymerase to improve processivity of the RdRP by unwinding double-stranded RNA replication intermediates and removing secondary structure downstream of the RTC[107]. As such, it is not unexpected that SARS-CoV nsp13 has been shown to interact with nsp12[108]. An outstanding question regarding the function of nsp13 during infection is the polarity of the enzyme. Nsp13 unwinds in a 5'-to-3' direction while the RdRP moves in a 3'-to-5' direction[106]. Consequently, it is unclear how these two proteins work concurrently. One hypothesis is that the helicase unwinds RNA behind the initial RTC but ahead of RTCs that may follow[100,106,109,110]. Additionally, the triphosphatase activity of nsp13 contributes to the capping of the viral genome[100].

Viruses within the *Nidovirales* encode for an exonuclease. Notably, exonuclease activity is exclusive to the order[111]. While nsp14 is not required for replication, nsp14 mutant viruses accumulate mutations at a higher rate than WT viruses[112]. Considering that CoVs have large genomes relative to other RNA viruses and the RdRPs are inherently error prone[113], this supports the idea that the exonuclease activity of nsp14 is important for proofreading during viral genome replication. Nsp14 interacts with the cofactor nsp10 to carry out its exonuclease activity[114,115]. As with nsp13, nsp14 has a secondary function as a 7-methyltransferase that coordinates capping of the viral genome[116,117]. Interestingly, while nsp10 interaction is required for exonuclease activity, this interaction is dispensable for nsp14 methyltransferase activity[114]. CoVs also encode for a second ribonuclease, nsp15, discussed in more detail below.

Nsp16, the final gene in ORF1b expresses a 2'-O-methyltransferase[118] and carries out the final step in capping of the viral genome[119]. This activity is contingent upon activation of nsp16 by nsp10[115]. While nsp16-mediated capping is known to be important to produce infectious virions, the relevance of nsp10 as a regulator in this process is currently understood[115].

Together, these nonstructural proteins make up the replication-transcription complex (RTC). However, it is still unclear how these proteins come together structurally to form the RTC though the RTC is believed to be comprised primarily of nsps7-16 as they are most enzymatically relevant with respect to the viral genome. One recent study used a computational approach and the known structures of many of the proteins to generate a model of the RTC[120]. This model is centered around a hexamer of the viral endonuclease, nsp15, with the other nsps surrounding it. Despite this finding, it remains to be seen whether this complex is formed as described in infected cells.

After initial translation of the non-structural proteins, the replication complex generates negative-sense RNAs which serve as templates for the positive-sense RNA genome. CoVs have evolved to transcribe a portion of the genome by discontinuous viral transcription[121-123]. During negative-sense RNA synthesis, the RTC is interrupted by TRS elements located in the 3' third of the genome at which point the RTC re-initiates RNA synthesis at the leader sequence in the 5' untranslated region (UTR) of the genome[122]. This results in a series of nested negative-sense sub-genomic RNAs (sgRNAs) that are

coterminal with the 3' end of the positive sense genome and each have a leader sequence of approximately 70 nucleotides at the 3' end of each negative sense subgenomic RNA[124,125]. The negative sense subgenomic RNAs are then used as templates for the production of positive-sense sgRNAs. The first ORF in each of these sgRNAs is translated into the structural proteins for the virus[122].

Among these structural proteins are spike (S), envelope (E), membrane (M), and nucleocapsid (N)[126]. N binds to viral RNA at the TRS elements and folds the viral genome into a helical structure[127,128]. Following translation, the structural proteins first translocate to the ER and subsequently to the membrane of the ER-to-Golgi intermediate compartment (ERGIC)[129,130]. Viral genomes are encapsidated by N-protein and bud into the ERGIC membrane at sites where these S, E, and M proteins are present thereby forming the envelope of the virion[130]. After assembly of the mature virion, viral particles exit the Golgi body and the cell by exocytosis. Recently, a secondary method of coronaviruses egress has been identified. Ghosh et al. found that MHV and SARS-CoV-2 can exit the cell through the lysosomal trafficking pathway. During this non-lytic egress process, the viruses translocate through lysosomes that are then deacidified and in which lysosomal proteases are inactive[131]. It is yet to be established how virions are transported to lysosomes and whether this is a mechanism that is conserved across all CoVs or limited to the two betacoronaviruses tested in this study. Following egress, the virus can go onto infect neighboring cells.

Coronaviruses as Etiologic Agents of Pandemics

In the past two decades, we have seen the emergence of three pandemics caused by CoVs. The first SARS outbreak began with a patient who was traveling from the Guangdong province of China to Hong Kong in February of 2003[12,132]. The index patient was staying the Metropole Hotel in Hong Kong before he was admitted to a hospital on February 22nd. The patient died the following day marking the first death from SARS-CoV infection. From this initial case, four healthcare workers at the hospital and two of the patient's family members became ill as well as twelve other individuals associated with the Metropole Hotel[12]. As hotel guests traveled home or to their next destinations, SARS-

CoV continue to spread with them. The disease eventually spread to 30 countries which would ultimately lead to more than 8000 cases and 900 deaths[132].

Patients infected with SARS-CoV generally presented with a persistent fever, dry cough, and myalgia[133]. In roughly 15% of these cases, patients progressed to acute respiratory distress syndrome (ARDS). ARDS is results in rapid respiratory failure and persistent hypoxemia which necessitates the support of a ventilator to supply oxygen[61,134]. Notably, this marked the first pandemic caused by a huCoV. The rapid spread of the virus from the inciting event as well as the severity of ARDS led to increased worldwide focus on CoV research.

SARS-CoV was initially believed to have transmitted to humans from palm civets due to 99.8% sequence identity between viruses isolated from humans and palm civets found in Chinese wet markets[11,135]. However, further surveillance did not detect the virus in wild and farmed palm civets[136]. Furthermore, infections of healthy civets with SARS-CoV resulted in overt symptoms in the animals further suggesting that they are unlikely to be the natural reservoir hosts[137]. Investigations turned to bats as a potential host reservoir as several viruses from different families have been isolated from bats[138]. One study showed that the genetic diversity of SARS-related coronaviruses (SARSr-CoVs) was far greater in bats than in palm civets indicating that SARS-CoV was likely circulating in bats prior to jumping to animals[139]. Interestingly, a surveillance study of a single cave in the Yunnan province of China sampled SARSr-CoVs from bats from 2011 to 2015. The viruses isolated from this investigation shared 93-96% sequence similarity to the civet and human isolates of SARS-CoV whereas to SARSr-CoVs in bats found in other regions of China only share 88-93% identity to SARS-CoV providing a potential location of origin of the virus[140]. Because these SARSr-CoVs were co-circulating in a single cave in the Yunnan province of China, there would be a higher likelihood of recombination occurring between them. Further recombination analysis showed that the closest progenitor to SARS-CoV was likely a recombinant strain from three SARSr-CoVs isolated during this study[140]. Presumably, this progenitor virus was first transmitted to palm civets before ultimately infecting humans.

Approximately 10 years following the SARS-CoV outbreak, the first cases of Middle Eastern respiratory syndrome (MERS) were found to be caused by the Middle

Eastern respiratory syndrome coronavirus (MERS-CoV)[13,141]. By 2015, MERS-CoV had infected 1100 people and resulted in over 400 deaths[142]. While there were far fewer cases of MERS than SARS, the 2012 spread of MERS-CoV in the Middle East represented the first coronavirus outbreak since the SARS-CoV pandemic in 2002. One reason as to why there were fewer cases of MERS relative to SARS is likely due to the inability of MERS-CoV to spread as well as SARS-CoV. One way in which virus transmissibility is measured is through its basic reproduction number or R_0 (R-naught) value. Three parameters determine the R_0 of an infectious agent: the duration of infectivity in a patient, the likelihood of transmission upon contact with another individual, and the contact rate between individuals[143,144]. As the R_0 of a pathogen increases, so too does the severity and length of the potential epidemic. MERS-CoV was estimated to have an R_0 value of 0.69 whereas SARS-CoV-1 had an R_0 of 0.80 indicating that the latter can spread more easily throughout the population[145]. This points to one possible explanation as to why there were a greater number of SARS cases.

As with SARS-CoV, bats are the suspected reservoir for MERS-CoV. Four species of bats found in the Middle East, *Taphozous perforates, Eidolon helvum, Rhinopoma hardwickii, and Pipistrellus kuhlii*, all contained closely related CoVs to MERS-CoV. One CoV found in *T. perforates* shared 100% sequence identity with the RNA-dependent RNA polymerase (RdRp) found in MERS-CoV. While the RdRp is a conserved gene in CoVs, this level of sequence identity between CoVs found in different species was unprecedented[146,147]. Further environmental sampling showed that *Nycteris gambiensis* bats in Ghana and *Nyctinomops laticaudatus* bats in Mexico too carried CoVs that were phylogenetically related to MERS-CoV; the latter shared 96.5% sequence identity at the amino acid level with MERS-CoV[147,148]. The existence of closely related CoVs to MERS-CoV in the Middle East strongly indicate that MERS-CoV arose from a progenitor virus in bats. Moreover, the presence of CoVs outside the region suggests that there may viruses with pandemic potential elsewhere in the world.

Though bats are likely the reservoir for MERS-CoV, lack of human to bat contact would suggest that there is an intermediate host for the virus. Several studies showed that antibodies against MERS-CoV were circulating in dromedary camels in several Middle Eastern and African countries including Saudi Arabia, Qatar, and Egypt[146,149-157].

Interestingly, these antibodies were found in serum samples from camels in this region dating back to the early 21st century suggesting that MERS-CoV jumped from bats to camels far earlier than the 2014 outbreak in humans[150,152,154,158]. Based on these serology studies, dromedary camels were determined to be the intermediate hosts of MERS-CoV. While the method of transmission of MERS-CoV from camels to humans remains to be determined, the first case of MERS in Yemen was identified in a man who had consumed raw camel milk. Despite this isolated instance, the most likely route of transmission is still considered to be respiratory[149].

COVID-19 Pandemic

Following the MERS-CoV outbreak in the Middle East, SARS-CoV-2 was identified in 2019. This novel CoV resulted in a devastating global pandemic in the years that followed. In late 2019, several cases of pneumonia were identified in Wuhan, China[52,59]. The disease outbreak is presumed to have started from the Huanan seafood market as many of the patients were associated the market[159-161]. Shortly afterwards, evidence for human-human transmission was found in Shenzhen, China. Six family members were admitted to the hospital after experiencing fever and symptoms of respiratory infection. Of these six patients, five had travelled to Wuhan. However, the sixth patient had not travelled and only became infected several days after contact with the other family members[16]. With the novel CoV able to spread by human-human transmission, the outbreak rapidly grew to infect over 2500 people in China along with 33 other individuals in 10 countries by the end of January 2020[52]. Along with pneumonia, these patients were presenting symptoms that included fever, dry cough, difficulty breathing and in serious cases, respiratory failure from alveolar damage[162,163]. Seven patients with severe pneumonia were connected to the seafood market where the outbreak began[52]. Samples from these individuals were analyzed, and it was found that one of these samples contained a virus that shared 79.6% sequence identity with SARS-CoV[52]. At the time, the World Health Organization (WHO) termed this the novel coronavirus 2019 (2019-nCoV).

Virus from further patient samples were sequenced and showed a 99.9% identity to 2019-nCoV at the nucleotide level. Though these viruses only shared approximately 80% identity with SARS-CoV, 2019-nCoV appeared to have several conserved replicase

domains that suggest that it is a SARSr-CoV[52]. Additional sequence analysis showed that 2019-nCoV shared 96.2% sequence identity with the bat CoV, RaTG13. Additionally, the nonstructural proteins in RaTG13 have 98.5% identity with those from the novel CoV[164]. Furthermore, the spike (S) protein from 2019-nCoV was highly divergent from other CoVs with the exception of RaTG13 to which it shared 93.1% sequence identity due to three short insertions in the *S* gene found only in these two CoVs[52]. Based on these data, RaTG13 is considered to be the closest known relative to the novel coronavirus once again suggesting that a pandemic causing CoV likely arose from bats[52].

Since then, the WHO has termed the novel coronavirus, SARS-CoV-2. By the end of 2022, over 650 million cases of SARS-CoV-2 had been reported, and the virus had claimed the lives of over 6 million people[165].

Viral Endonucleases

Many viruses have evolved to encode for viral endonucleases to aid in various processes during their replication cycle. These enzymes can be found in distantly related virus families including the *Herpesviridae*, *Orthomyxoviridae*, and *Coronaviridae*.

Herpes simplex virus-1 (HSV-1) uses the riboendonuclease, vhs, to degrade host and cellular mRNAs[166-169]. In conjunction with other viral proteins, the activity of vhs transitions the infected cell from host gene expression to viral gene expression[168]. HSV-1 lacking vhs are more sensitive to IFN-α and IFN-ß suggesting the endonuclease is important for immune evasion during infection[170]. *In vivo* studies have demonstrated that vhs is important for replication in a STAT1-dependent manner[171]. This provides further evidence that vhs is an important factor in antagonizing the immune response during infection.

KSHV encodes the viral protein SOX; in combination with the host exonuclease XrnI, SOX induces host translation shutoff[172-174]. Interestingly, SOX has two known functions, the first being an exonuclease that acts on viral DNA[175,176]. This DNase activity is important for infectious virion production and occurs specifically in the nucleus[174-176]. SOX also functions as a host translational shutoff factor. This activity is carried out by endonuclease activity that cleaves mRNA internally thereby allowing the cellular exonuclease XrnI to cleave these substrates from the accessible 5’ end of the mRNA.

SOX appears to preferentially cleave mRNAs generated by polymerase-II over other RNAs. Ultimately, SOX endonuclease activity leads to cellular translational shutoff and transitions the cell to translation of viral proteins during infection[172].

Nsp15

CoVs and other Nidoviruses encode for a uridine-specific endonuclease (nsp15 in CoVs) that is required for efficient viral replication and evasion of the host immune response[177,178]. Nsp15 was first described as a viral protein in SARS-CoV-1 and was shown to be conserved across other CoVs including MHV, 229E, and IBV[111]. The endonuclease is unique to the *Nidovirales* and is considered to be a marker for viruses in this order (nsp11 in arteriviruses)[179,180].

Nsp15 is unique in its catalytic function as a viral endonuclease because it is a uridine-specific endonuclease that is dependent on Mn^{2+} ions to cleave RNA[181,182]. Early studies of nsp15 catalytic activity in 229E and SARS-CoV-1 indicated not only that Mn^{2+} is required for optimal nsp15 activity, but also that another divalent cation cannot be substituted for Mn^{2+} [179,182]. Increasing concentrations of Mn^{2+} change the conformation of nsp15 suggesting that Mn^{2+} might be required to induce this conformational change in order to mediate cleavage[182]. However, a recent report indicated that SARS-CoV-2 nsp15 does not undergo a conformational change when exposed to Mn^{2+}[183]. Furthermore, Mn^{2+} has not yet been resolved in crystal structures of nsp15. While it is possible that Mn^{2+} impacts various species of nsp15 differently, ultimately, more work needs to be done to assess the role of Mn^{2+} in nsp15 cleavage.

Nsp15 oligomerizes as a hexamer and is catalytically most active in this state. A mutation at a glutamate residue at the N-terminal end of the protein prevents hexamer formation, and this mutant is impaired for cleavage activity relative to WT nsp15[184,185]. One exception to this is that porcine *Deltacoronavirus* (PDCov) nsp15 can only form dimers *in* vitro. As with other nsp15 homologs, this oligomerization appears to be driven by the N-terminal end of the protein, and PDCoV dimerization is also important for optimal cleavage activity[186]. While nsp15 requires Mn^{2+} for its catalytic activity, Mn^{2+} is not required for nsp15 oligomerization[184].

Initial characterization of nsp15 catalytic activity showed that it cleaves specifically after uridines. Snijder et al. found that SARS-CoV-1 nsp15 was homologous to an endonuclease (XendoU) from *Xenopus laevis*[111] which has been shown to cleave RNA substrates at poly uridine (polyU) sequences though no further similarities between the enzymes have been characterized[181]. This was the first indication that the viral endonuclease also had cleavage specificity. Since then, several studies have shown that nsp15 does indeed cleave specifically 3' of uridines[179,182,184,187,188]. Unlike XendoU, nsp15 does not require a polyU stretch for cleavage and can cleave after single uridine residues[179]. However, while previous studies canonically described nsp15 to cleave following uridines, recent investigations have shown that nsp15 has suboptimal cleavage efficiency following cytosines as well[188,189]. The difference in cleavability of the two ribonucleotides is presumed to be due to the inability of cytosines to seat as well in the catalytic pocket of nsp15 due to their larger amino group relative to the oxygen atom in the same location for uridines[189]. In addition to being able to cleave both uridines and cytosines, recent studies have shown that MHV nsp15 cleaves these ribonucleotides within certain dinucleotide contexts[188]. RNA sequencing analysis identified multiple cleavage sites at UA and CA sites throughout the MHV genome though the functional relevance of this activity has not yet been determined[188].

The specificity of uridine cleavage is driven by residues S294 and N278 in SARS-CoV-2 nsp15[190]. As modeled by cryo-electron microscopy (cryo-EM), both S294 and N278 are within hydrogen bonding distance of the scissile uracil base[190]. Previous work with SARS-CoV-1 nsp15 showed that mutating S294 to an alanine disrupts uridine cleavage specificity to the point where both uridines and cytidines are cleaved with similar efficiency[189].

A structural study of nsp15 identified a structure by cryo-EM of a dsRNA molecule interacting with the nsp15 hexamer[191]. The authors observed that a single uridine from the dsRNA substrate is positioned in the active site of one of the nsp15 monomers. The uridine is oriented away from the rest of the RNA in a "flipped" position relative to the substrate. This marked the first example of SARS-CoV-2 nsp15 cleaving a dsRNA substrate[191]. However, nsp15 cannot cleave dsRNAs as rapidly as ssRNAs[191]. The proposed base-flipping mechanism in this study is described in the context of paired

uridines where the W333 residue is important in mediating cleavage of dsRNA and likely the flipping of the uridine into the active site of the endonuclease. It remains to be seen whether this mechanism is specific to paired uridines or if nucleotides constrained by RNA secondary structure are also susceptible to base-flipping.

Additionally, Frazier et al. showed that, in the context of trinucleotide substrates, SARS-CoV-2 nsp15 cleaves substrates with uridines followed by a purine more proficiently than those within uridines followed by a pyrimidine[187]. Together, these studies indicate that while it has a strong preference for cleaving substrates 3' of uridines, nsp15 can be more promiscuous in the RNAs that it cleaves. To date, RNA sequencing analysis has not yet revealed an extended motif for nsp15 cleavage or binding[187,188].

One study of note looked at SARS-CoV-1 nsp15 cleavage activity using the stem-loop II motif (s2m) as a substrate[192]. The s2m motif is an RNA structural element found in the 3' UTR of SARS-CoV-1 and other RNA viruses including astroviruses, caliciviruses, and reoviruses[193]. The structure of s2m in SARS-CoV-1 has been resolved by x-ray crystallography, and though there are many uridines in s2m, only two are unpaired[194]. Bharadwaj et al. found that SARS-CoV-1 nsp15 specifically cleaves only one of these unpaired uridines indicating that the endonuclease does not ubiquitously cleave 3' of all accessible uridines. Rather, there are likely other factors that determine nsp15 cleavage specificity.

Whether this cleavage specificity has functional relevance during infection has yet to be determined. Recent work showed that in the astrovirus VA1, s2m is required to produce infectious virus[195]. Beyond this, however, though it is present in several different virus families, the role of s2m is largely unknown. Notably, newer variants of SARS-CoV-2 have partial or complete deletions of s2m suggesting that the virus can replicate in the absence of s2m[196]. Indeed, recent studies show that SARS-CoV-2 s2m is dispensable for replication *in vivo* and *in vitro*[197]. Interestingly, SARS-CoV-2 s2m contains two mutations relative to SARS-CoV-1 s2m resulting in a change in the secondary structure of the RNA. Despite the conservation of the s2m element across virus families, further studies are necessary to clarify its function as well as whether mutations between different s2m elements are relevant to viral fitness.

Nsp15 is an important factor during CoV replication. MHV lacking nsp15 catalytic activity is unable to replicate as efficiently as WT MHV specifically in the context of IFN signaling[178]. This and other previous studies strongly suggest that nsp15 antagonizes the host immune response. *In vitro* work has shown that nsp15 diminishes mitochondrial antiviral-signaling protein (MAVS)-mediated IFN signaling and prevents translocation of IRF3 to the nucleus[198,199]. Additionally, various substrates of nsp15 have been identified *in vitro and in vivo*. Nsp15 encoded by porcine epidemic diarrhea virus (PEDV) has been shown to cleave IRF3 and TBK1 *in vitro*[200]. In an infection setting, nsp15 encoded by MHV cleaves negative-sense viral RNA. Nsp15 targets polyU RNA at the 5' end of the negative-sense viral genome thereby preventing the formation of dsRNA and subsequent MDA5-mediated immune activation[177]. Other targets of nsp15 cleavage during infection have not yet been identified. Comar et al. showed in a recent study that while MERS-CoV nsp15 is not required for replication on its own, knocking out the endonuclease along with the NS4a and NS4b accessory proteins resulted in approximately a two-log decrease in viral replication[201]. NS4a and NS4b encode for a dsRNA binding protein and a phosphodiesterase respectively. Notably, SARS-CoV-2 lacks an NS4a homolog, and the NS4b homolog is expressed exclusively in the cytoplasm. In contrast MERS-CoV NS4b is expressed in both the cytoplasm and the nucleus as well in MERS-CoV[201]. Nevertheless, because it is conserved across CoVs, nsp15 represents an attractive target for pan-CoV antiviral therapeutics.

The structure of nsp15 has been solved by X-ray crystallography and cryo-electron microscopy for various CoVs[185,190,191,202,203]. As stated above, nsp15 is enzymatically functional as a hexamer, which is formed by the assembly of two back-to-back trimers[202,203]. Recent work has also identified a structure in which nsp15 is interacting with a dsRNA substrate. This structure indicates that the RNA spans across the two nsp15 trimers but is only cleaved in the active site in one trimer[191]. Because the nsp15 hexamer forms a structure with three-fold symmetry along the vertical axis (Figure 1.3), it has been proposed that multiple RNAs can bind to and be cleaved by nsp15 in isolation. However, this has not yet been shown structurally, and further work needs to be done to determine whether multiple RNAs can interface with the nsp15 hexamer.

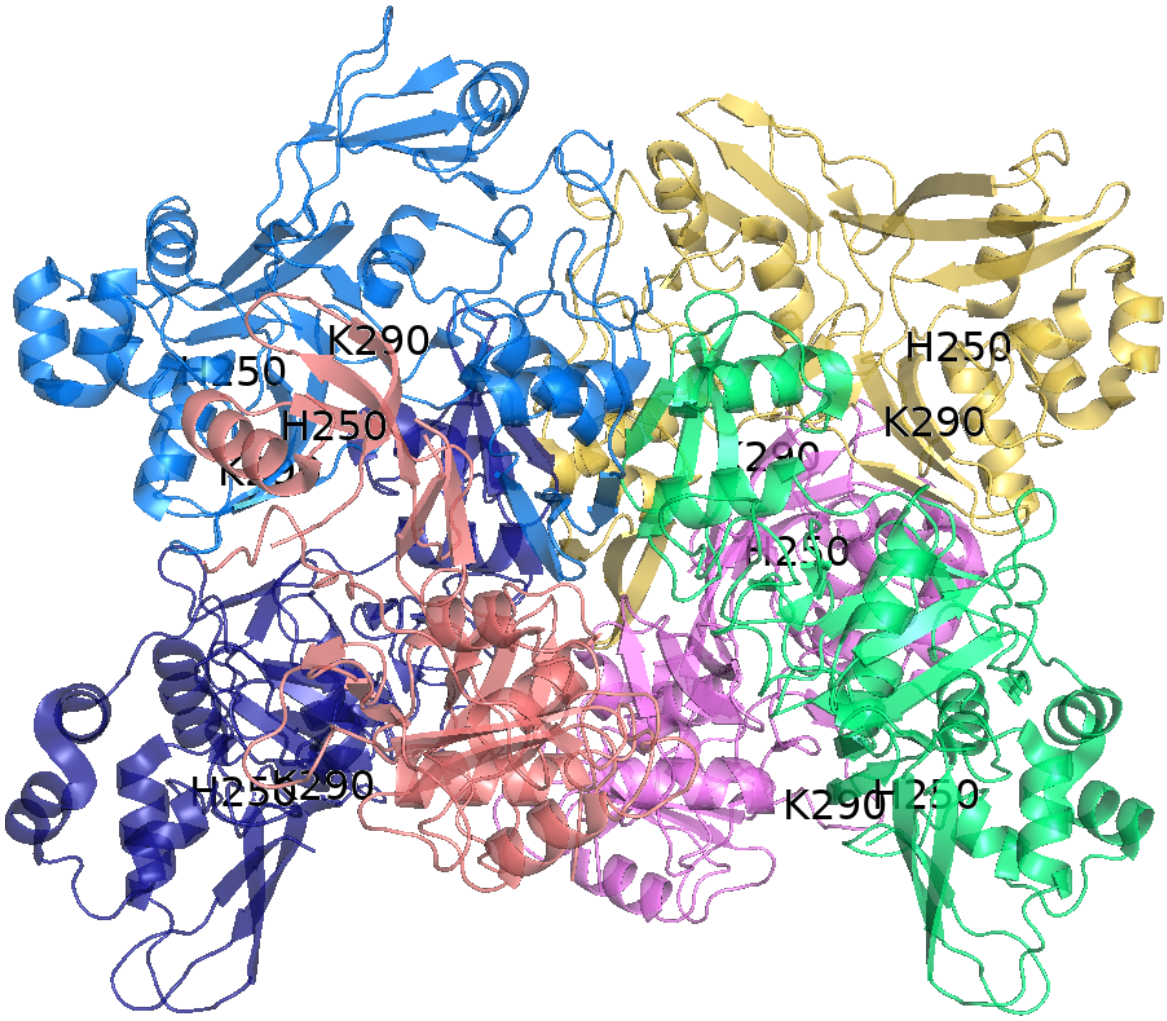


Figure 1.3 Nsp15 oligomerizes as a hexamer in its functional state. Nsp15 forms a hexamer of back-to-back trimers. Each nsp15 monomer is colored separately with the two trimers shown on the left (blue, pink, purple monomers) and right (gold, green, pink monomers). Two of the catalytic residues H250 and K290 are labeled for each monomer. RNA substrates interface with Nsp15 across the two trimers. While the hexamer has three-fold symmetry, nsp15 has not yet been shown to interact with multiple RNAs simultaneously.

Thesis work

Though the enzymatic function and structure of nsp15 has been established, the structural characteristics of nsp15 substrates beyond double-strandedness are yet to be determined. In my thesis, I sought to examine the specificity of nsp15 cleavage in RNA beyond uridine specificity in order to test the hypothesis that RNA secondary structure plays a central role in SARS-CoV-2 nsp15 cleavage activity. Through these studies, I found that RNAs with higher thermodynamic stability were protected from nsp15 cleavage and positioning of the scissile uridine within loop structures significantly impacts endonuclease activity *in vitro*. Further characterization of the s2m RNA from the 3' UTR of SARS-CoV (described in further detail below) revealed that uracil position within the pentaloop greatly impacts nsp15 cleavage efficiency. Understanding the replication process of CoVs, including the enzymatic function of nsp15, will provide valuable insight for the development of drug targets and a deeper understanding of CoV pathogenesis.

Chapter II: Regulation of coronavirus nsp15 cleavage specificity by RNA structure

Adapted from accepted manuscript Salukhe, I., et al. (2023). "Regulation of coronavirus nsp15 cleavage specificity by RNA structure." PLoS One **18**(8): e0290675.

Introduction

Coronaviruses (CoVs) are a diverse group of enveloped, positive sense single-stranded RNA (+ssRNA) viruses within the *Nidovirales* order. CoVs infect a wide range of host species and are associated with mild to severe disease in livestock and humans[22,52,204]. The *Coronaviridae* family is composed of four genera: Alpha-, Beta-, Gamma-, and Deltacoronaviruses[21]. Since the turn of the century, several pathogenic CoVs have emerged in the human population including SARS-CoV-1, MERS-CoV, and SARS-CoV-2 with the latter being the causative agent of the COVID-19 pandemic[22,52].

Replication of the large (~30kb) CoV genome is facilitated by multiple nonstructural proteins (nsp1-16) which are produced from two open reading frames encoded by the first two thirds of the viral genome. These proteins encode several key functions necessary for synthesis and post-translational processing of the viral genome, including RNA-dependent RNA polymerase (nsp12), helicase (nsp13), RNA M7 and 2'-O-methyltransferase (nsp14, nsp16/nsp10), protease (nsp3, nsp5), and exonuclease (nsp14) activities. Structural proteins which assemble to form the virion, as well as accessory proteins, are encoded in a set of nested subgenomic RNAs which are coterminal with the 3' end of the genome.

The viral replication/transcription complex (RTC) is housed within membrane structures derived from the host endoplasmic reticulum (ER) which are induced by expression of nsp3, 4, and 6[48-51,78-81]. Like other +ssRNA viruses, these double-membrane vesicles (DMVs) provide a microenvironment that supports efficient viral replication and sequester double-stranded RNA (dsRNA) replication intermediates from detection by host pathogen recognition receptors (PRRs)[82,83]. Previous studies have demonstrated a key role for the interferon (IFN) response in restriction of CoVs. Accordingly, CoVs have evolved several mechanisms to evade antiviral immunity (e.g., 2'-O-methylation) and encode several proteins which antagonize the IFN response (nsp1,

ORF6)[77,205,206]. Nsp15 is a uridine-specific endonuclease (endoU) encoded by CoVs and other nidoviruses which has been implicated in evasion of host-mediated RNA sensing and IFN activation[177,178]. While nsp15 has been shown to be non-essential for viral replication of several CoVs, absence of nsp15 endoU activity (Δnsp15) leads to greater accumulation of dsRNA replication intermediates which trigger IFN activation in immune competent cells such as macrophages[177,178,207]. During CoV infection, dsRNA sensing occurs primarily via MDA-5, though other RNA sensors [e.g., protein kinase R (PKR)] have also been implicated[178,208].

While originally described as a uridine-specific endonuclease, MHV nsp15 has been shown to preferentially cleave pyrimidine-adenine dinucleotides (U↓A or C↓A)[188]. More recently, viral targets of MHV nsp15 have been identified in both the polyU tract of 5' negative-sense RNA, as well as at several sites throughout the CoV genome[177,188]. Although multiple nsp15 cleavage sites have been mapped throughout the MHV genome, no other sequence determinants have yet been identified, and despite the abundance of pyrimidine-adenine dinucleotides across the MHV genome, only a fraction of these sites is cleaved by nsp15. This indicates that other RNA determinants likely affect cleavage specificity of viral RNA. Earlier *in vitro* studies have shown that SARS-CoV-1 nsp15 cleaves a single unpaired uridine in the stem-loop II motif (s2m) RNA structural element found in the 3'-UTR of the viral genome despite the presence of other uridines, paired and unpaired. This would suggest that the structural context in which the scissile uridine is presented may play a role in determining nsp15 cleavage specificity[192]. As the comparable s2m sequence in SARS-CoV-2 changes the secondary structure of the RNA, we chose to use SARS-CoV-1 s2m as a model RNA to explore the secondary structure requirements for SARS-CoV-2 nsp15 cleavage. Our findings reveal that thermodynamically stable RNAs are protected from SARS-CoV-2 nsp15 cleavage and cleaved less efficiently. We show for the first time that the thermodynamic stability of RNA structure as well as positioning and sequence context of the scissile uridine play a role in determining SARS-CoV-2 nsp15 cleavage specificity. This is consistent with the results of previous studies of SARS-CoV-1 nsp15[192], suggesting that SARS-CoV-2, and CoVs more broadly, may use RNA structure to regulate nsp15 cleavage specificity. These data

shed further light on requirements for viral endoU target recognition and demonstrate an underappreciated role for viral RNA structure in evasion of innate immunity.

Results

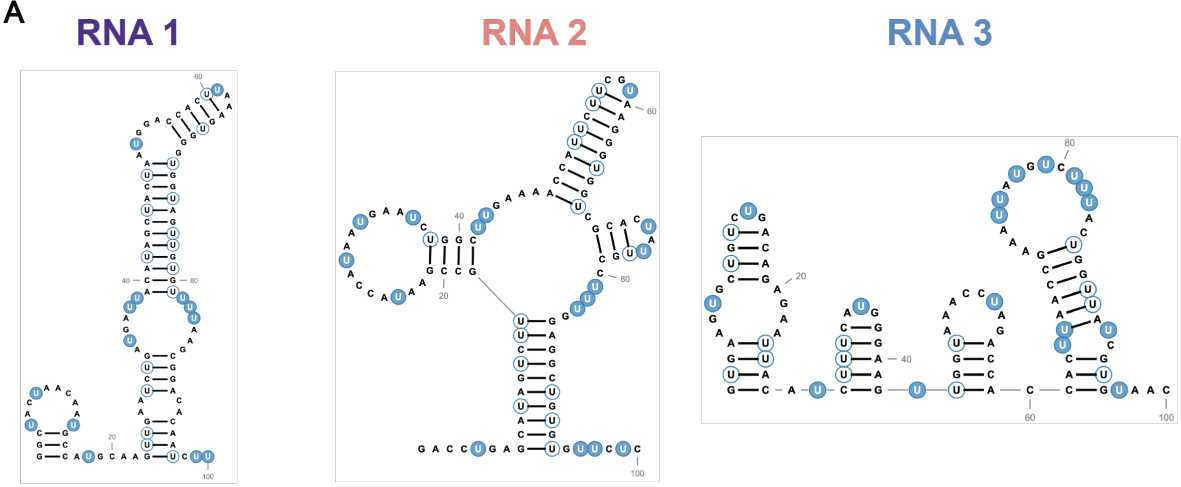
RNA secondary structure modulates nsp15 cleavage efficiency

Nsp15 cleavage sites have previously been mapped in the MHV genome following viral infection[177,188]. Despite the abundance of pyrimidine-adenine dinucleotides throughout the genome, MHV nsp15 was found to cleave only a fraction of these sites suggesting that other determinants play a role in CoV nsp15 cleavage specificity. Using the MEME-suite XSTREME motif analysis and discovery tool[209] we analyzed 200 nt sequences corresponding to 100 nt upstream and downstream of the MHV nsp15 cleavage sites identified by Ancar et al. to identify shared motifs that might contribute to specific recognition or cleavage of these sequences by nsp15. Although some motifs were found to be enriched in sequences surrounding the MHV nsp15 cleavage sites, they were not uniformly conserved across all cleavage sites and were also found at other sites not previously identified as nsp15 targets. However, other motifs and RNA structures that are not identified by XSTREME may be present, and further investigation at these sites is needed to determine a link between nsp15 motifs and cleavage.

As RNA structure has previously been implicated in nsp15 cleavage specificity[192], we next compared the underlying predicted secondary structure of the same 200 nt fragments using LocRNA software[210,211]. We did not identify any conserved structural elements at nsp15 cleavage sites in MHV nor did we identify conserved structural elements upstream or downstream of these sites. Despite the absence of any conserved sequence or structural motifs, we hypothesized that the presence of thermodynamically stable RNA elements would prevent cleavage of RNA by nsp15.

To test this hypothesis, we used RNAfold[212] to identify regions across the SARS-CoV-2 genome predicted to have varying degrees of thermodynamic stability (RNA 1, RNA 2, RNA 3; Fig 2.1, Table 2.1)[213,214]. We identified representative regions of high, moderate, and low thermodynamic stability as indicated by their relative ΔG values of the minimum free energy (MFE) (Fig 2.1). RNAs with higher ΔG values have more stable predicted structures than those with lower ΔG values[215].

A



B

RNA	SARS-CoV-2 gene	Thermodynamic Stability	ΔG of Minimum Free Energy (kcal/mol)	Frequency of MFE	Ensemble Diversity
1	*nsp3*	High	-27.8	15.0%	4.81
2	*nsp2*	Moderate	-23.4	0.9%	25.80
3	*nsp12*	Low	-14.2	1.2%	39.93

Figure 2.1 The SARS-CoV-2 genome contains regions of varying RNA structure. **A.** Predicted minimum free energy structures (RNAfold) of 100 nt RNAs from the SARS-CoV-2 genome of varying thermodynamic stability (high, medium, and low). Unpaired uridines are highlighted in blue, and paired uridines are outlined in blue. All RNA structures were predicted in RNAfold and designed in RNA2Drawer. **B.** The locations within the SARS-CoV-2 genome are indicated for each structure, and thermodynamic stability is indicated as -ΔG. Ensemble diversity is listed as predicted by RNAfold.

Table 2.1 Sequences of RNAs 1-3 from SARS-CoV-2

RNA	Sequence
RNA 1 (3496-3595)	GGCUACUAAC AAUGCCAUGC AAGUUGAAUC UGAUGAUUAC AUAGCUACUA AUGGACCACU UAAAGUGGGU GGUAGUUGUG UUUUAAGCGG ACACAAUCUU
RNA 2 (1401-1500)	GACCUGAGCA UAGUCUUGCC GAAUACCAUA AUGAAUCUGG CUUGAAAACC AUUCUUCGUA AGGGUGGUCG CACUAUUGCC UUUGGAGGCU GUGUGUUCUC
RNA 3 (16700-16799)	GUGAAGUGCU GUCUGACAGA GAAUUACAUC UUUCAUGGGA AGUUGGUAAA CCUAGACCAC CACUUAACCG AAAUUAUGUC UUUACUGGUU AUCGUGUAAC

Additionally, we compared the predicted RNA structures to previously published SHAPE-MaP data (Fig 2.2)[216]. RNA 1, the most thermodynamically stable RNA, shared significant similarity to these data. However, we observed less correlation between the RNAfold predictions and SHAPE-MaP data for RNA 2 and particularly RNA 3. This is to be expected for these two RNAs, as they have higher ensemble diversities (Fig 2.1) and, as such, are predicted to adopt a larger number of conformers than RNA 1. Therefore, it follows that the predicted and empirical structures are not strongly correlated.

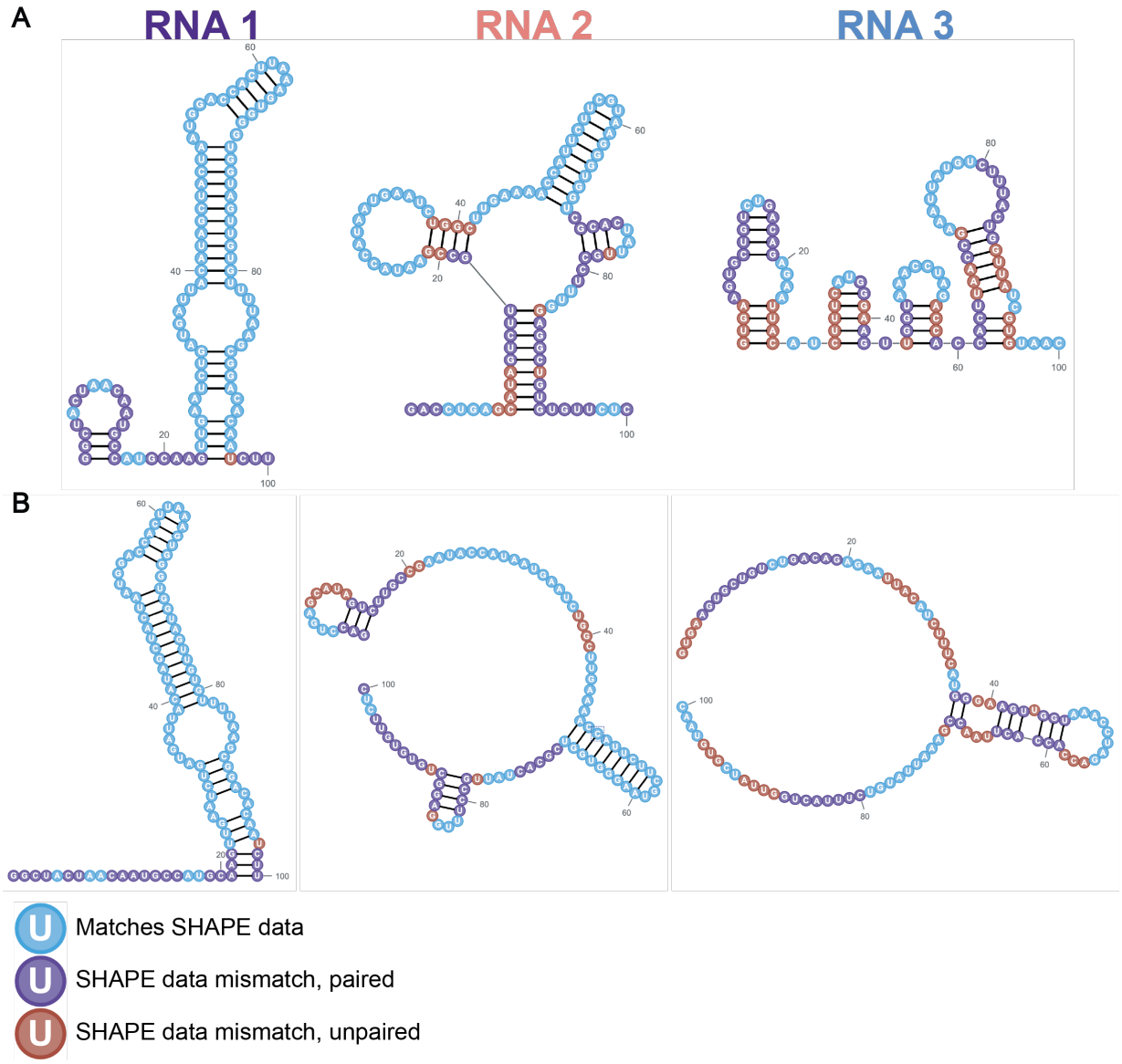


Figure 2.2 Correlation between RNAfold predictions and SHAPE-MaP data. A. Minimum free energy structures of RNAs 1, 2, and 3 as predicted by RNAfold. **B.** Published structures of RNA 1, 2, and 3 as resolved by SHAPE-MaP (Huston et al. *Molecular Cell* 2021). Nucleotides highlighted in blue match with SHAPE data while bases highlighted in purple and red do not. Purple nucleotides are paired with a different nucleotide whereas red nucleotides are unpaired in SHAPE structures.

100 nt RNAs corresponding to these regions were synthesized *in vitro* and cleavage efficiency of these RNAs was assessed using a SARS-CoV-2 nsp15 endonuclease assay[217] (Fig 2.3). Folded RNAs were incubated with recombinant SARS-CoV-2 nsp15 for various times (1, 15, 60, and 240 minutes) and cleavage products were analyzed by polyacrylamide gel electrophoresis (Fig 2.3). Cleavage was quantified by densitometry by comparing the abundance of full-length RNA at each time point to a denatured nsp15 control in which the RNA was incubated with non-functional, denatured nsp15 (no cleavage, (-) far-right lane) (Fig 2.3). As we hypothesized, we observed that RNAs with relatively moderate or low predicted thermodynamic stability were cleaved more rapidly than RNA 1 with relatively high thermodynamically stability (Fig 2.3, compare abundance of full-length transcript at 1, 15 and 60 min [red arrow]; Fig 2.3 compare percentage uncleaved RNA). Notably, RNA with relatively low thermodynamic stability showed rapid cleavage as early as 1 min following the addition of nsp15. While each RNA yielded different cleavage products owing to the differences in RNA sequence, we observed the appearance of distinct cleavage products that either decreased (green arrows) or increased (blue arrows) in abundance over time (Fig 2.3). These likely represent cleavage intermediates or end products respectively and suggest that other structural features or specific sequences also likely contribute to the higher efficiency of nsp15 cleavage of some RNA sequences over others. Collectively, these data suggest that RNA secondary structure impacts nsp15-dependent RNA cleavage.

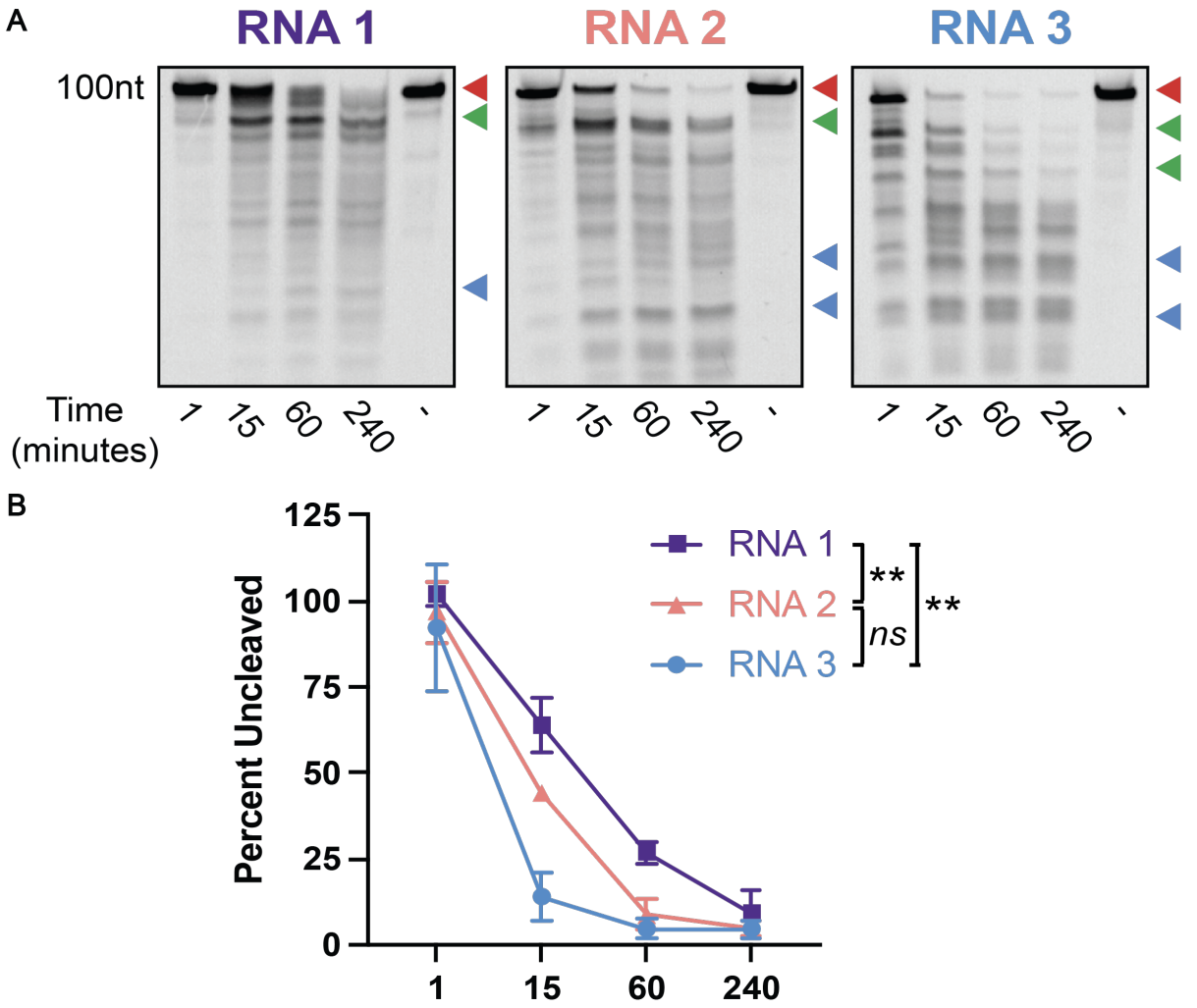
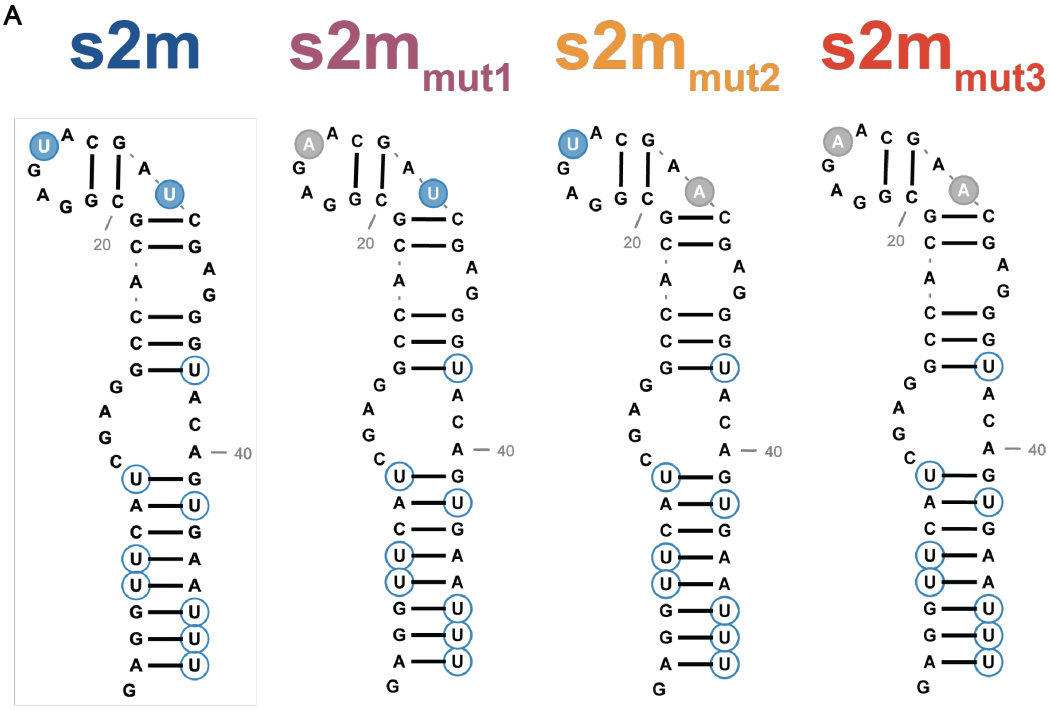


Figure 2.3 Secondary structure protects RNA from cleavage by SARS-CoV-2 nsp15. A. Endonuclease assays of RNAs from (A). Full-length (uncut) RNAs are indicated by the red arrows. Diminishing cleavage products (cleavage intermediates) are indicated by green arrows. Accumulating cleavage products (cleavage end products) indicated by blue arrows. Representative images from one of three total experiments are shown. **B.** Quantitation of data from (A). The percentage of full-length RNA remaining was measured by densitometry. Most thermodynamically stable RNA (RNA 1) was cleaved least rapidly relative to other RNAs. Percentage of uncut RNA was calculated by normalizing to a denatured nsp15 control (indicated by (-) in the far-right lane). Area under the curve (AUC) was calculated for each RNA and one-way ANOVA with multiple comparisons was performed on the AUC. Data represents three independent experiments.

SARS-CoV-2 nsp15 cleaves unpaired bases in a structured RNA

In vitro biochemical studies by Bhardwaj et al. previously showed that SARS-CoV-1 nsp15 specifically cleaves 3' of unpaired uridines in s2m, a conserved structural element found in the 3'UTR of the SARS-CoV-1 genome and some other related CoVs[192,194,218]. In addition to our results described above (Fig 2.3), this study suggests that the context in which the cleaved uridine is presented may be important for determining the cleavage specificity of nsp15. In the case of s2m, the two unpaired uridines (U25 and U30) are located in two unstructured loops (Fig 2.4, highlighted in blue). Structural resolution of s2m has shown that U25, which is located within the pentaloop of s2m, is highly disordered relative to other nucleotides, and that U30 is also moderately disordered[194]. The s2m pentaloop adopts the GNRNA sequence, and these loops have been shown to promote nucleotide extrusion or base-flipping[219,220]. of Interestingly, while there is significant homology between s2m from SARS-CoV-1 and SARS-CoV-2, SARS-CoV-2 s2m has two single nucleotide polymorphisms (SNPs) that change the structure of the RNA (Fig 2.5). Based on previous findings and the availability of structural information, we posited that SARS-CoV-1 s2m would be a tractable model RNA for studying the determinants of cleavage specificity of SARS-CoV-2 nsp15.

A

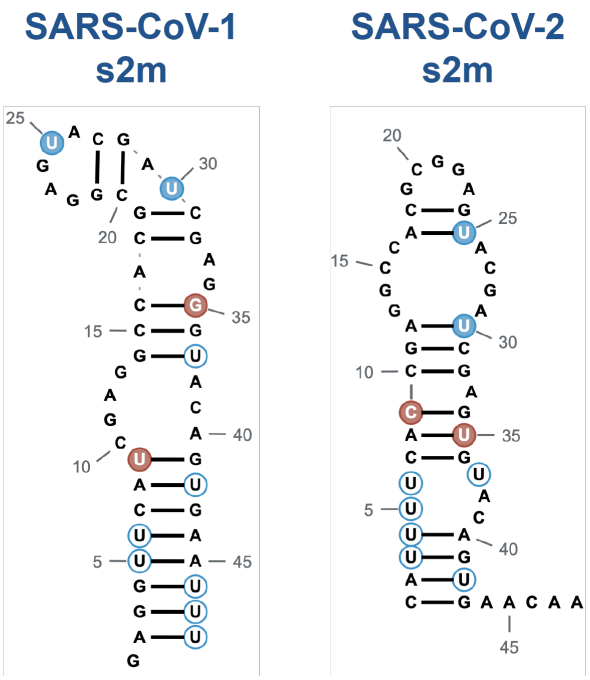


B

RNA	ΔG of Minimum Free Energy (kcal/mol)	Frequency of MFE	Ensemble Diversity
s2m	-9.4	25.2%	6.59
$s2m_{mut1}$	-9.4	27.5%	4.19
$s2m_{mut2}$	-9.4	26.1%	6.32
$s2m_{mut3}$	-9.4	28.7%	3.64

Figure 2.4 Thermodynamic stability of wt and mutant s2m RNAs. A. Wt and mutant s2m RNAs with unpaired uridines are highlighted in blue, and paired uridines are outlined in blue. All RNA structures were predicted in RNAfold and designed in RNA2Drawer. **B.** Thermodynamic stability and ensemble diversity of s2m and mutant RNAs indicated as predicted by RNAfold.

A



B

RNA	ΔG of Minimum Free Energy (kcal/mol)	Frequency of MFE	Ensemble Diversity
SARS-CoV-1 s2m	-9.4	25.2%	6.59
SARS-CoV-2 s2m	-6.9	12.4%	14.12

C

Figure 2.5 Alignment of SARS-CoV1 and SARS-CoV-2 s2m. A. Minimum free energy structures of SARS-CoV-1 and SARS-CoV-2 s2m. The U25 (pentaloop) and U30 (bulge) nucleotides are highlighted in blue. Two nucleotide changes between the RNAs in the stem (highlighted in red) change the MFE structure. **B.** Thermodynamic stability, frequency of MFE, and ensemble diversity of s2m from each virus. Increased ΔG and ensemble diversity in SARS-CoV-2 s2m indicate that it does not maintain the highly conserved structure of SARS-CoV-1 s2m. **C.** Alignment of s2m from both viruses with relevant polymorphisms in SARS-CoV-2 s2m indicated by red arrows.

We first sought to determine whether SARS-CoV-2 nsp15 exhibits cleavage specificity for unpaired uridines in s2m similar to what was previously observed for SARS-CoV-1 nsp15[192]. To identify whether nsp15 preferentially cleaves at U25, U30, or both, we designed the wild type (wt) SARS-CoV-1 s2m RNA, as well as single and double s2m mutants containing U-to-A substitutions at U25 and U30 (Fig 2.4) and compared cleavage efficiency of these RNAs via endonuclease assay (Fig 2.6). In contrast to wt s2m, cleavage of the double mutant ($s2m_{mut3}$) was noticeably impaired consistent with previous studies showing nsp15 cleavage preference for unpaired uridines[187,192]. However, in contrast to previous findings with SARS-CoV-1 nsp15, we observed that mutation of U30 (U30A; $s2m_{mut2}$) did not impact RNA cleavage, whereas mutation of U25 (U25A; $s2m_{mut1}$) did[192]. Although quantitative differences in full-length RNA cleavage were not found to be statistically significant (Fig 2.6), accumulation of cleavage products in the U25A ($s2m_{mut1}$) and double mutant ($s2m_{mut3}$) trended towards relative to wt and U30A ($s2m_{mut2}$) RNAs.

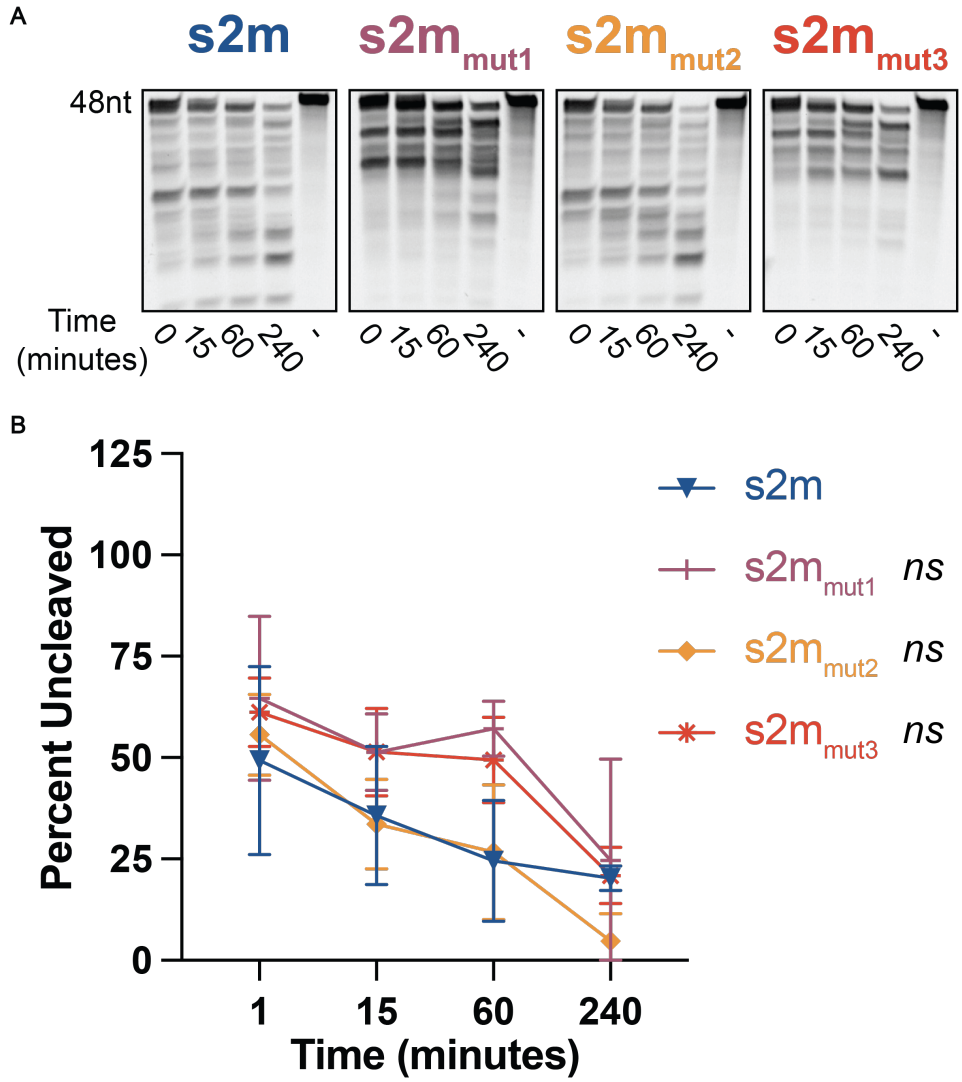


Figure 2.6 SARS-CoV-2 nsp15 cleaves structured RNAs at specific sites. A. Endonuclease assays of full-length s2m mutant RNAs. Full-length bands indicated at the top of the gel with cleavage products below. Representative images from one experiment. **B.** Percent of full-length RNA remaining as measured by densitometry. Percentage of uncut RNA was calculated by normalizing to a denatured nsp15 control (indicated by (-) in the far-right lane). RNAs with uridines present in the pentaloop (s2m and $s2m_{mut2}$) cleaved more efficiently than $s2m_{mut}1$ and $s2m_{mut3}$. AUC was calculated for each RNA and one-way ANOVA with multiple comparisons was performed on the AUC. AUC of s2m mutant RNAs are compared to wt s2m. Data represents three independent experiments.

We analyzed the dominant cleavage products from these assays by measuring these bands by densitometry (Fig 2.7, indicated by numbered arrows). Each product was normalized to its corresponding band at the one minute timepoint of the reaction. We found that the cleavage products from the s2m and $s2m_{mut2}$ RNAs not only shared a similar banding pattern, but the cleavage products were also visible at comparable times in the reaction (Fig 2.7). Interestingly, we also observed accumulation of distinct cleavage products in $s2m_{mut1}$ and $s2m_{mut3}$. Cleavage of these RNAs also result in a similar output when measuring the products by densitometry (Fig 2.7) further corroborating that differential cleavage of s2m is driven by U25 in the pentaloop. This may indicate that step wise cleavage of RNA and RNA intermediates is altered in the absence of the scissile U25. However, while mutation of U25 was not predicted to impact the structure of s2m using RNAfold analysis (Fig 2.4), we cannot rule out the possibility that mutation of U25 leads to changes in s2m RNA structure which impact nsp15-mediated cleavage.

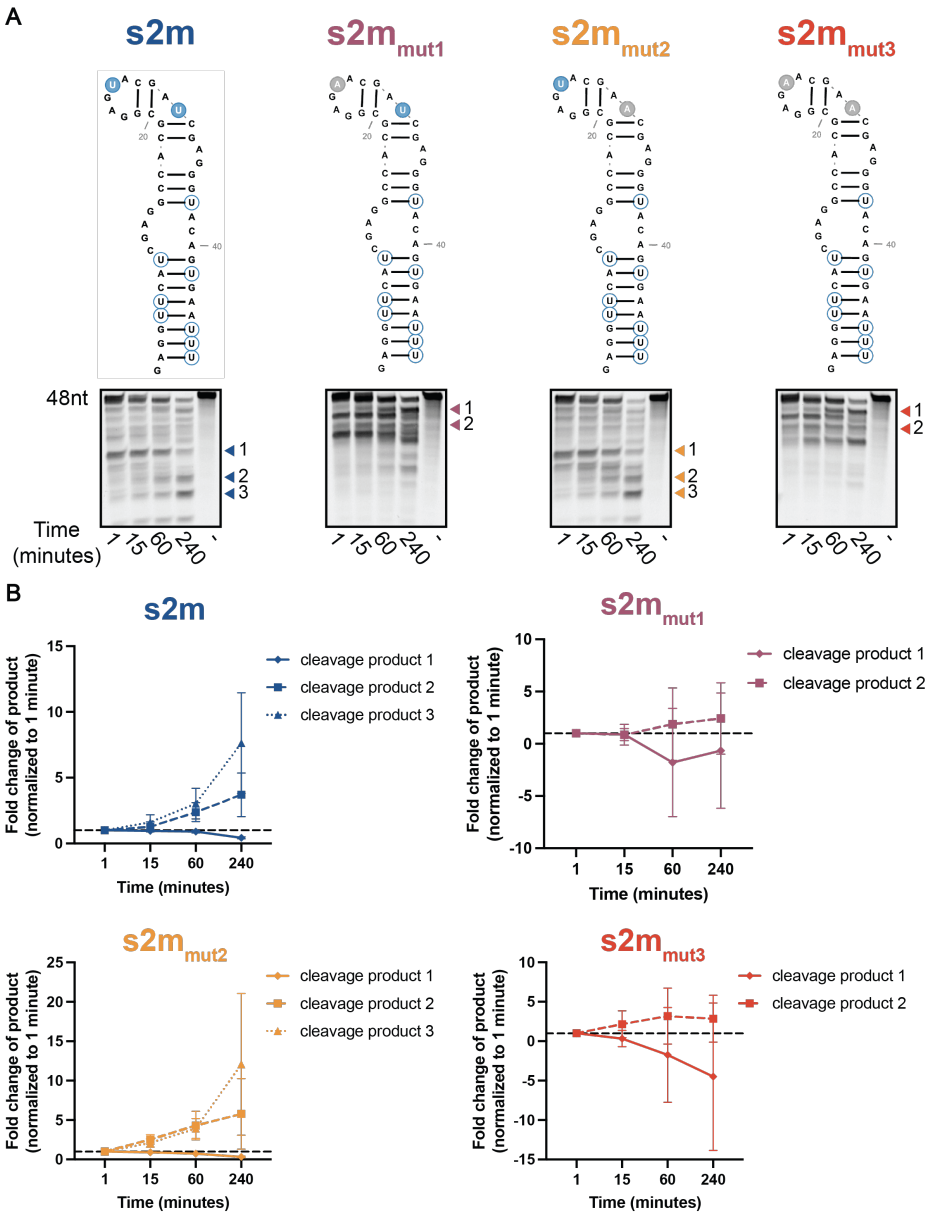
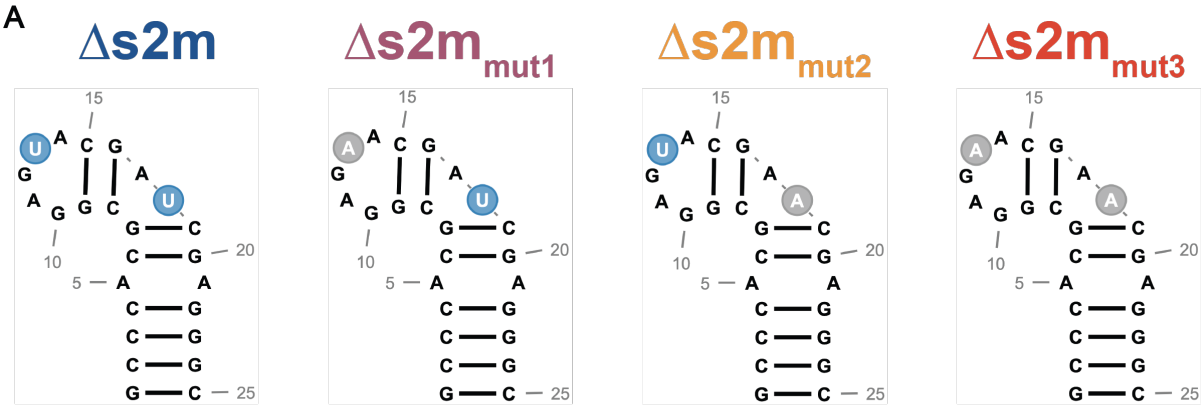


Figure 2.7 Cleavage products of s2m RNAs differ based on mutations in the pentaloop and bulge. Endonuclease assays of full-length s2m mutant RNAs as shown in Fig 2. Full-length bands indicated at the top of the gel with cleavage products below. Selected cleavage products are numbered and indicated by arrows. Representative images from one experiment. **B.** Fold change in each cleavage product relative to the amount present at the one minute time point. Cleavage products above the dashed lines indicate accumulating products while those below are diminishing over the course of the reaction. Data represents three independent experiments.

Due to the presence of multiple uridines in s2m, accurate identification of cleavage products and intermediates was challenging. In order to eliminate extraneous cleavage products to clarify interpretation of endonuclease data and further dissect RNA structural determinants of cleavage efficiency, we designed truncated s2m RNAs (Δs2m) which lack the base of the stem and all uridines except U25 and U30 (Fig 2.8). While there are nucleotide changes in the stem of truncated RNAs, these mutations help to maintain the pentaloop and bulge structure found in s2m. Similar to the experiments described above, we also engineered Δs2m mutants with U-to-A substitutions at U13 ($\Delta s2m_{mut1}$) and U18 ($\Delta s2m_{mut2}$) (equivalent to U25 and U30 in wt s2m), or both U13 and U18 ($\Delta s2m_{mut3}$). As with the full-length s2m RNAs, the free energies and ensemble diversities of the ∆s2m mutant RNAs were similar to each other (Fig 2.8).

A



B

RNA	ΔG of Minimum Free Energy (kcal/mol)	Frequency of MFE	Ensemble Diversity
Δs2m	-7.9	75.7%	1.07
$\Delta s2m_{mut1}$	-7.9	75.8%	1.07
$\Delta s2m_{mut2}$	-7.9	73.7%	1.14
$\Delta s2m_{mut3}$	-7.9	73.7%	1.13

Figure 2.8 Thermodynamic stability of Δs2m RNAs. A. Wt and mutant Δs2m RNAs with unpaired uridines are highlighted in blue. All RNA structures were predicted in RNAfold and designed in RNA2Drawer. **B.** Thermodynamic stability and ensemble diversity of Δs2m and mutant RNAs indicated as predicted by RNAfold.

We analyzed cleavage of wt and mutant Δs2m RNAs over time by endonuclease assay as described above (Fig 2.9). In keeping with experiments with full-length s2m RNA, we observed rapid accumulation of cleavage products in wt Δs2m and $\Delta s2m_{mut2}$ RNAs. In contrast, mutation of both U13 and U18 (U25 and U30 equivalent), or U13 (U25 equivalent) alone, resulted in delayed and reduced nsp15 cleavage of RNA. Interestingly, we still observed some modest cleavage of $\Delta s2m_{mut1}$ suggesting that while U13 is the predominant cleavage site in the Δs2m RNA, nsp15 is also able to recognize and cleave at other sites. While mutation of both U13 and U18 led to a further reduction in nsp15-

mediated cleavage, some cleavage products were still detected in the double mutant lending further support to the hypothesis that nsp15 can also cleave non-uridine bases, albeit at lower efficiency. Indeed, while nsp15 displays a strong preference of cleavage 3' of uridines, previous work has shown that cleavage at other bases (especially cytosine) also occurs[187].

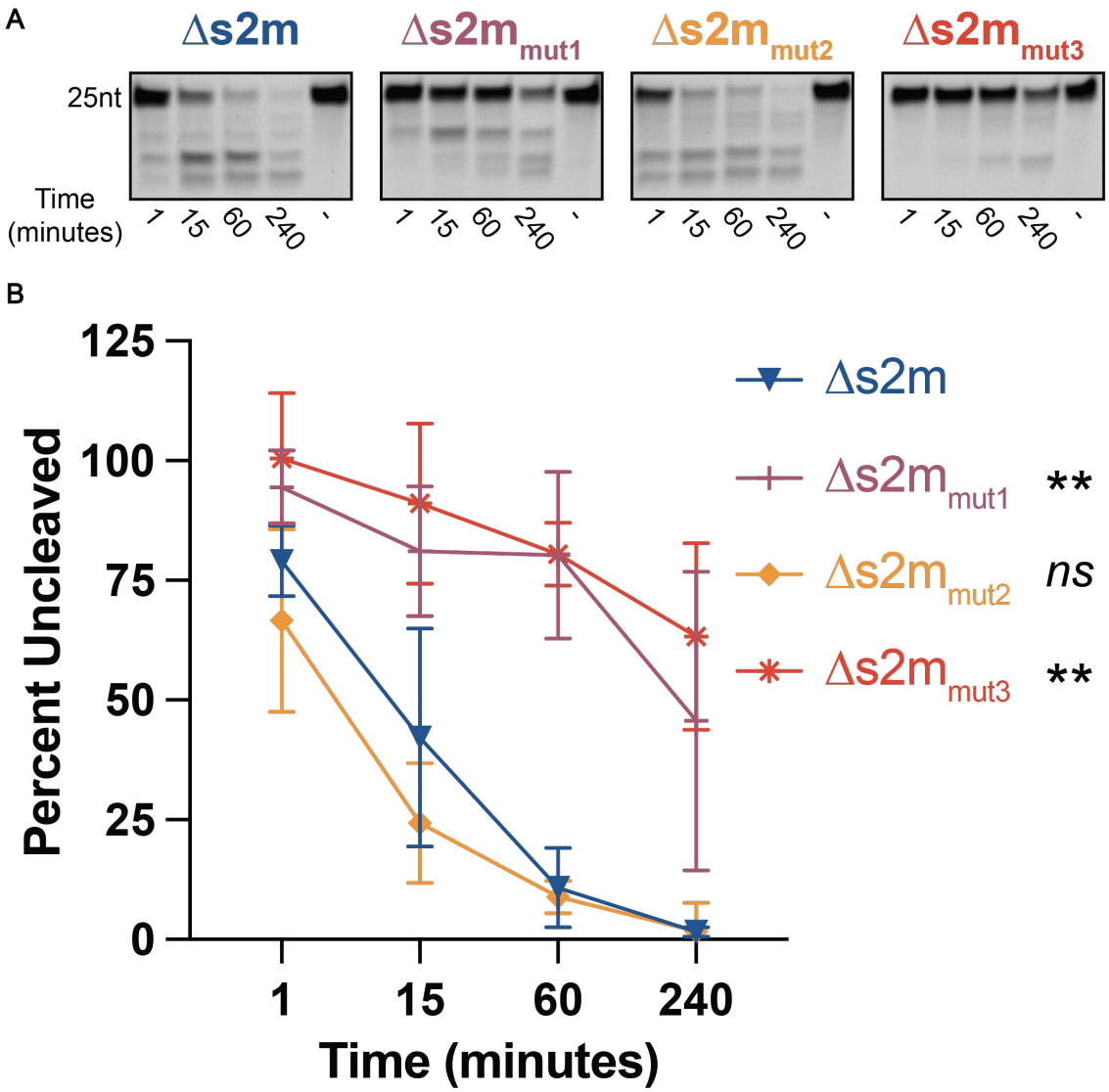


Figure 2.9 SARS-CoV-2 nsp15 cleaves structured RNAs at specific sites. A. Endonuclease assays of modified s2m RNAs. Full-length bands indicated at the top of the gel with cleavage products below. Representative images from one experiment. **B.** Percent of full-length RNA remaining as measured by densitometry. Percentage of uncut RNA was calculated by normalizing to a denatured nsp15 control (indicated by (-) in the far-right lane). RNAs with uridines present in the pentaloop (Δs2m and $\Delta s2m_{mut2}$) cleaved more efficiently than $\Delta s2m_{mut}1$ and $\Delta s2m_{mut3}$. AUC was calculated for each RNA and one-way ANOVA with multiple comparisons was performed on the AUC. AUC of Δs2m mutant RNAs are compared to wt Δs2m. Data represents three independent experiments.

We hypothesized that RNA secondary structure contributes to nsp15 cleavage specificity by presenting the scissile uridine in a structural context favorable for nsp15 cleavage. However, an alternative explanation could be that RNA structure serves to sterically hinder or promote nsp15-RNA binding. To address whether the differences we observed in cleavage of wt and mutant Δs2m RNAs was due to differences in protein-RNA binding alone, we compared nsp15-RNA binding affinities using differential radial capillary action of ligand assay (DRaCALA)[221,222]. Here, 5'-end radiolabeled Δs2m RNAs were incubated with increasing concentrations of recombinant mutant nsp15 (nsp15 (K290A)), complexes applied to nitrocellulose membrane, and the fraction of bound vs unbound RNA quantified (Fig 2.10). Nsp15 (K290A) contains a mutation of one of the conserved catalytic residues necessary for endoU activity and was used to eliminate endonuclease activity and contains a mutation of one of the conserved catalytic residues necessary for endoU activity[185,192,223,224]. We compared binding of wt and mutant Δs2m RNA, in addition to full-length wt s2m. The latter was chosen as a control, as it is presently unknown what the minimum length requirement is for RNA recognition and binding by nsp15. We observed binding of nsp15 (K290A) to wt Δs2m (K_D=1.5μM). When we compared binding of K290A to each mutant Δs2m RNA, we observed only slight decreases in binding affinity (approximately 1.3-fold; $\Delta s2m_{mut1}$ K_D=2.0μM, $\Delta s2m_{mut2}$ K_D=1.9μM, $\Delta s2m_{mut3}$ K_D=8.5μM), which were not found to be statistically significant. Notably, we observed a modest increase in binding affinity of K290A for the full-length s2m RNA (approximately 4-fold; s2m K_D=0.3μM) suggesting that other RNA-protein contacts in longer RNAs may modestly contribute to binding affinity. As nsp15 exhibited similar binding affinities for all Δs2m RNAs we concluded that differences in cleavage efficiency of Δs2m RNAs is not due to differences in protein-RNA binding but likely driven by positioning of the scissile uridine in these differently structured RNAs.

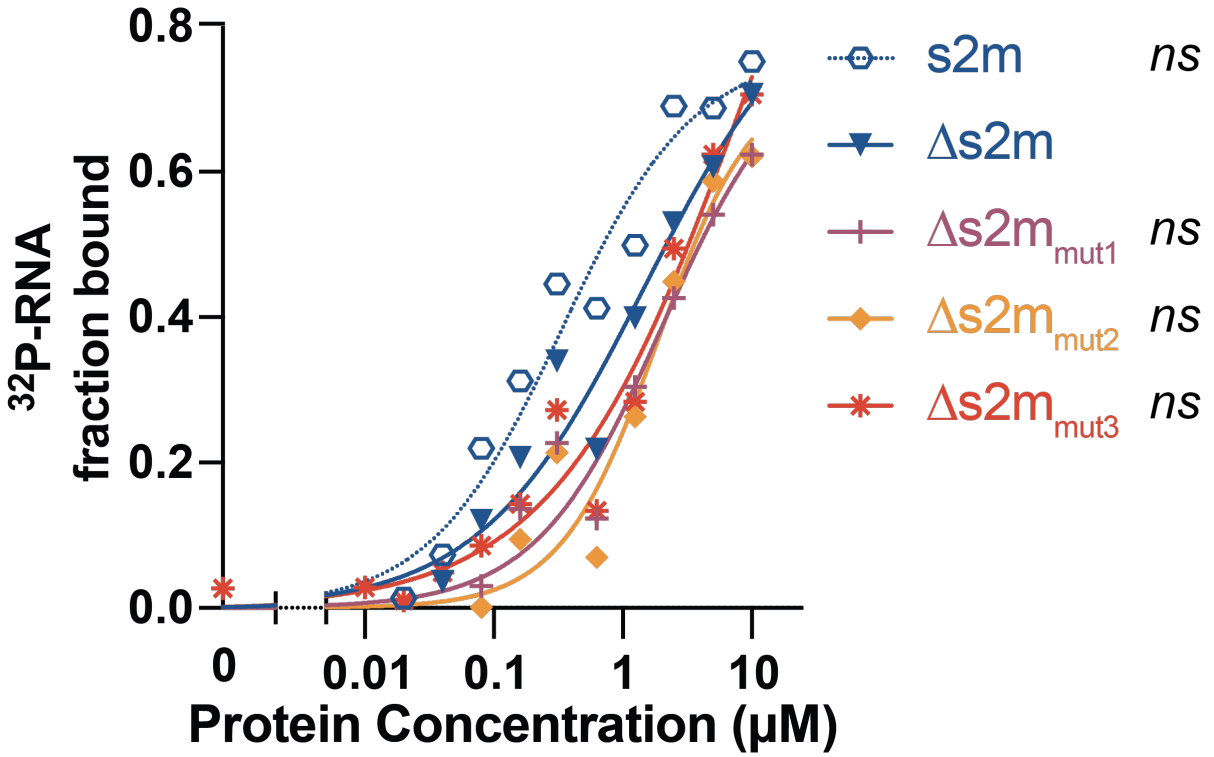


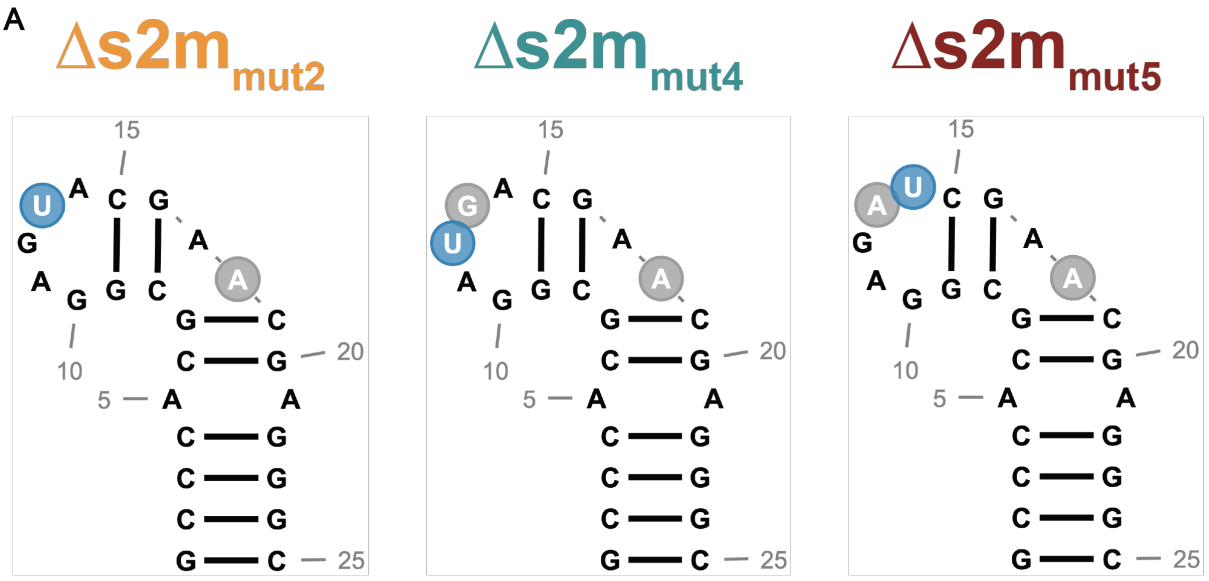
Figure 2.10 SARS-CoV-2 nsp15 binds to Δs2m RNAs with similar affinity. Nsp15 binding curves of Δs2m RNAs as compared to wt s2m. Δs2m RNA binding curves overlap with each other while wt s2m curve is shifted left indicating improved binding with nsp15. Dissociation constant (K_D) was calculated for each RNA and one-way ANOVA with multiple comparisons was performed on the K_D. AUC of wt full-length s2m and Δs2m mutant RNAs are compared to wt Δs2m. Data represents three independent experiments.

Flexible uridine nucleotides in structured pentaloops are susceptible to SARS-CoV-2 nsp15 cleavage

Recent structural studies have suggested that SARS-CoV-2 nsp15 uses a base flipping mechanism to position the scissile uridine in the catalytic pocket, thus enabling cleavage of dsRNA[191]. The crystal structure of s2m shows that U25 in the GNRNA pentaloop adopts a similar flipped out orientation[194] suggesting that this mechanism could also account for the specificity of nsp15 cleavage for U25 in s2m. RNA pentaloops adopt

a characteristic structure which consists of a sheared G-A base pair that closes the pentaloop and induces base stacking of first N, R, and A nucleotides. Notably, this conformation induces extrusion (or 'flipping') of the second N (GNR**N**A) in the pentaloop[219,220]. Thus, we postulated that the position of the scissile uridine in the pentaloop (GAG**U**A) might be necessary to confer the correct orientation required for nsp15 cleavage. To test whether the relative position of U25 is necessary for cleavage specificity, we constructed two additional Δs2m mutants in which the position of the scissile uridine has been altered (Fig 2.11) and compared cleavage efficiency in the endoU assay (Fig 2.12). In $\Delta s2m_{mut4}$ the adjacent guanidine and uridine in the pentaloop were swapped (GA**UG**A), such that the scissile uridine now sits at the apex of the pentaloop. In $\Delta s2m_{mut5}$ the scissile uridine was positioned downstream at the 3' end of the pentaloop and adjacent to the closing GC base pair of the helix (GAGA**U**). The second uridine in the bulge of Δs2m (U18) was also mutated in both RNAs to assess the impact of U13 (U25 equivalent in full-length s2m) cleavage only.

A



B

RNA	ΔG of Minimum Free Energy (kcal/mol)	Frequency of MFE	Ensemble Diversity
$\Delta s2m_{mut2}$	-7.9	73.7%	1.14
$\Delta s2m_{mut4}$	-7.9	73.7%	1.13
$\Delta s2m_{mut5}$	-6.9	36.5%	2.32

Figure 2.11 Thermodynamic stability of Δs2m uridine position mutant RNAs. A. Δs2m mutant RNAs with unpaired uridines are highlighted in blue. All RNA structures were predicted in RNAfold and designed in RNA2Drawer. **B.** Thermodynamic stability and ensemble diversity of Δs2m mutant RNAs indicated as predicted by RNAfold.

We predicted that in $\Delta s2m_{mut4}$ positioning of the scissile uridine at the apex would have minimal effect on cleavage but that placement of uridine adjacent to the helix would place structural constraints on this base such that it would be unable to adopt the necessary conformation (flipped out) required for positioning within the catalytic pocket of nsp15. When we compared cleavage of each mutant, we observed that $\Delta s2m_{mut4}$ (GA**U**GA) was cleaved to completion earlier than both $\Delta s2m_{mut2}$ (control) and $\Delta s2m_{mut5}$ (GAGA**U**) (Fig 2.12).

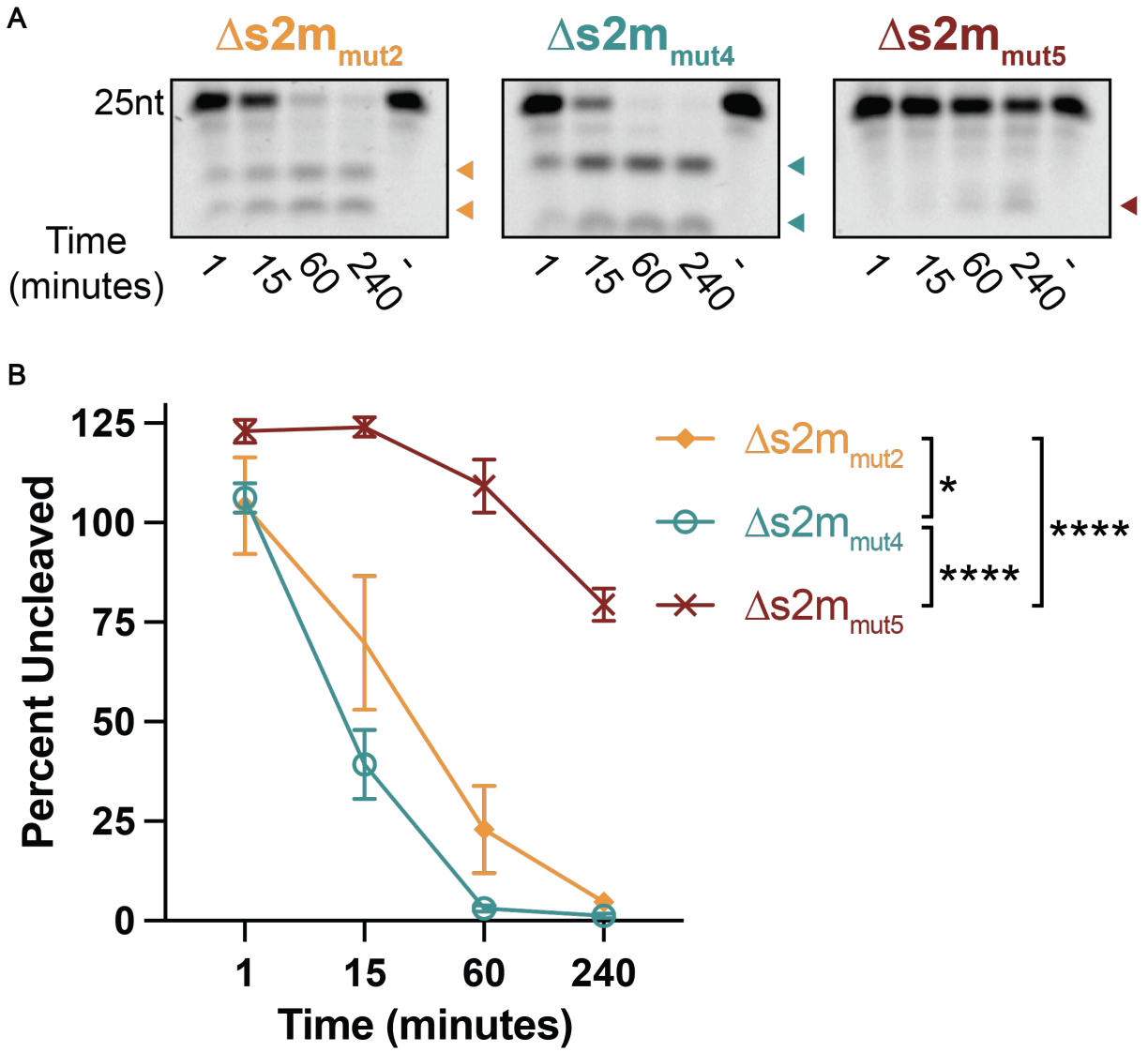
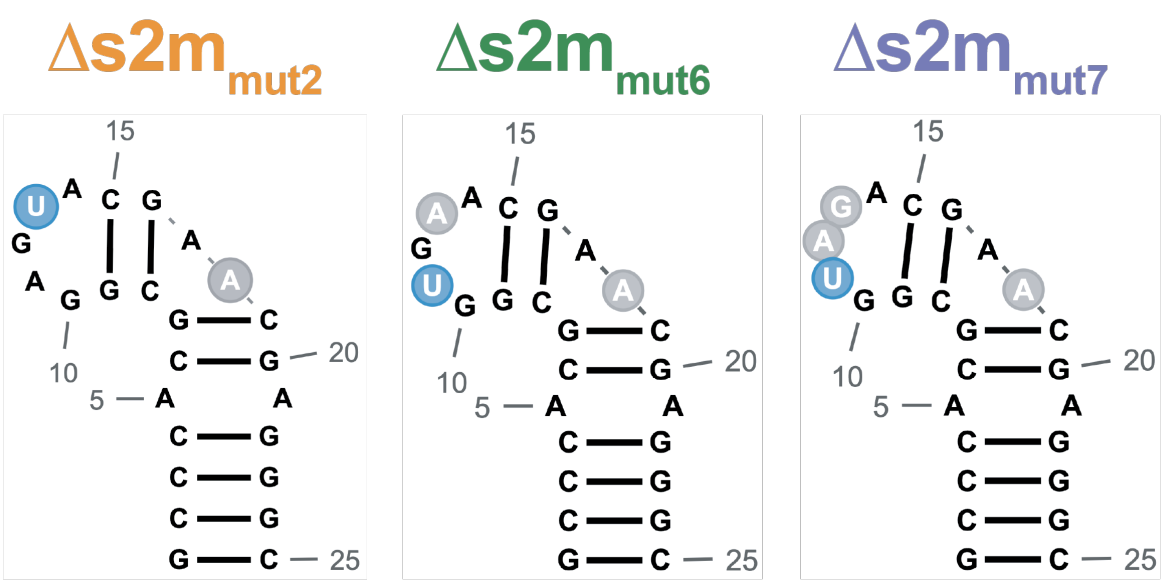


Figure 2.12 Uracil position in loop structures impacts SARS-CoV-2 nsp15 cleavage efficiency. A. Endonuclease assays of Δs2m pentaloop mutant RNAs. Full-length bands indicated at the top of the gel with cleavage products below. Representative images from one experiment. **B.** Percent of full-length RNA remaining as measured by densitometry. Percentage of uncut RNA was calculated by normalizing to a denatured nsp15 control (indicated by (-) in the far-right lane). RNAs with uridines positioned closer to the apex of the pentaloop cleaved more rapidly than $\Delta s2m_{mut5}$ in which the uridine is at the end of the pentaloop. AUC was calculated for each RNA and one-way ANOVA with multiple comparisons was performed on the AUC. Data represents three independent experiments.

We also designed $\Delta s2m_{mut6}$ and $\Delta s2m_{mut7}$ with a uridine in the second position to assess cleavage of these RNAs by nsp15 (Fig 2.13). We observed a slight, but consistent increase in nsp15 cleavage activity in $\Delta s2m_{mut6}$ and $\Delta s2m_{mut7}$ relative to $\Delta s2m_{mut2}$ (Fig 2.14). In contrast, we observed almost no cleavage of $\Delta s2m_{mut5}$ supporting the hypothesis that the scissile uridine might be more structurally constrained at this position and unable to engage with the catalytic pocket of nsp15. From these data, we concluded that uridine position within the pentaloop influences nsp15 cleavage activity, likely through structural constraints.

A



B

RNA	ΔG of Minimum Free Energy (kcal/mol)	Frequency of MFE	Ensemble Diversity
Δs2m$_{mut2}$	-7.9	73.7%	1.14
Δs2m$_{mut6}$	-7.9	72.0%	1.34
Δs2m$_{mut7}$	-7.9	73.4%	1.17

Figure 2.13 Thermodynamic stability of further Δs2m uridine position mutant RNAs. A. Δs2m mutant RNAs with unpaired uridines are highlighted in blue. All RNA structures were predicted in RNAfold and designed in RNA2Drawer. **B.** Thermodynamic stability and ensemble diversity of Δs2m mutant RNAs indicated as predicted by RNAfold.

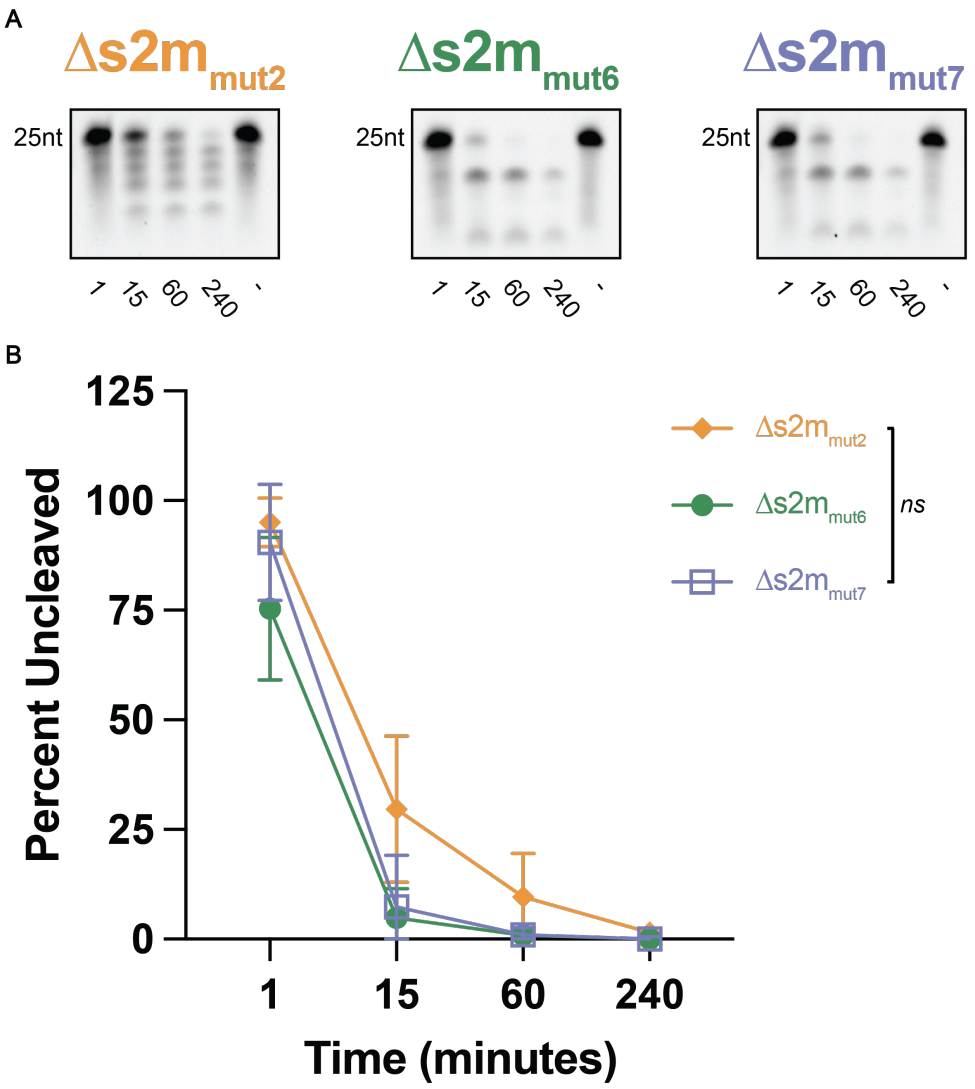


Figure 2.14 Uridine positioning in the s2m pentaloop impacts nsp15 cleavage. A. Endonuclease assays of additional Δs2m mutant RNAs. Reactions were carried out at a nsp15 hexamer:RNA ratio of 1:6. RNAs were run on a 22.5% TBE-Urea Gel. **B.** Thermodynamic stability, frequency of MFE, and ensemble diversity of Δs2m mutant RNAs indicated as predicted by RNAfold. **C.** Percent of full-length RNA remaining as measured by densitometry. Percentage of uncut RNA was calculated by normalizing to a denatured nsp15 control (indicated by (-) in the far-right lane). RNAs with uridines in the second position of the pentaloop ($\Delta s2m_{mut6/7}$) were consistently cleaved more rapidly than $\Delta s2m_{mut2}$ in which the uridine is in the fourth position. P-values were calculated using one-way ANOVA with multiple comparisons analysis of the AUC. Data represents three independent experiments.

Discussion

Nsp15 is a viral endonuclease conserved across the *Nidovirales* and plays a crucial role in CoV evasion of host innate immunity. The crystal structures of several CoV nsp15 proteins have been solved to date, including that of SARS-CoV-2[185,186,190,202,203,223,225]. Structural and biochemical analyses have revealed key aspects of nsp15 function, including the requirement of Mn^{2+} for endonuclease activity at neutral pH as well as formation of higher order oligomers (hexamers) which is also necessary for endoU function[184,190,192,223-225]. Nsp15 primarily functions to evade innate immune responses by cleaving dsRNA replicative intermediates and preventing activation of MDA-5 and possibly other RNA sensors in the host cell[177,178,188,207,208,226-229]. Some studies also suggest that CoV nsp15 may play a role in viral transcriptional regulation during viral replication[188] and that endoU-independent activities additionally contribute to modulation of host responses[200,228,230-232].

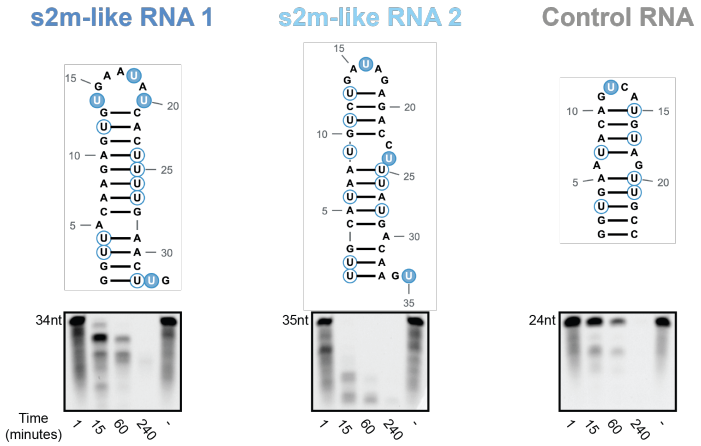
At present, the RNA targets of nsp15 endoU from across the *Coronaviridae* are poorly defined. This can be attributed in part to the fact that *in vitro* studies have used different RNA species (dsRNA vs ssRNA) and different RNA sequences and lengths, making cross-comparison of these data challenging. Nsp15 has been shown to cleave 3' of pyrimidines, in particular uridine, and exhibit a preference for weak bases (A or U) 3' of the scissile nucleotide. However, to date no specific nsp15 recognition motifs have been described. Indeed, our analysis of previously identified MHV nsp15 cleavage sites[188] failed to identify novel sequence motifs in and around the cleavage site. In the absence of a sequence motif, we postulate that RNA secondary structure may function to regulate both RNA binding and cleavage efficiency of nsp15. To this end, recent work has shown that SARS-CoV-2 nsp15 manipulates the scissile uridine for cleavage of dsRNA[191]. However, the extent to which RNA secondary structure is involved in modulating SARS-CoV-2 nsp15 cleavage activity remains to be determined. Thus, in this study we sought to determine the role of RNA structure in the regulation of SARS-CoV-2 nsp15-mediated cleavage and define specific structural features of RNAs that contribute to this. We demonstrated that increased thermodynamic stability of RNA is associated with decreased nsp15 cleavage efficiency (Fig 2.3). Notably, these experiments were done with three RNAs representative of high, moderate, and low RNA secondary structure.

However, examining nsp15 cleavage of other structured RNAs would corroborate these results. Based on our preliminary results, we postulated that stable secondary structure may contribute to nsp15 cleavage by two modes that are not necessarily mutually exclusive: 1) steric hinderance and 2) inducing optimal positioning of the scissile uridine in the nsp15-RNA complex. To test this second hypothesis, we used s2m as a model RNA to explore the structural determinants of nsp15 cleavage specificity.

Using an *in vitro* biochemical approach, we first demonstrated that nsp15 exhibits specificity for cleavage of the uridine nucleotide located in the GNRNA pentaloop of s2m (Fig 2.6). This contrasts with previous data from Bhardwaj et al., in which the authors observed preferential cleavage of U30 located in the bulge sequence downstream of the pentaloop[192]. It should be noted that these previous studies were performed with SARS-CoV-1 nsp15, which shares 88% identity with SARS-CoV-2 nsp15. One explanation for the discrepancy in our data could be structural differences between these proteins which lead to altered cleavage specificity. We began these studies with SARS-CoV-2 nsp15 as we had it on hand, but our aim in the future is to run these endonuclease assays with SARS-CoV-1 nsp15.

We also tested s2m-like RNAs that shared a similar loop size and sequence to wt s2m and found that the s2m-like RNAs were cleaved far more rapidly that a control RNA (Fig 2.15). However, we cannot rule out that other variables may be playing a role in this difference including the positioning and the number of uridines within the loops of these RNAs. Furthermore, the control RNA contains a tetraloop as opposed to pentaloop or heptaloop. As we have not thoroughly tested the influence of loop size on nsp15 cleavage, it is possible that nsp15 cleaves tetraloops with a lower efficiency than pentaloops.

A



B

RNA	ΔG of Minimum Free Energy (kcal/mol)	Frequency of MFE	Ensemble Diversity
s2m-like RNA 1	-10.3	69.8%	0.65
s2m-like RNA 2	-8.6	63.4%	0.97
Control RNA	-6.1	67.8%	0.80

C

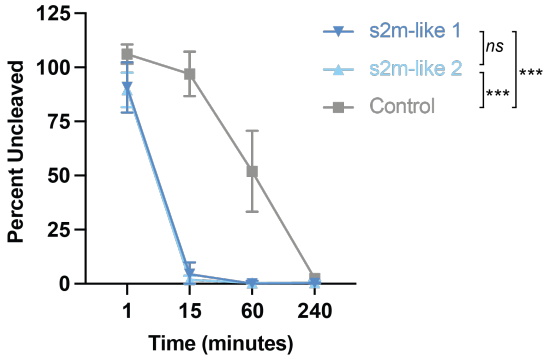


Figure 2.15 RNAs with similar characteristics to s2m are rapidly cleaved by nsp15. **A.** Minimum free energy structures of s2m-like and control RNAs. Representative images from endonuclease assays are indicated below each RNA. Reactions were carried out at a nsp15 hexamer:RNA ratio of 1:120. RNAs were run on a 22.5% TBE-Urea Gel. **B.** Thermodynamic stability and ensemble diversity of s2m-like RNAs indicated as predicted by RNAfold. **C.** Percent of full-length RNA remaining as measured by densitometry. Percentage of uncut RNA was calculated by normalizing to a denatured nsp15 control (indicated by (-) in the far-right lane). AUC was calculated for each RNA, and one-way ANOVA with multiple comparisons was performed on the AUC. Data represents three independent experiments.

Notably, the specificity for cleavage of the pentaloop uridine (GAG**U**A) was not due to differences in nsp15-RNA binding (Fig 2.10), thus we postulated that the sequence context of the scissile uridine might be important for this specificity. The characteristic structure adopted by GNRNA pentaloops induces extrusion, or base flipping, of the scissile uridine away from the body of the RNA[219,220], and this base flipping mechanism has been suggested to be important for correct positioning of uridine in the catalytic pocket of nsp15[191]. Pentaloop motifs have also been implicated in base editing mediated by Adenosine Deaminases that Act on RNA 2 (ADAR2)[233], though in this instance the GCUMA pentaloop is located downstream of the edited A and is implicated in recruitment of ADAR2 via interactions between the dsRNA binding motif of ADAR2 and the target RNA. We speculated that if structural conformation of the scissile uridine was important for base recognition and cleavage, then altering the position (and conformation) of the uridine base would lead to altered cleavage efficiency. Indeed, repositioning of the scissile uridine to the apex of the pentaloop ($\Delta s2m_{mut4}$) slightly enhanced cleavage efficiency, whereas repositioning adjacent to the closing G-C base pair ($\Delta s2m_{mut5}$) significantly reduced RNA cleavage (Fig 2.12). Interestingly, based on the anticipated cleavage sites of $\Delta s2m_{mut2}$ and $\Delta s2m_{mut4}$, we expected to see a single nucleotide difference in the bands between these RNAs, However, there appears to be a larger size discrepancy between these products (Fig 2.16). This may represent more significant differences in cleavage between the RNAs potentially at noncanonical sites. Further studies are needed to confirm the site of cleavage in $\Delta s2m_{mut4}$. The marked decrease in $\Delta s2m_{mut5}$ cleavage is likely due to altered positioning of the scissile uridine which we predict is no longer extruded from the pentaloop, as well as decreased flexibility of this base which is now structurally constrained by the neighboring G-C base pair. However, RNAfold predictions indicate that there is potential wobble base pairing between the G-U nucleotides in the pentaloop of $\Delta s2m_{mut5}$. As a result, it is possible that occlusion of the scissile uridine may be contributing to the lack of cleavage of $\Delta s2m_{mut5}$ (Fig 2.12).

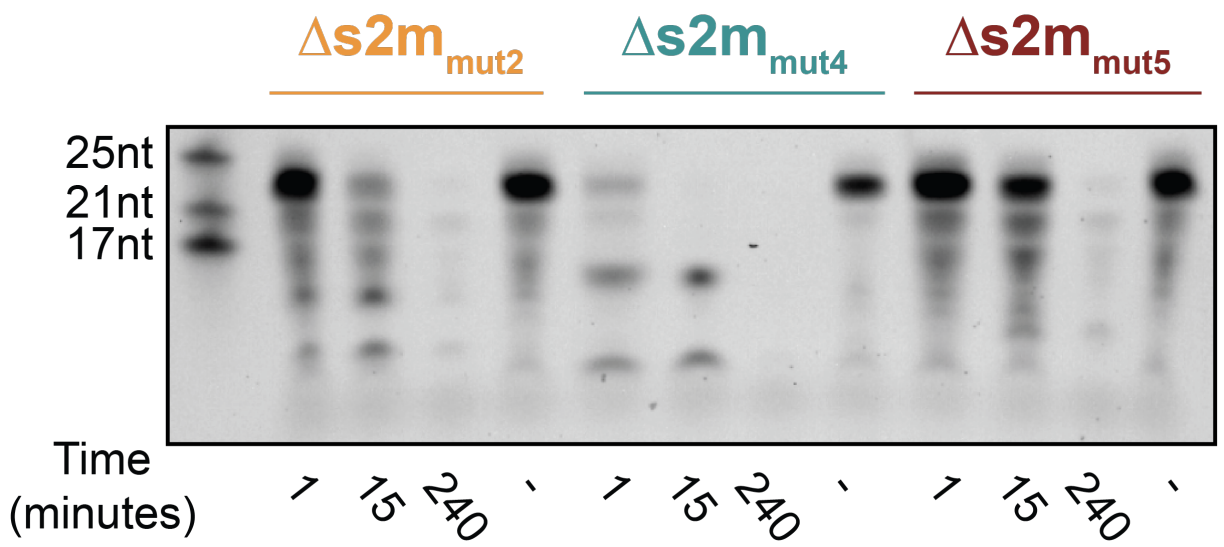


Figure 2.16 $\Delta s2m_{mut2}$ and $\Delta s2m_{mut4}$ show distinct cleavage patterns despite having near identical secondary structure. Timecourse of $\Delta s2m_{mut2}$, $\Delta s2m_{mut4}$, $\Delta s2m_{mut5}$, RNAs (from Fig 3) run on a single gel. Reactions were carried out at a nsp15 hexamer:RNA ratio of 1:6. RNAs were run on a 22.5% TBE-Urea Gel. Cleavage products run side by side show distinct banding patterns despite the presumed cleavage sites in each RNA being only one nucleotide apart indicating potential alternative cleavage sites.

To further address the importance of uridine position within the s2m pentaloop, we designed $\Delta s2m_{mut6}$ and $\Delta s2m_{mut7}$ RNAs which contained the uridine in the second position of the pentaloop. These RNAs were cleaved slightly but consistently more rapidly than $\Delta s2m_{mut2}$ (Fig 2.14). This provides further credence to the hypothesis that uridine position within pentaloops impacts the ability of nsp15 to cleave RNA though further mutants can be tested to corroborate these results. Previous work has shown that trinucleotide sequence context is important for nsp15 activity and that G-U$^{\downarrow}$A trinucleotides were cleaved most efficiently[187]. $\Delta s2m_{mut6}$ (G-U$^{\downarrow}$G) and $\Delta s2m_{mut7}$ (G-U$^{\downarrow}$A) were cleaved at similar rates (Fig 2.14) whereas $\Delta s2m_{mut7}$ and $\Delta s2m_{mut2}$ both share the same sequence context (G-U$^{\downarrow}$A) but were cleaved at slightly different rates. This may point to a role not only for trinucleotide sequence in nsp15 cleavage, but specifically trinucleotide

sequences in the context of RNA structure leading to differential enzymatic activity. Future combinatorial studies looking at both sequence and structural context will further elucidate the substrate specificity of nsp15.

Another possibility is that cleavage efficiency is impacted by altering the sequence context of the scissile uridine. Nsp15 has been shown to preferentially cleave uridines flanked by a weak base (A/U) which is found in wt Δs2m and $\Delta s2m_{mut2}$ ($U^{\downarrow}A$) but is changed to a strong base in $\Delta s2m_{mut5}$ ($U^{\downarrow}C$). However, we observed a modest increase in cleavage efficiency of $\Delta s2m_{mut4}$ which also has a strong base at the 3' position ($U^{\downarrow}G$) and lends support to the hypothesis that the structural context of the scissile uridine, not sequence context, may be more relevant in dictating cleavage specificity and efficiency.

We have listed the MFE of all the RNAs used in our study. While the MFE is the most stable conformation of the RNA, notably, RNAs do not always adopt the MFE structure. Regardless of whether a particular RNA forms the MFE as its true structure, RNAs are dynamic molecules that shift between multiple conformations. To ameliorate this issue, we compared the 100 nt RNAs from SARS-CoV-2 to established SHAPE-MaP data and found that the most thermodynamically stable RNA highly correlated with published structural data (Fig 2.2)[216]. For the same reason, we chose to use s2m as a model RNA for our studies because it has an established crystal structure[194]. Nonetheless, it is important to note that these substrates are likely undergoing conformational changes throughout the course of the reaction. Despite this, we show large differences in cleavage of RNAs that have even single nucleotide changes (Fig 2.6 and 2.9) thus indicating that nsp15 preferentially acts on certain substrates.

Furthermore, in our endonuclease assays, we used several RNAs with loop structures. While loops are commonly thought to be regions lacking structure, unpaired nucleotides in loops can form tertiary structures through non-Watson-Crick base pairing interactions with each other[234]. While RNA prediction software such as RNAfold cannot model these tertiary structures, prediction software is accurate when a large percentage of the queried RNA contain traditional Watson-Crick base pairing[234]. Because the structure of s2m has been verified by x-ray crystallography[194], and ∆s2m RNAs are based off the full-length s2m and have a high percentage of base pairing in the stem, we are confident in the predicted structure of these RNAs.

Collectively, our data suggests that RNA secondary structure plays two distinct roles in determining nsp15 cleavage specificity. First, thermodynamically stable structures likely sterically hinder nsp15-RNA interactions, thus preventing engagement and cleavage of some RNA sequences. Second, susceptible bases are presented in a structural context which facilitates optimal positioning of the scissile nucleotide in the catalytic pocket of nsp15 (e.g., extrusion of the scissile U25 in the GNRNA pentaloop). However, these two mechanisms are not necessarily mutually exclusive. In fact, it is likely that there are multiple factors that play into nsp15 cleavage specificity. However, several questions remain unanswered, most notably, what are the biologically relevant targets (host and viral) of nsp15 cleavage during CoV infection and are viral targets conserved across different CoVs? Two studies to date have identified cleavage sites within the positive-sense genomic RNA and 3' polyU tails of the negative-sense RNA of MHV[177,188]. Despite this, no recognition motif or other regulatory element associated with nsp15 cleavage has been identified. While the *in vitro* assays performed here and in other studies provide a tractable system for dissecting the molecular mechanism of nsp15-RNA interactions and cleavage determinants, a recent computational model suggests that nsp15 is part of a larger replication-transcription complex (RTC) which contains other viral proteins and host factors. How nsp15 protein-protein interactions and RNA interactions with other replicase proteins (e.g., nsp12, nsp14, nsp16) impacts nsp15 cleavage of RNA remains unsolved. Indeed, recent models of the RTC suggest the hexameric structure of nsp15 provides the central structure and arrangement of a hexameric RTC[120]. In this model where nsp15 forms part of a multimeric RTC, it is possible there may be differences in its cleavage of viral RNAs due to its 3-D interaction with the nascent RNA as opposed to when nsp15 is free from the RTC as was the case in this study. The importance of examining nsp15-RNA interactions in their relevant biological context is further highlighted by recent studies showing that the requirement for Mn^{2+} and the structural conformation of nsp15 is highly dependent on the pH of the environment[235]. Given that the redox state of the viral RTC of +ssRNA viruses is altered during replication, the conditions under which nsp15 endoU activity is examined is highly relevant. To address this our future studies will explore the impact of pH and redox state on nsp15 structure and function *in vitro*.

Overall, our findings suggest a role for RNA secondary structure in regulation of nsp15 cleavage. Identification of bona fide nsp15 cleavage sites in SARS-CoV-2 RNA during replication and examination of how perturbation of these sites impacts the outcome of replication and pathogenesis will provide key insights into the broader roles for nsp15 endoU activity during infection (e.g., viral transcriptional/translational regulation, host translation). Furthermore, defining key structural features of RNAs which are resistant or susceptible to nsp15 cleavage may provide avenues for development of antiviral therapeutics.

Chapter III: Future Directions and Outstanding Questions

Summary of work

Our work described in this thesis shows that RNA secondary structure impacts nsp15 activity. We first generated multiple RNA segments from the SARS-CoV-2 genome with varying thermodynamic stability and found that the RNA with the highest thermodynamic stability (RNA 1) was cleaved slower than two RNAs with lower thermodynamic stability (RNA 2 and 3). This is presumably due to an increased availability of unpaired uridines in RNAs 2 and 3 as they have a greater number of predicted structures that they can adopt relative to RNA 1. Consequently, they likely transition between these predicted structures during which previously base-paired uridines become unpaired thereby presenting accessible uridines to nsp15 for cleavage. Conversely, RNA 1 has low structural diversity and is not as dynamic in changing structures. As a result, the already paired uridines are obscured from nsp15. These results suggest that RNAs with higher thermodynamic stability are protected from nsp15 cleavage relative to RNAs with lower order structure.

We then turned to a previously identified RNA structure from SARS-CoV termed s2m. The s2m RNA is found in the 3'UTR of SARS-CoV genome and has been confirmed to be a stem loop structure by x-ray crystallography[194]. S2m has two unpaired uridines: one in the pentaloop of the RNA and one in a bulge below the pentaloop. Because of its established secondary structure, we chose to use this RNA as a model for our structural studies. Previous work with SARS-CoV nsp15 indicated that the enzyme cleaved s2m at the unpaired uridine in the bulge of the RNA. To validate these data with SARS-CoV-2 nsp15, we designed s2m mutant RNAs in which the unpaired uridines were mutated to adenosines that would not be able to be cleaved by nsp15. Our research revealed that SARS-CoV-2 nsp15 exhibits a distinct preference for cleaving s2m within the pentaloop of the RNA contrary to earlier findings with SARS-CoV nsp15. We further investigated these results using modified s2m RNAs that maintained the pentaloop and the bulge while eliminating the other uridines within the stem of WT s2m. Using the same set of mutants in the truncated s2m RNAs, we corroborated that SARS-CoV-2 nsp15 preferentially cleaves the unpaired uridine in the pentaloop of the RNA over the uridine in the bulge.

Finally, we sought to determine whether uridine positioning in the s2m pentaloop impacted nsp15 cleavage efficiency. Using the ∆s2m mut2 mutants as a baseline since it contains uridine in the fourth and wild-type position, we designed further ∆s2m mutants with the uridine placed in the third and fifth positions while maintaining the stem loop feature. We found that uridine in the third and fourth positions of the pentaloop were cleaved far more rapidly than when uridine was in the fifth position of the s2m pentaloop. From these data, we conclude that positioning of the scissile uridine impacts cleavage, likely due to structural constraints.

Outstanding Questions

What is the role of Mn^{2+} in nsp15 catalysis?

Nsp15 has been shown to have optimal cleavage activity in the presence of Mn^{2+}[182]. However, the exact role that nsp15 plays during nsp15 catalysis has not yet been characterized. Notably, none of the resolved nsp15 structures have identified the position of Mn2+ ions. One hypothesis is that Mn^{2+} stabilizes the nsp15 hexamer in an orientation that facilitates RNA cleavage[236]. Critically, the existing studies investigating the dependence of nsp15 on Mn^{2+} has been carried out *in vitro*, employing concentrations of Mn^{2+} that exceed physiological levels by 50-500 fold[183]. As such, further investigations are necessary to determine the role of Mn^{2+} during catalysis, particularly during infection.

Does nsp15 play different roles at acidic and neutral pH?

The experiments performed in this thesis were carried out at a neutral pH of 7.5. However, a recent study showed that nsp15 adopts different conformations at pH 6.0 and pH 7.5[237]. At pH 6.0, nsp15 is in a more open state relative to nsp15 at pH 7.5. Specifically, the internal barrel of the nsp15 hexamer is contracted at neutral pH relative to acidic pH[237]. This suggests that nsp15 may play a different role during infection at the two pH conditions potentially present in different areas across the cell.

The pH of the DMVs in which CoVs replicate has not yet been studied. We hypothesize that these replication compartments are acidic and that nsp15 adopts its open conformation in DMVs. A computational model has shown that the SARS-CoV-2 RTC is centered around the nsp15 hexamer[120]. While this model has not yet been

experimentally validated, this scaffolding function represents a possible role of nsp15 outside of its endonuclease activity. Conversely, the pH of the cytoplasm is known to be more neutral at 7.5. In this state, we propose that nsp15 functions as a host immune antagonist through its endonuclease activity.

What are the viral RNA targets of nsp15 during infection?

The work in this thesis sheds light on the structural characteristics of RNAs that nsp15 cleaves. However, very little work has been done to identify the RNAs that nsp15 cleaves *in vivo*, particularly in the context of highly pathogenic human CoVs. Two recent studies have shown that MHV nsp15 cleaves viral RNA[177,188]. Hackbart et al. showed that nsp15 primarily cleaves polyU RNA at the 5' end of the negative-sense RNA genome, whereas Ancar et al. found that nsp15 cleaves viral RNA at various locations across the MHV genome[177,188]. As these studies were performed with MHV, it remains to be seen whether nsp15 cleaves similar sites in the SARS-CoV-2 genome or if nsp15 targets a different set of loci in highly pathogenic human CoVs.

Furthermore, cellular targets of nsp15 during infection have not yet been explored. Gaglia et al. developed a method to map cellular mRNAs that were cleaved by the KSHV ribonuclease, SOX[238]. This protocol can be adapted to identify cleavage sites of nsp15 in cellular RNAs. Briefly, total RNA would be harvested from cells infected with SARS-CoV-2 and the polyA RNAs would be enriched. These RNAs are then ligated to a 5' adapter that contains an MmeI restriction site. Following this, cDNA is generated from these RNAs and the RNAs are digested with MmeI. The digested RNAs are ligated to a 3' RNA adapter, and the samples are prepared for high-throughput sequencing. These data will provide further insight into the role of nsp15 as an innate immune antagonist.

Two recent studies have analyzed the targets of MHV nsp15 during infection[177,188]. Hackbart et al. show that MHV nsp15 cleave polyU RNA found at the 5' end of the negative sense RNA genome. The authors did this by immunoprecipitating double stranded RNA (dsRNA), sequencing the samples by RNA-seq and mapping the reads to the MHV genome. When comparing the nsp15 catalytic mutant to WT MHV, the mutant samples were enriched for dsRNA. Further PCR analysis indicated that nsp15 cleaves polyU RNA that accumulates at the 5' end of the RNA genome. A caveat with this study

is that despite dsRNA accumulating across the viral genome in the nsp15 mutant, the authors did not investigate other RNAs that nsp15 might be cleaving.

In a second study from 2020 looking at targets of MHV nsp15, Ancar et al. found specific sites of cleavage in the across the positive sense viral genome. In addition to canonical cleavage sites after uridines, the authors found that nsp15 cleaves following cytosines as well. This finding was corroborated by Frazier et al. in 2021[187]. Further analysis of these cleavage sites showed that nsp15 cleaved between the specific dinucleotides UA, CA, CC, and AC. Interestingly, this study did not identify nsp15 cleavage sites in the negative sense viral RNA. It is unclear why these two studies found different targets of nsp15 cleavage in MHV. It's possible that the differences in these results may be due to the separate sequencing techniques used to identify cleavage sites. Further studies of nsp15 cleavage sites in other CoVs may provide insight into which of these sites are important for viral replication as well as if there are differences in nsp15 cleavage across CoVs.

Does nsp15 interact with other viral proteins in the replication-transcription complex?

A report from 2017 showed that nsp15 associates with the RTC proteins nsp8 and nsp12 in an RNA-independent manner[239]. However, no structures of the RTC have been identified with nsp15 suggesting that the interaction between nsp15 and the RTC may only be transient. However, as previously stated, a recent study generated a computational model which centers the RTC around the nsp15 hexamer[120]. While these results need to be corroborated experimentally, this represents a potential scaffolding function for nsp15 during infection. Furthermore, nsp15 interactions with other viral proteins may also contribute to regulating nps15 cleavage during infection.

Can small molecule inhibitors of nsp15 be used as therapeutics against CoV infection?

As pathogenic CoVs have emerged over the past twenty years and with SARS-CoV-2 becoming endemic in the population, therapeutics are needed to combat current infections and control potential future outbreaks. To date, therapeutics against CoVs have

focused on inhibiting some of the essential viral proteins including the viral RdRp and proteases[97,98,240]; however, other viral proteins have been overlooked. Because it is conserved across CoVs, nsp15 represents an attractive target for therapeutic intervention. Our previous work has identified Exebryl-1 as an inhibitor against nsp15[217]. Unfortunately, Exebryl-1 was only able to inhibit nsp15 at 1-10μM *in vitro* and at 65.6μM during SARS-CoV-2 infection[217]. As such, there is still a need to find nsp15 inhibitors that are functional at low micromolar or nanomolar concentrations. This work is currently ongoing and will be published in a subsequent manuscript.

Concluding remarks

CoVs have long circulated among the population; however, the advent of highly pathogenic CoVs beginning with SARS-CoV in 2002 has spotlighted CoVs in the public eye and accelerated research in the field. As a family of viruses within the *Nidovirales*, CoVs have among the largest genomes of RNA viruses. Many of the nsps encoded by CoVs are well studied with established enzymatic activities. One protein that has been overlooked is nsp15. Though the enzymatic activity of nsp15 as a uridine-specific endonuclease has been characterized, the substrate specificity of nsp15 remains understudied. The work in this thesis shows that RNA secondary structure plays a role in protecting substrates from nsp15 cleavage. Specifically, RNAs with high thermodynamic stability are not cleaved by nsp15 as rapidly as RNAs lacking stability. Additionally, nsp15 cleaves uridines at specific sites within pentaloops. This represents a step toward determining the species of RNA that nsp15 catalyzes during viral infection.

Chapter IV: Materials and Methods

Generation of SARS-CoV-2 RNA segments

100nt segments of SARS-CoV-2 genome were amplified using the primers listed in Table 4.1. Amplicons were purified using Mag-Bind Total Pure NGS beads (Omega Bio-Tek). 1.5x bead volume relative to amplification reactions was used and purified as per manufacturer's instructions. Purified amplicons were subcloned into a vector containing the HDV ribozyme[241]. These constructs were linearized and purified prior to transcription.

Table 4.1 Primers used for generation of SARS-CoV-2 RNA segments

RNA	Forward Primer	Reverse Primer
RNA 1	CATGAGCTTA AGTAATACGA CTCACTATAG GCTACTAACA ATGCCAT	CATGAGGCTA GCAAGATTGT GTCCGCTTAA AA
RNA 2	CATGAGCTTA AGTAATACGA CTCACTATAG ACCTGAGCAT AGTCTTG	CATGAGGCTA GCGAGAACAC ACAGCCTC
RNA 3	CATGAGCTTA AGTAATACGA CTCACTATAG TGAAGTGCTG TCTGAC	CATGAGGCTA GCGTTACACG ATAACCAGTA AAG

Approximately 5 pmol of purified DNA was used as a template for T7 transcription reactions. RNAs were transcribed using the HiScribe T7 High Yield RNA Synthesis Kit (New England Biolabs). Reactions were set up as per manufacturer's instructions with the addition of 4 units of Murine RNase Inhibitor (NEB) and 5% DMSO. Transcriptions were incubated at 42°C overnight after which, the reactions were treated with DNaseI (NEB) as per manufacturer's instructions. DNase-treated transcriptions were then purified using Mag-Bind Total Pure NGS beads as described above.

Transcriptions were gel purified by running samples on urea gels. 15% TBE-Urea gels were made using SequaGel UreaGel 29:1 Denaturing Gel System (National Diagnostics). Urea was added to the samples to a final concentration of 4M before running on the gels at 200V for 90 minutes. Gels were briefly stained with 0.02% Methylene Blue in TBE until ladder was visible. Gels were imaged on Gel Doc XR+ (Bio-Rad), and bands corresponding to 100nt fragments were cut out and shredded by centrifugation. RNAs were eluted overnight on rotator in 3x volume of elution buffer (10mM Tris-HCl pH 7.5,

1mM EDTA, 300mM NaCl, 0.1% SDS) relative to gel weight. Gel slurries were applied to cellulose acetate columns and frozen at -80°C for 10 minutes. Samples were centrifuged at 10,000x g for 3 minutes. Eluate was adjusted up to 600µL with nuclease-free water and purified by acid phenol-chloroform extraction.

All other RNAs (s2m and ∆s2m and corresponding mutant RNAs) were synthesized by Integrated DNA Technologies. RNA sequences can be found in Table 4.2.

Table 4.2 RNAs generated by Integrated DNA Technologies

RNA	RNA sequence
s2m-like RNA 1	GGUUACAAGAGUGUGAAUAUCACUUUUGAACUUG
s2m-like RNA 2	UUGCAUAAUGUCUGAUAGAGACCUUUAUGACAAGU
control RNA	GGUGAAUACAGUCAUGUAGUUGCC
s2m	GAGGUUCAUCGAGGCCACGCGGAGUACGAUCGAGGGUACAGUGAAUUU
$s2m_{mut\ 1}$	GAGGUUCAUCGAGGCCACGCGGAGAACGAUCGAGGGUACAGUGAAUUU
$s2m_{mut\ 2}$	GAGGUUCAUCGAGGCCACGCGGAGUACGAACGAGGGUACAGUGAAUUU
$s2m_{mut\ 3}$	GAGGUUCAUCGAGGCCACGCGGAGAACGAACGAGGGUACAGUGAAUUU
∆s2m	GCCCACGCGGAGUACGAUCGAGGGC
$\Delta s2m_{mut\ 1}$	GCCCACGCGGAGAACGAUCGAGGGC
$\Delta s2m_{mut\ 2}$	GCCCACGCGGAGUACGAACGAGGGC
$\Delta s2m_{mut\ 3}$	GCCCACGCGGAGAACGAACGAGGGC
$\Delta s2m_{mut\ 4}$	GCCCACGCGGAUGACGAACGAGGGC
$\Delta s2m_{mut\ 5}$	GCCCACGCGGAGAUCGAACGAGGGC
$\Delta s2m_{mut\ 6}$	GCCCACGCGGUGAACGAACGAGGGC
$\Delta s2m_{mut\ 7}$	GCCCACGCGGUAGACGAACGAGGGC
$\Delta s2m_{mut\ 8}$	GCCCACGCGCAGAACGAACGAGGGC

Expression of recombinant nsp15

Recombinant nsp15 was expressed and purified as previously described[217]. Briefly, a pET28a(+) construct consisting of a codon-optimized SARS-CoV-2 nsp15 gene sequence with a N-terminal hexahistidine tag was expressed in Rosetta BL21(DE3) *E.*

coli competent cells using autoinduction methods. The K290A mutant was generated using a Q5 site-directed mutagenesis kit (NEB) and the forward and reverse primers: CGGCAGCAGCGCATGCGTGTGC and GTCTGCGCATCGGTAATAAAATAG. Recombinant protein was purified using 5 mL HisTrap FF immobilized metal ion affinity chromatography columns (GE Healthcare) and fractionated using a Superdex 200 size-exclusion chromatography column and an NGC Chromatography System (Bio-Rad). Fractions containing purified nsp15 were flash frozen in liquid nitrogen and stored at -80°C.

Endonuclease Assay

RNAs were diluted in folding buffer (50mM Tris-HCl pH 7.5, 50mM KCl, 5mM $MgCl_2$, 5mM DTT). RNAs were folded by incubating at 95^{o}C for 10 minutes then cooled to 25^{o}C at a rate of 0.1^{o}C/sec (Fig 4.1). Purified nsp15 was diluted in assay buffer (50mM Tris-HCl pH 7.5, 50mM KCl, 5mM $MnCl_2$, 5mM $MgCl_2$, 5mM DTT) to the appropriate concentration for indicated nsp15 hexamer:RNA ratios. To denature nsp15, the protein was incubated at 95^{o}C for 5 minutes. Denatured wt protein showed no cleavage when compared to the nsp15 catalytic mutant protein (Fig 4.2). Folded RNA was aliquoted into individual tubes, and diluted nsp15 was mixed in equal amounts to each RNA simultaneously. Reactions were sampled and arrested at indicated timepoints with a final concentration of 20mM EDTA before flash freezing in liquid nitrogen. Denatured nsp15 reactions were incubated for 240 minutes.

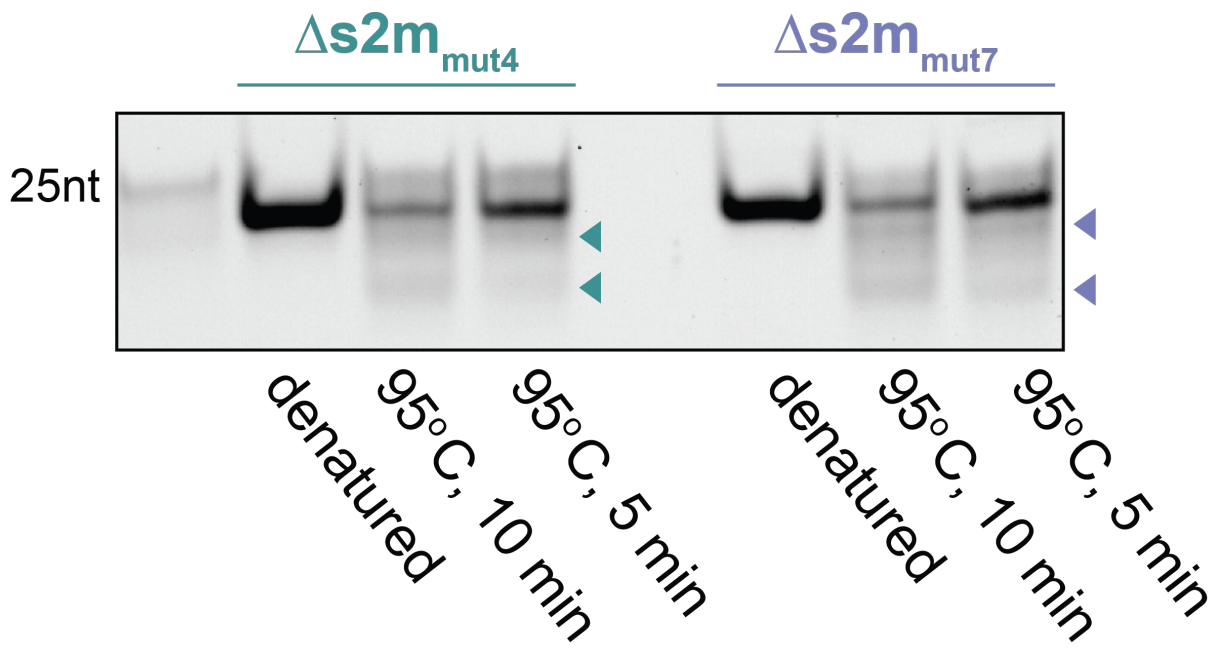


Fig. 4.1 TBE-urea PAGE analysis of folded vs unfolded RNAs. $\Delta s2m_{mut4}$ and $\Delta s2m_{mut7}$ RNAs were denatured and folded as outlined in the materials and methods prior to being electrophoresed on a native gel. Denatured RNA was heated in 2x RNA Loading Dye (NEB) at 70°C for 10 minutes. Folded RNAs show the presence of conformers as indicated by arrows.

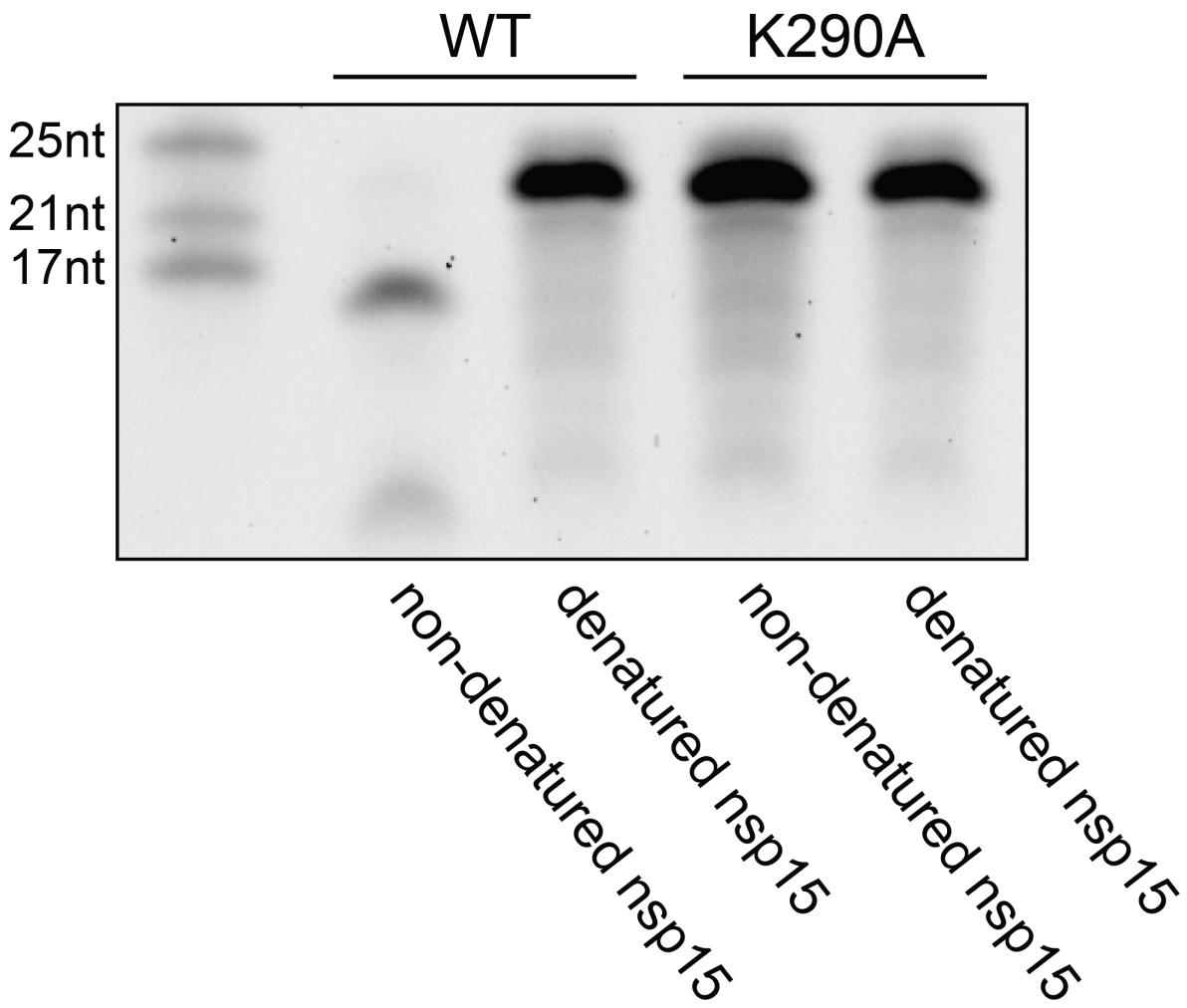


Fig. 4.2 Denatured nsp15 does not cleave RNA substrates. $\Delta s2m_{mut7}$ RNA was incubated with wt nsp15, catalytically inactive nsp15 (K290A), or denatured wt or K290A nsp15. For denatured samples, nsp15 was denatured by heating recombinant protein for 5 minutes at 95°C prior to the addition of folded RNA. Endonuclease reactions were incubated for 15 minutes then terminated with the addition of EDTA to a final concentration of 20mM and flash frozen.

TBE-Urea gels were made using SequaGel UreaGel 29:1 Denaturing Gel System (National Diagnostics). 15% or 22.5% gels were poured as indicated in each experiment. RNA samples were thawed and mixed with an equal volume of 2x RNA Loading Dye

(NEB) and incubated at 70°C for 10 minutes prior to loading onto the gels. Gels were run at 200V for 1 hour in 1x TBE. Gels were stained with SYBR Green II (ThermoFisher) diluted 1:5000 in TBE for 30 minutes while shaking and imaged on Gel Doc XR+ (Bio-Rad). Full-length bands at each time point were calculated by densitometry using Image Lab software (Bio-Rad). Density of bands were normalized to denatured nsp15 control for each reaction.

Differential Radial Capillary Action of Ligand Assay (DRaCALA)

DRaCALAs were performed as described previously[221]. Briefly, RNAs were dephosphorylated and 5' radiolabeled with 30μCi ATP. Radiolabeled RNAs were set up in reactions with the following components: DRaCALA buffer (50mM HEPES pH 7.0, 100mM KCl), 5mM $MgCl_2$, 50ng/μL yeast tRNA, 1mM DTT, 5% glycerol. RNAs were heated to 95°C for 2 minutes and cooled at RT for 10 minutes. Serial dilutions of catalytic mutant nsp15 K290A were added to the reactions which were incubated for 15 minutes at RT and spotted onto a 0.45μm nitrocellulose membrane. Blots were exposed to a phosphor screen overnight and imaged on Sapphire Molecular Imager (Azure Biosystems). Binding was measured by densitometry as previously described[222] using Fiji.

Statistical analysis

For quantification of band density in endonuclease assays, the area under the curve (AUC) was calculated for each replicate using GraphPad Prism software. Statistical analysis of the AUC was performed using one-way ANOVA with multiple comparisons, whereby the mean of each column was compared to the mean of all other columns. *$p<0.05$ (*), $p<0.01$ (**), $p<0.001$ (***), $p<0.0001$(****)* For statistical analysis of DRaCALA, the dissociation constant (K_D) was calculated for each RNA and a one-way ANOVA with multiple comparisons was used to calculate *p-values* for each comparison.

Chapter V: Referenced Literature

1 Manaker, R. A., Piczak, C. V., Miller, A. A. & Stanton, M. F. A hepatitis virus complicating studies with mouse leukemia. *J Natl Cancer Inst* **27**, 29-51 (1961).

2 Cheever, F. S., Daniels, J. B. & et al. A murine virus (JHM) causing disseminated encephalomyelitis with extensive destruction of myelin. *J Exp Med* **90**, 181-210, doi:10.1084/jem.90.3.181 (1949).

3 Bailey, O. T., Pappenheimer, A. M., Cheever, F. S. & Daniels, J. B. A Murine Virus (Jhm) Causing Disseminated Encephalomyelitis with Extensive Destruction of Myelin : Ii. Pathology. *J Exp Med* **90**, 195-212, doi:10.1084/jem.90.3.195 (1949).

4 Doyle, L. P. & Hutchings, L. M. A transmissible gastroenteritis in pigs. *J Am Vet Med Assoc* **108**, 257-259 (1946).

5 Beach, J. R. & Schalm, O. W. A Filterable Virus, Distinct from that of Laryngotracheitis, the Cause of a Respiratory Disease of Chicks1. *Poultry Science* **15**, 199-206, doi:https://doi.org/10.3382/ps.0150199 (1936).

6 Tyrrell, D. A. & Bynoe, M. L. Cultivation of viruses from a high proportion of patients with colds. *Lancet* **1**, 76-77, doi:10.1016/s0140-6736(66)92364-6 (1966).

7 Hamre, D. & Procknow, J. J. A new virus isolated from the human respiratory tract. *Proc Soc Exp Biol Med* **121**, 190-193, doi:10.3181/00379727-121-30734 (1966).

8 Picault, J. P. *et al.* Isolation, characterisation and preliminary cross-protection studies with a new pathogenic avian infectious bronchitis virus (strain PL-84084). *Avian Pathol* **15**, 367-383, doi:10.1080/03079458608436300 (1986).

9 Lavi, E., Gilden, D. H., Highkin, M. K. & Weiss, S. R. The organ tropism of mouse hepatitis virus A59 in mice is dependent on dose and route of inoculation. *Lab Anim Sci* **36**, 130-135 (1986).

10 Davelaar, F. G., Kouwenhoven, B. & Burger, A. G. Occurrence and significance of infectious bronchitis virus variant strains in egg and broiler production in the Netherlands. *Vet Q* **6**, 114-120, doi:10.1080/01652176.1984.9693924 (1984).

11 Guan, Y. *et al.* Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* **302**, 276-278, doi:10.1126/science.1087139 (2003).

12 CDC. *Update: Outbreak of Severe Acute Respiratory Syndrome --- Worldwide, 2003*, <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5212a1.htm#fig1> (2003).

13 Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. & Fouchier, R. A. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* **367**, 1814-1820, doi:10.1056/NEJMoa1211721 (2012).

14 van Boheemen, S. *et al.* Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. *mBio* **3**, doi:10.1128/mBio.00473-12 (2012).

15 Guo, Y. R. *et al.* The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status. *Mil Med Res* **7**, 11, doi:10.1186/s40779-020-00240-0 (2020).

16 Chan, J. F. *et al.* A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* **395**, 514-523, doi:10.1016/S0140-6736(20)30154-9 (2020).

17 Taharaguchi, S., Soma, T. & Hara, M. Prevalence of feline coronavirus antibodies in Japanese domestic cats during the past decade. *J Vet Med Sci* **74**, 1355-1358, doi:10.1292/jvms.11-0577 (2012).

18 Fulton, R. W. *et al.* Bovine coronavirus (BCV) infections in transported commingled beef cattle and sole-source ranch calves. *Can J Vet Res* **75**, 191-199 (2011).

19 Wang, L., Byrum, B. & Zhang, Y. New variant of porcine epidemic diarrhea virus, United States, 2014. *Emerg Infect Dis* **20**, 917-919, doi:10.3201/eid2005.140195 (2014).

20 Wood, E. N. An apparently new syndrome of porcine epidemic diarrhoea. *Vet Rec* **100**, 243-244, doi:10.1136/vr.100.12.243 (1977).

21 Luk, H. K. H., Li, X., Fung, J., Lau, S. K. P. & Woo, P. C. Y. Molecular epidemiology, evolution and phylogeny of SARS coronavirus. *Infect Genet Evol* **71**, 21-30, doi:10.1016/j.meegid.2019.03.001 (2019).

22 Helmy, Y. A. *et al.* The COVID-19 Pandemic: A Comprehensive Review of Taxonomy, Genetics, Epidemiology, Diagnosis, Treatment, and Control. *J Clin Med* **9**, doi:10.3390/jcm9041225 (2020).

23 Ignjatovic, J. & Sapats, S. Avian infectious bronchitis virus. *Rev Sci Tech* **19**, 493-508, doi:10.20506/rst.19.2.1228 (2000).

24 Schalk A, F. An apparently new respiratory disease of baby chicks. *J. Am. Vet. Med. Assoc.* **78**, 413-423 (1931).

25 Bushnell, L. D. & Brandly, C. A. Laryngotracheitis in Chicks*. *Poultry Science* **12**, 55-60, doi:https://doi.org/10.3382/ps.0120055 (1933).

26 ICTV. *Family: Coronaviridae*
Chapter Version: ICTV Ninth Report; 2009 Taxonomy Release, <https://ictv.global/report_9th/RNApos/Nidovirales/Coronaviridae> (2011).

27 Bumstead, N., Huggins, M. B. & Cook, J. K. Genetic differences in susceptibility to a mixture of avian infectious bronchitis virus and Escherichia coli. *Br Poult Sci* **30**, 39-48, doi:10.1080/00071668908417123 (1989).

28 Meulemans, G., Carlier, M. C., Gonze, M., Petit, P. & Vandenbroeck, M. Incidence, characterisation and prophylaxis of nephropathogenic avian infectious bronchitis viruses. *Vet Rec* **120**, 205-206, doi:10.1136/vr.120.9.205 (1987).

29 Cook, J. K., Smith, H. W. & Huggins, M. B. Infectious bronchitis immunity: its study in chickens experimentally infected with mixtures of infectious bronchitis virus and Escherichia coli. *J Gen Virol* **67 (Pt 7)**, 1427-1434, doi:10.1099/0022-1317-67-7-1427 (1986).

30 Karimi, V. *et al.* Efficacy of H120 and Ma5 avian infectious bronchitis vaccines in early challenge against QX strain. *Virusdisease* **29**, 123-126, doi:10.1007/s13337-017-0414-4 (2018).

31 de Wit, J. J., de Jong, M. C., Pijpers, A. & Verheijden, J. H. Transmission of infectious bronchitis virus within vaccinated and unvaccinated groups of chickens. *Avian Pathol* **27**, 464-471, doi:10.1080/03079459808419370 (1998).

32 MacNamara, K. C., Chua, M. M., Phillips, J. J. & Weiss, S. R. Contributions of the viral genetic background and a single amino acid substitution in an immunodominant CD8+ T-cell epitope to murine coronavirus neurovirulence. *J Virol* **79**, 9108-9118, doi:10.1128/JVI.79.14.9108-9118.2005 (2005).

33 Compton, S. R., Ball-Goodrich, L. J., Paturzo, F. X. & Macy, J. D. Transmission of enterotropic mouse hepatitis virus from immunocompetent and immunodeficient mice. *Comp Med* **54**, 29-35 (2004).

34 Compton, S. R. *et al.* Pathogenesis of enterotropic mouse hepatitis virus in immunocompetent and immunodeficient mice. *Comp Med* **54**, 681-689 (2004).

35 Barthold, S. W., Beck, D. S. & Smith, A. L. Enterotropic coronavirus (mouse hepatitis virus) in mice: influence of host age and strain on infection and disease. *Lab Anim Sci* **43**, 276-284 (1993).

36 Bradburne, A. F. & Somerset, B. A. Coronative antibody tires in sera of healthy adults and experimentally infected volunteers. *J Hyg (Lond)* **70**, 235-244, doi:10.1017/s0022172400022294 (1972).

37 Bradburne, A. F., Bynoe, M. L. & Tyrrell, D. A. Effects of a "new" human respiratory virus in volunteers. *Br Med J* **3**, 767-769, doi:10.1136/bmj.3.5568.767 (1967).

38 McIntosh, K. *et al.* Coronavirus infection in acute lower respiratory tract disease of infants. *J Infect Dis* **130**, 502-507, doi:10.1093/infdis/130.5.502 (1974).

39 El-Sahly, H. M., Atmar, R. L., Glezen, W. P. & Greenberg, S. B. Spectrum of clinical illness in hospitalized patients with "common cold" virus infections. *Clin Infect Dis* **31**, 96-100, doi:10.1086/313937 (2000).

40 Falsey, A. R. *et al.* The "common cold" in frail older persons: impact of rhinovirus and coronavirus in a senior daycare center. *J Am Geriatr Soc* **45**, 706-711, doi:10.1111/j.1532-5415.1997.tb01474.x (1997).

41 Nicholson, K. G., Kent, J., Hammersley, V. & Cancio, E. Acute viral infections of upper respiratory tract in elderly people living in the community: comparative, prospective, population based study of disease burden. *BMJ* **315**, 1060-1064, doi:10.1136/bmj.315.7115.1060 (1997).

42 Falsey, A. R., Walsh, E. E. & Hayden, F. G. Rhinovirus and coronavirus infection-associated hospitalizations among older adults. *J Infect Dis* **185**, 1338-1341, doi:10.1086/339881 (2002).

43 Atmar, R. L. *et al.* Respiratory tract viral infections in inner-city asthmatic adults. *Arch Intern Med* **158**, 2453-2459, doi:10.1001/archinte.158.22.2453 (1998).

44 Johnston, S. L. *et al.* Community study of role of viral infections in exacerbations of asthma in 9-11 year old children. *BMJ* **310**, 1225-1229, doi:10.1136/bmj.310.6989.1225 (1995).

45 Nicholson, K. G., Kent, J. & Ireland, D. C. Respiratory viruses and exacerbations of asthma in adults. *BMJ* **307**, 982-986, doi:10.1136/bmj.307.6910.982 (1993).

46 Tyrrell, D. A., Bynoe, M. L. & Hoorn, B. Cultivation of "difficult" viruses from patients with common colds. *Br Med J* **1**, 606-610, doi:10.1136/bmj.1.5592.606 (1968).

47 Ziebuhr, J., Snijder, E. J. & Gorbalenya, A. E. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J Gen Virol* **81**, 853-879, doi:10.1099/0022-1317-81-4-853 (2000).

48 Prentice, E., Jerome, W. G., Yoshimori, T., Mizushima, N. & Denison, M. R. Coronavirus replication complex formation utilizes components of cellular autophagy. *J Biol Chem* **279**, 10136-10141, doi:10.1074/jbc.M306124200 (2004).

49 Gosert, R., Kanjanahaluethai, A., Egger, D., Bienz, K. & Baker, S. C. RNA replication of mouse hepatitis virus takes place at double-membrane vesicles. *J Virol* **76**, 3697-3708, doi:10.1128/jvi.76.8.3697-3708.2002 (2002).

50 van der Meer, Y. *et al.* Localization of mouse hepatitis virus nonstructural proteins and RNA synthesis indicates a role for late endosomes in viral replication. *J Virol* **73**, 7641-7657, doi:10.1128/JVI.73.9.7641-7657.1999 (1999).

51 Shi, S. T. *et al.* 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, doi:10.1128/JVI.73.7.5957-5969.1999 (1999).

52 Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273, doi:10.1038/s41586-020-2012-7 (2020).

53 Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292 e286, doi:10.1016/j.cell.2020.02.058 (2020).

54 Letko, M., Marzi, A. & Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol* **5**, 562-569, doi:10.1038/s41564-020-0688-y (2020).

55 Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280 e278, doi:10.1016/j.cell.2020.02.052 (2020).

56 Shang, J. *et al.* Cell entry mechanisms of SARS-CoV-2. *Proc Natl Acad Sci U S A* **117**, 11727-11734, doi:10.1073/pnas.2003138117 (2020).

57 Li, M. Y., Li, L., Zhang, Y. & Wang, X. S. Expression of the SARS-CoV-2 cell receptor gene ACE2 in a wide variety of human tissues. *Infect Dis Poverty* **9**, 45, doi:10.1186/s40249-020-00662-x (2020).

58 Hikmet, F. *et al.* The protein expression profile of ACE2 in human tissues. *Mol Syst Biol* **16**, e9610, doi:10.15252/msb.20209610 (2020).

59 Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497-506, doi:10.1016/S0140-6736(20)30183-5 (2020).

60 Li, W. *et al.* Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450-454, doi:10.1038/nature02145 (2003).

61 Hui, D. S., Wong, P. C. & Wang, C. SARS: clinical features and diagnosis. *Respirology* **8 Suppl**, S20-24, doi:10.1046/j.1440-1843.2003.00520.x (2003).

62 Gheware, A. *et al.* ACE2 protein expression in lung tissues of severe COVID-19 infection. *Sci Rep* **12**, 4058, doi:10.1038/s41598-022-07918-6 (2022).

63 Chu, V. C., McElroy, L. J., Chu, V., Bauman, B. E. & Whittaker, G. R. The avian coronavirus infectious bronchitis virus undergoes direct low-pH-dependent fusion activation during entry into host cells. *J Virol* **80**, 3180-3188, doi:10.1128/JVI.80.7.3180-3188.2006 (2006).

64 Kawase, M., Shirato, K., Matsuyama, S. & Taguchi, F. Protease-mediated entry via the endosome of human coronavirus 229E. *J Virol* **83**, 712-721, doi:10.1128/JVI.01933-08 (2009).

65 Nash, T. C. & Buchmeier, M. J. Entry of mouse hepatitis virus into cells by endosomal and nonendosomal pathways. *Virology* **233**, 1-8, doi:10.1006/viro.1997.8609 (1997).

66 Finkel, Y. *et al.* The coding capacity of SARS-CoV-2. *Nature* **589**, 125-130, doi:10.1038/s41586-020-2739-1 (2021).

67 Bonilla, P. J., Hughes, S. A. & Weiss, S. R. Characterization of a second cleavage site and demonstration of activity in trans by the papain-like proteinase of the murine coronavirus mouse hepatitis virus strain A59. *J Virol* **71**, 900-909, doi:10.1128/JVI.71.2.900-909.1997 (1997).

68 Liu, D. X., Shen, S., Xu, H. Y. & Wang, S. F. Proteolytic mapping of the coronavirus infectious bronchitis virus 1b polyprotein: evidence for the presence of four cleavage sites of the 3C-like proteinase and identification of two novel cleavage products. *Virology* **246**, 288-297, doi:10.1006/viro.1998.9199 (1998).

69 Liu, D. X., Xu, H. Y. & Brown, T. D. Proteolytic processing of the coronavirus infectious bronchitis virus 1a polyprotein: identification of a 10-kilodalton polypeptide and determination of its cleavage sites. *J Virol* **71**, 1814-1820, doi:10.1128/JVI.71.3.1814-1820.1997 (1997).

70 Liu, D. X., Brierley, I., Tibbles, K. W. & Brown, T. D. A 100-kilodalton polypeptide encoded by open reading frame (ORF) 1b of the coronavirus infectious bronchitis virus is processed by ORF 1a products. *J Virol* **68**, 5772-5780, doi:10.1128/JVI.68.9.5772-5780.1994 (1994).

71 Harcourt, B. H. *et al.* Identification of severe acute respiratory syndrome coronavirus replicase products and characterization of papain-like protease activity. *J Virol* **78**, 13600-13612, doi:10.1128/JVI.78.24.13600-13612.2004 (2004).

72 Liu, D. X. & Brown, T. D. Characterisation and mutational analysis of an ORF 1a-encoding proteinase domain responsible for proteolytic processing of the infectious bronchitis virus 1a/1b polyprotein. *Virology* **209**, 420-427, doi:10.1006/viro.1995.1274 (1995).

73 Ng, L. F. & Liu, D. X. Further characterization of the coronavirus infectious bronchitis virus 3C-like proteinase and determination of a new cleavage site. *Virology* **272**, 27-39, doi:10.1006/viro.2000.0330 (2000).

74 Ng, L. F. & Liu, D. X. Identification of a 24-kDa polypeptide processed from the coronavirus infectious bronchitis virus 1a polyprotein by the 3C-like proteinase and determination of its cleavage sites. *Virology* **243**, 388-395, doi:10.1006/viro.1998.9058 (1998).

75 Ng, L. F. & Liu, D. X. Membrane association and dimerization of a cysteine-rich, 16-kilodalton polypeptide released from the C-terminal region of the coronavirus infectious bronchitis virus 1a polyprotein. *J Virol* **76**, 6257-6267, doi:10.1128/jvi.76.12.6257-6267.2002 (2002).

76 Denison, M. R. & Perlman, S. Translation and processing of mouse hepatitis virus virion RNA in a cell-free system. *J Virol* **60**, 12-18, doi:10.1128/JVI.60.1.12-18.1986 (1986).

77 Kamitani, W. *et al.* Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. *Proc Natl Acad Sci U S A* **103**, 12885-12890, doi:10.1073/pnas.0603144103 (2006).

78 Oudshoorn, D. *et al.* Expression and Cleavage of Middle East Respiratory Syndrome Coronavirus nsp3-4 Polyprotein Induce the Formation of Double-Membrane Vesicles That Mimic Those Associated with Coronaviral RNA Replication. *mBio* **8**, doi:10.1128/mBio.01658-17 (2017).

79 Lundin, A. *et al.* Targeting membrane-bound viral RNA synthesis reveals potent inhibition of diverse coronaviruses including the middle East respiratory syndrome virus. *PLoS Pathog* **10**, e1004166, doi:10.1371/journal.ppat.1004166 (2014).

80 Angelini, M. M., Akhlaghpour, M., Neuman, B. W. & Buchmeier, M. J. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. *mBio* **4**, doi:10.1128/mBio.00524-13 (2013).

81 Ricciardi, S. *et al.* The role of NSP6 in the biogenesis of the SARS-CoV-2 replication organelle. *Nature* **606**, 761-768, doi:10.1038/s41586-022-04835-6 (2022).

82 Overby, A. K., Popov, V. L., Niedrig, M. & Weber, F. Tick-borne encephalitis virus delays interferon induction and hides its double-stranded RNA in intracellular membrane vesicles. *J Virol* **84**, 8470-8483, doi:10.1128/JVI.00176-10 (2010).

83 Knoops, K. *et al.* SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biol* **6**, e226, doi:10.1371/journal.pbio.0060226 (2008).

84 Subissi, L. *et al.* One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proc Natl Acad Sci U S A* **111**, E3900-3909, doi:10.1073/pnas.1323705111 (2014).

85 Zhai, Y. *et al.* Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. *Nat Struct Mol Biol* **12**, 980-986, doi:10.1038/nsmb999 (2005).

86 Imbert, I. *et al.* A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J* **25**, 4933-4942, doi:10.1038/sj.emboj.7601368 (2006).

87 Kuchta, R. D., Reid, B. & Chang, L. M. DNA primase. Processivity and the primase to polymerase alpha activity switch. *J Biol Chem* **265**, 16158-16165 (1990).

88 te Velthuis, A. J., Arnold, J. J., Cameron, C. E., van den Worm, S. H. & Snijder, E. J. The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Res* **38**, 203-214, doi:10.1093/nar/gkp904 (2010).

89 te Velthuis, A. J., van den Worm, S. H. & Snijder, E. J. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Res* **40**, 1737-1747, doi:10.1093/nar/gkr893 (2012).

90 Tvarogova, J. *et al.* Identification and Characterization of a Human Coronavirus 229E Nonstructural Protein 8-Associated RNA 3'-Terminal Adenylyltransferase Activity. *J Virol* **93**, doi:10.1128/JVI.00291-19 (2019).

91 Lehmann, K. C. *et al.* Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. *Nucleic Acids Res* **43**, 8416-8434, doi:10.1093/nar/gkv838 (2015).

92 Lee, H. J. *et al.* The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* **180**, 567-582, doi:10.1016/0042-6822(91)90071-i (1991).

93 Cheng, A. *et al.* Expression, purification, and characterization of SARS coronavirus RNA polymerase. *Virology* **335**, 165-176, doi:10.1016/j.virol.2005.02.017 (2005).

94 Knipe, D. M. & Howley, P. *Fields Virology*. (Wolters Kluwer Health, 2013).

95 Bruenn, J. A. A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. *Nucleic Acids Res* **31**, 1821-1829, doi:10.1093/nar/gkg277 (2003).

96 Ruan, Z. *et al.* SARS-CoV-2 and SARS-CoV: Virtual screening of potential inhibitors targeting RNA-dependent RNA polymerase activity (NSP12). *J Med Virol* **93**, 389-400, doi:10.1002/jmv.26222 (2021).

97 Agostini, M. L. *et al.* Coronavirus Susceptibility to the Antiviral Remdesivir (GS-5734) Is Mediated by the Viral Polymerase and the Proofreading Exoribonuclease. *mBio* **9**, doi:10.1128/mBio.00221-18 (2018).

98 Khalili, J. S., Zhu, H., Mak, N. S. A., Yan, Y. & Zhu, Y. Novel coronavirus treatment with ribavirin: Groundwork for an evaluation concerning COVID-19. *J Med Virol* **92**, 740-746, doi:10.1002/jmv.25798 (2020).

99 Ivanov, K. A. & Ziebuhr, J. Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities. *J Virol* **78**, 7833-7838, doi:10.1128/JVI.78.14.7833-7838.2004 (2004).

100 Ivanov, K. A. *et al.* Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. *J Virol* **78**, 5619-5632, doi:10.1128/JVI.78.11.5619-5632.2004 (2004).

101 van Dinten, L. C., van Tol, H., Gorbalenya, A. E. & Snijder, E. J. The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis. *J Virol* **74**, 5213-5223, doi:10.1128/jvi.74.11.5213-5223.2000 (2000).

102 den Boon, J. A. *et al.* Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J Virol* **65**, 2910-2920, doi:10.1128/JVI.65.6.2910-2920.1991 (1991).

103 Andrew M.Q. King, M. J. A., Eric B. Carstens, Elliot J. Lefkowitz. *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*. (2011).

104 Deng, Z. *et al.* Structural basis for the regulatory function of a complex zinc-binding domain in a replicative arterivirus helicase resembling a nonsense-mediated mRNA decay helicase. *Nucleic Acids Res* **42**, 3464-3477, doi:10.1093/nar/gkt1310 (2014).

105 Gorbalenya, A. E. & Koonin, E. V. Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Res* **17**, 8413-8440, doi:10.1093/nar/17.21.8413 (1989).

106 Seybert, A., Hegyi, A., Siddell, S. G. & Ziebuhr, J. The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5'-to-3' polarity. *RNA* **6**, 1056-1068, doi:10.1017/s1355838200000728 (2000).

107 Kadare, G. & Haenni, A. L. Virus-encoded RNA helicases. *J Virol* **71**, 2583-2590, doi:10.1128/JVI.71.4.2583-2590.1997 (1997).

108 von Brunn, A. *et al.* Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFeome. *PLoS One* **2**, e459, doi:10.1371/journal.pone.0000459 (2007).

109 Tanner, J. A. *et al.* The severe acute respiratory syndrome (SARS) coronavirus NTPase/helicase belongs to a distinct class of 5' to 3' viral helicases. *J Biol Chem* **278**, 39578-39582, doi:10.1074/jbc.C300328200 (2003).

110 Snijder, E. J., Decroly, E. & Ziebuhr, J. The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing. *Adv Virus Res* **96**, 59-126, doi:10.1016/bs.aivir.2016.08.008 (2016).

111 Snijder, E. J. *et al.* Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J Mol Biol* **331**, 991-1004, doi:10.1016/s0022-2836(03)00865-9 (2003).

112 Eckerle, L. D., Lu, X., Sperry, S. M., Choi, L. & Denison, M. R. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J Virol* **81**, 12135-12144, doi:10.1128/JVI.01296-07 (2007).

113 Steinhauer, D. A., Domingo, E. & Holland, J. J. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene* **122**, 281-288, doi:10.1016/0378-1119(92)90216-c (1992).

114 Bouvet, M. *et al.* RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proc Natl Acad Sci U S A* **109**, 9372-9377, doi:10.1073/pnas.1201130109 (2012).

115 Bouvet, M. *et al.* Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes. *J Biol Chem* **289**, 25783-25796, doi:10.1074/jbc.M114.577353 (2014).

116 Chen, Y. *et al.* Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc Natl Acad Sci U S A* **106**, 3484-3489, doi:10.1073/pnas.0808790106 (2009).

117 Jin, X. *et al.* Characterization of the guanine-N7 methyltransferase activity of coronavirus nsp14 on nucleotide GTP. *Virus Res* **176**, 45-52, doi:10.1016/j.virusres.2013.05.001 (2013).

118 Decroly, E. *et al.* Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2'O)-methyltransferase activity. *J Virol* **82**, 8071-8084, doi:10.1128/JVI.00407-08 (2008).

119 Bouvet, M. *et al.* In vitro reconstitution of SARS-coronavirus mRNA cap methylation. *PLoS Pathog* **6**, e1000863, doi:10.1371/journal.ppat.1000863 (2010).

120 Perry, J. K. *et al.* An atomistic model of the coronavirus replication-transcription complex as a hexamer assembled around nsp15. *J Biol Chem* **297**, 101218, doi:10.1016/j.jbc.2021.101218 (2021).

121 Sawicki, S. G. & Sawicki, D. L. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J Virol* **64**, 1050-1056, doi:10.1128/JVI.64.3.1050-1056.1990 (1990).

122 Sawicki, S. G. & Sawicki, D. L. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv Exp Med Biol* **380**, 499-506, doi:10.1007/978-1-4615-1899-0_79 (1995).

123 Zuniga, S., Sola, I., Alonso, S. & Enjuanes, L. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J Virol* **78**, 980-994, doi:10.1128/jvi.78.2.980-994.2004 (2004).

124 Wang, D. *et al.* The SARS-CoV-2 subgenome landscape and its novel regulatory features. *Mol Cell* **81**, 2135-2147 e2135, doi:10.1016/j.molcel.2021.02.036 (2021).

125 Brant, A. C., Tian, W., Majerciak, V., Yang, W. & Zheng, Z. M. SARS-CoV-2: from its discovery to genome structure, transcription, and replication. *Cell Biosci* **11**, 136, doi:10.1186/s13578-021-00643-z (2021).

126 de Haan, C. A., Volders, H., Koetzner, C. A., Masters, P. S. & Rottier, P. J. Coronaviruses maintain viability despite dramatic rearrangements of the strictly conserved genome organization. *J Virol* **76**, 12491-12502, doi:10.1128/jvi.76.24.12491-12502.2002 (2002).

127 Grossoehme, N. E. *et al.* Coronavirus N protein N-terminal domain (NTD) specifically binds the transcriptional regulatory sequence (TRS) and melts TRS-cTRS RNA duplexes. *J Mol Biol* **394**, 544-557, doi:10.1016/j.jmb.2009.09.040 (2009).

128 Barcena, M. *et al.* Cryo-electron tomography of mouse hepatitis virus: Insights into the structure of the coronavirion. *Proc Natl Acad Sci U S A* **106**, 582-587, doi:10.1073/pnas.0805270106 (2009).

129 Krijnse-Locker, J., Ericsson, M., Rottier, P. J. & Griffiths, G. Characterization of the budding compartment of mouse hepatitis virus: evidence that transport from the RER to the Golgi complex requires only one vesicular transport step. *J Cell Biol* **124**, 55-70, doi:10.1083/jcb.124.1.55 (1994).

130 de Haan, C. A. & Rottier, P. J. Molecular interactions in the assembly of coronaviruses. *Adv Virus Res* **64**, 165-230, doi:10.1016/S0065-3527(05)64006-7 (2005).

131 Ghosh, S. *et al.* beta-Coronaviruses Use Lysosomes for Egress Instead of the Biosynthetic Secretory Pathway. *Cell* **183**, 1520-1535 e1514, doi:10.1016/j.cell.2020.10.039 (2020).

132 Cherry, J. D. & Krogstad, P. SARS: the first pandemic of the 21st century. *Pediatr Res* **56**, 1-5, doi:10.1203/01.PDR.0000129184.87042.FC (2004).

133 Peiris, J. S. *et al.* Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* **361**, 1319-1325, doi:10.1016/s0140-6736(03)13077-2 (2003).

134 Ware, L. B. & Matthay, M. A. The acute respiratory distress syndrome. *N Engl J Med* **342**, 1334-1349, doi:10.1056/NEJM200005043421806 (2000).

135 Song, H. D. *et al.* Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. *Proc Natl Acad Sci U S A* **102**, 2430-2435, doi:10.1073/pnas.0409608102 (2005).

136 Kan, B. *et al.* Molecular evolution analysis and geographic investigation of severe acute respiratory syndrome coronavirus-like virus in palm civets at an animal market and on farms. *J Virol* **79**, 11892-11900, doi:10.1128/JVI.79.18.11892-11900.2005 (2005).

137 Wu, D. *et al.* Civets are equally susceptible to experimental infection by two different severe acute respiratory syndrome coronavirus isolates. *J Virol* **79**, 2620-2625, doi:10.1128/JVI.79.4.2620-2625.2005 (2005).

138 Eaton, B. T. Introduction to Current focus on Hendra and Nipah viruses. *Microbes Infect* **3**, 277-278, doi:10.1016/s1286-4579(01)01380-6 (2001).

139 Li, W. *et al.* Bats are natural reservoirs of SARS-like coronaviruses. *Science* **310**, 676-679, doi:10.1126/science.1118391 (2005).

140 Hu, B. *et al.* Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog* **13**, e1006698, doi:10.1371/journal.ppat.1006698 (2017).

141 Bermingham, A. *et al.* Severe respiratory illness caused by a novel coronavirus, in a patient transferred to the United Kingdom from the Middle East, September 2012. *Euro Surveill* **17**, 20290 (2012).

142 Lu, G., Wang, Q. & Gao, G. F. Bat-to-human: spike features determining 'host jump' of coronaviruses SARS-CoV, MERS-CoV, and beyond. *Trends Microbiol* **23**, 468-478, doi:10.1016/j.tim.2015.06.003 (2015).

143 Achaiah, N. C., Subbarajasetty, S. B. & Shetty, R. M. R(0) and R(e) of COVID-19: Can We Predict When the Pandemic Outbreak will be Contained? *Indian J Crit Care Med* **24**, 1125-1127, doi:10.5005/jp-journals-10071-23649 (2020).

144 Delamater, P. L., Street, E. J., Leslie, T. F., Yang, Y. T. & Jacobsen, K. H. Complexity of the Basic Reproduction Number (R(0)). *Emerg Infect Dis* **25**, 1-4, doi:10.3201/eid2501.171901 (2019).

145 Breban, R., Riou, J. & Fontanet, A. Interhuman transmissibility of Middle East respiratory syndrome coronavirus: estimation of pandemic risk. *Lancet* **382**, 694-699, doi:10.1016/S0140-6736(13)61492-0 (2013).

146 Memish, Z. A. *et al.* Middle East respiratory syndrome coronavirus in bats, Saudi Arabia. *Emerg Infect Dis* **19**, 1819-1823, doi:10.3201/eid1911.131172 (2013).

147 Annan, A. *et al.* Human betacoronavirus 2c EMC/2012-related viruses in bats, Ghana and Europe. *Emerg Infect Dis* **19**, 456-459, doi:10.3201/eid1903.121503 (2013).

148 Anthony, S. J. *et al.* Coronaviruses in bats from Mexico. *J Gen Virol* **94**, 1028-1038, doi:10.1099/vir.0.049759-0 (2013).

149 Sharif-Yakan, A. & Kanj, S. S. Emergence of MERS-CoV in the Middle East: origins, transmission, treatment, and perspectives. *PLoS Pathog* **10**, e1004457, doi:10.1371/journal.ppat.1004457 (2014).

150 Meyer, B. *et al.* Antibodies against MERS coronavirus in dromedary camels, United Arab Emirates, 2003 and 2013. *Emerg Infect Dis* **20**, 552-559, doi:10.3201/eid2004.131746 (2014).

151 Haagmans, B. L. *et al.* Middle East respiratory syndrome coronavirus in dromedary camels: an outbreak investigation. *Lancet Infect Dis* **14**, 140-145, doi:10.1016/S1473-3099(13)70690-X (2014).

152 Corman, V. M. *et al.* Antibodies against MERS coronavirus in dromedary camels, Kenya, 1992-2013. *Emerg Infect Dis* **20**, 1319-1322, doi:10.3201/eid2008.140596 (2014).

153 Chu, D. K. *et al.* MERS coronaviruses in dromedary camels, Egypt. *Emerg Infect Dis* **20**, 1049-1053, doi:10.3201/eid2006.140299 (2014).

154 Alagaili, A. N. *et al.* Middle East respiratory syndrome coronavirus infection in dromedary camels in Saudi Arabia. *mBio* **5**, e00884-00814, doi:10.1128/mBio.00884-14 (2014).

155 Reusken, C. B. *et al.* Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. *Lancet Infect Dis* **13**, 859-866, doi:10.1016/S1473-3099(13)70164-6 (2013).

156 Perera, R. A. *et al.* Seroepidemiology for MERS coronavirus using microneutralisation and pseudoparticle virus neutralisation assays reveal a high prevalence of antibody in dromedary camels in Egypt, June 2013. *Euro Surveill* **18**, pii=20574, doi:10.2807/1560-7917.es2013.18.36.20574 (2013).

157 Ithete, N. L. *et al.* Close relative of human Middle East respiratory syndrome coronavirus in bat, South Africa. *Emerg Infect Dis* **19**, 1697-1699, doi:10.3201/eid1910.130946 (2013).

158 Corman, V. M. *et al.* Rooting the phylogenetic tree of middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat. *J Virol* **88**, 11297-11303, doi:10.1128/JVI.01498-14 (2014).

159 Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265-269, doi:10.1038/s41586-020-2008-3 (2020).

160 Singhal, T. A Review of Coronavirus Disease-2019 (COVID-19). *Indian J Pediatr* **87**, 281-286, doi:10.1007/s12098-020-03263-6 (2020).

161 Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565-574, doi:10.1016/S0140-6736(20)30251-8 (2020).

162 Angeles Montero-Fernandez, M. & Pardo-Garcia, R. Histopathology features of the lung in COVID-19 patients. *Diagn Histopathol (Oxf)* **27**, 123-127, doi:10.1016/j.mpdhp.2020.11.009 (2021).

163 Prieto-Perez, L. *et al.* Histiocytic hyperplasia with hemophagocytosis and acute alveolar damage in COVID-19 infection. *Mod Pathol* **33**, 2139-2146, doi:10.1038/s41379-020-0613-1 (2020).

164 Mohammed, M. E. A. The percentages of SARS-CoV-2 protein similarity and identity with SARS-CoV and BatCoV RaTG13 proteins can be used as indicators of virus origin. *J Proteins Proteom* **12**, 81-91, doi:10.1007/s42485-021-00060-3 (2021).

165 Coronavirus World Map: Tracking the Global Outbreak. *The New York Times* (2023). <https://www.nytimes.com/interactive/2021/world/covid-cases.html>.

166 Esclatine, A., Taddeo, B., Evans, L. & Roizman, B. The herpes simplex virus 1 UL41 gene-dependent destabilization of cellular RNAs is selective and may be sequence-specific. *Proc Natl Acad Sci U S A* **101**, 3603-3608, doi:10.1073/pnas.0400354101 (2004).

167 Esclatine, A., Taddeo, B. & Roizman, B. The UL41 protein of herpes simplex virus mediates selective stabilization or degradation of cellular mRNAs. *Proc Natl Acad Sci U S A* **101**, 18165-18170, doi:10.1073/pnas.0408272102 (2004).

168 Everly, D. N., Jr., Feng, P., Mian, I. S. & Read, G. S. mRNA degradation by the virion host shutoff (Vhs) protein of herpes simplex virus: genetic and biochemical evidence that Vhs is a nuclease. *J Virol* **76**, 8560-8571, doi:10.1128/jvi.76.17.8560-8571.2002 (2002).

169 Oroskar, A. A. & Read, G. S. Control of mRNA stability by the virion host shutoff function of herpes simplex virus. *J Virol* **63**, 1897-1906, doi:10.1128/JVI.63.5.1897-1906.1989 (1989).

170 Suzutani, T. *et al.* The role of the UL41 gene of herpes simplex virus type 1 in evasion of non-specific host defence mechanisms during primary infection. *J Gen Virol* **81**, 1763-1771, doi:10.1099/0022-1317-81-7-1763 (2000).

171 Pasieka, T. J., Lu, B. & Leib, D. A. Enhanced pathogenesis of an attenuated herpes simplex virus for mice lacking Stat1. *J Virol* **82**, 6052-6055, doi:10.1128/JVI.00297-08 (2008).

172 Covarrubias, S. *et al.* Coordinated destruction of cellular messages in translation complexes by the gammaherpesvirus host shutoff factor and the mammalian exonuclease Xrn1. *PLoS Pathog* **7**, e1002339, doi:10.1371/journal.ppat.1002339 (2011).

173 Glaunsinger, B. & Ganem, D. Lytic KSHV infection inhibits host gene expression by accelerating global mRNA turnover. *Mol Cell* **13**, 713-723, doi:10.1016/s1097-2765(04)00091-7 (2004).

174 Rowe, M. *et al.* Host shutoff during productive Epstein-Barr virus infection is mediated by BGLF5 and may contribute to immune evasion. *Proc Natl Acad Sci U S A* **104**, 3366-3371, doi:10.1073/pnas.0611128104 (2007).

175 Hartenian, E., Mendez, A. S., Didychuk, A. L., Khosla, S. & Glaunsinger, B. A. DNA processing by the Kaposi's sarcoma-associated herpesvirus alkaline exonuclease SOX contributes to viral gene expression and infectious virion production. *Nucleic Acids Res* **51**, 182-197, doi:10.1093/nar/gkac1190 (2023).

176 Uppal, T., Meyer, D., Agarwal, A. & Verma, S. C. The DNase Activity of Kaposi's Sarcoma-Associated Herpesvirus SOX Protein Serves an Important Role in Viral Genome Processing during Lytic Replication. *J Virol* **93**, doi:10.1128/JVI.01983-18 (2019).

177 Hackbart, M., Deng, X. & Baker, S. C. Coronavirus endoribonuclease targets viral polyuridine sequences to evade activating host sensors. *Proc Natl Acad Sci U S A* **117**, 8094-8103, doi:10.1073/pnas.1921485117 (2020).

178 Deng, X. *et al.* Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages. *Proc Natl Acad Sci U S A* **114**, E4251-E4260, doi:10.1073/pnas.1618310114 (2017).

179 Ivanov, K. A. *et al.* Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc Natl Acad Sci U S A* **101**, 12694-12699, doi:10.1073/pnas.0403127101 (2004).

180 Zheng, Y. *et al.* Research Progress on NSP11 of Porcine Reproductive and Respiratory Syndrome Virus. *Vet Sci* **10**, doi:10.3390/vetsci10070451 (2023).

181 Laneve, P. *et al.* Purification, cloning, and characterization of XendoU, a novel endoribonuclease involved in processing of intron-encoded small nucleolar RNAs in Xenopus laevis. *J Biol Chem* **278**, 13026-13032, doi:10.1074/jbc.M211937200 (2003).

182 Bhardwaj, K., Guarino, L. & Kao, C. C. The severe acute respiratory syndrome coronavirus Nsp15 protein is an endoribonuclease that prefers manganese as a cofactor. *J Virol* **78**, 12218-12224, doi:10.1128/JVI.78.22.12218-12224.2004 (2004).

183 Huang, T. *et al.* Kinetic analysis of RNA cleavage by coronavirus Nsp15 endonuclease: Evidence for acid-base catalysis and substrate-dependent metal ion activation. *J Biol Chem* **299**, 104787, doi:10.1016/j.jbc.2023.104787 (2023).

184 Guarino, L. A. *et al.* Mutational analysis of the SARS virus Nsp15 endoribonuclease: identification of residues affecting hexamer formation. *J Mol Biol* **353**, 1106-1117, doi:10.1016/j.jmb.2005.09.007 (2005).

185 Xu, X. *et al.* New antiviral target revealed by the hexameric structure of mouse hepatitis virus nonstructural protein nsp15. *J Virol* **80**, 7909-7917, doi:10.1128/JVI.00525-06 (2006).

186 Zheng, A. *et al.* 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, doi:10.1074/jbc.RA118.003756 (2018).

187 Frazier, M. N. *et al.* Characterization of SARS2 Nsp15 nuclease activity reveals it's mad about U. *Nucleic Acids Res* **49**, 10136-10149, doi:10.1093/nar/gkab719 (2021).

188 Ancar, R. *et al.* Physiologic RNA targets and refined sequence specificity of coronavirus EndoU. *RNA* **26**, 1976-1999, doi:10.1261/rna.076604.120 (2020).

189 Bhardwaj, K. *et al.* Structural and functional analyses of the severe acute respiratory syndrome coronavirus endoribonuclease Nsp15. *J Biol Chem* **283**, 3655-3664, doi:10.1074/jbc.M708375200 (2008).

190 Pillon, M. C. *et al.* Cryo-EM structures of the SARS-CoV-2 endoribonuclease Nsp15 reveal insight into nuclease specificity and dynamics. *Nat Commun* **12**, 636, doi:10.1038/s41467-020-20608-z (2021).

191 Frazier, M. N. *et al.* Flipped over U: structural basis for dsRNA cleavage by the SARS-CoV-2 endoribonuclease. *Nucleic Acids Res*, doi:10.1093/nar/gkac589 (2022).

192 Bhardwaj, K., Sun, J., Holzenburg, A., Guarino, L. A. & Kao, C. C. RNA recognition and cleavage by the SARS coronavirus endoribonuclease. *J Mol Biol* **361**, 243-256, doi:10.1016/j.jmb.2006.06.021 (2006).

193 Tengs, T. & Jonassen, C. M. Distribution and Evolutionary History of the Mobile Genetic Element s2m in Coronaviruses. *Diseases* **4**, doi:10.3390/diseases4030027 (2016).

194 Robertson, M. P. *et al.* The structure of a rigorously conserved RNA element within the SARS virus genome. *PLoS biology* **3**, e5, doi:10.1371/journal.pbio.0030005 (2005).

195 Janowski, A. The highly conserved stem-loop II RNA element (s2m) is critical for the lifecycle of astrovirus VA1 but is dispensable for SARS-CoV-2. *Journal of Pediatric Infectious Disease Society* (2022).

196 Yeh, T. Y. & Contreras, G. P. Emerging viral mutants in Australia suggest RNA recombination event in the SARS-CoV-2 genome. *Med J Aust* **213**, 44-44 e41, doi:10.5694/mja2.50657 (2020).

197 Jiang, H. *et al.* The Highly Conserved Stem-Loop II Motif Is Dispensable for SARS-CoV-2. *J Virol* **97**, e0063523, doi:10.1128/jvi.00635-23 (2023).

198 Yuen, C. K. *et al.* SARS-CoV-2 nsp13, nsp14, nsp15 and orf6 function as potent interferon antagonists. *Emerg Microbes Infect* **9**, 1418-1428, doi:10.1080/22221751.2020.1780953 (2020).

199 Shemesh, M. *et al.* SARS-CoV-2 suppresses IFNbeta production mediated by NSP1, 5, 6, 15, ORF6 and ORF7b but does not suppress the effects of added interferon. *PLoS Pathog* **17**, e1009800, doi:10.1371/journal.ppat.1009800 (2021).

200 Wu, Y. *et al.* Porcine Epidemic Diarrhea Virus nsp15 Antagonizes Interferon Signaling by RNA Degradation of TBK1 and IRF3. *Viruses* **12**, doi:10.3390/v12060599 (2020).

201 Comar, C. E. *et al.* MERS-CoV endoribonuclease and accessory proteins jointly evade host innate immunity during infection of lung and nasal epithelial cells. *Proc Natl Acad Sci U S A* **119**, e2123208119, doi:10.1073/pnas.2123208119 (2022).

202 Kim, Y. *et al.* Crystal structure of Nsp15 endoribonuclease NendoU from SARS-CoV-2. *Protein Sci* **29**, 1596-1605, doi:10.1002/pro.3873 (2020).

203 Zhang, L. *et al.* Structural and Biochemical Characterization of Endoribonuclease Nsp15 Encoded by Middle East Respiratory Syndrome Coronavirus. *J Virol* **92**, doi:10.1128/JVI.00893-18 (2018).

204 Weiss, S. R. & Navas-Martin, S. Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiol Mol Biol Rev* **69**, 635-664, doi:10.1128/MMBR.69.4.635-664.2005 (2005).

205 Miorin, L. *et al.* SARS-CoV-2 Orf6 hijacks Nup98 to block STAT nuclear import and antagonize interferon signaling. *Proc Natl Acad Sci U S A* **117**, 28344-28354, doi:10.1073/pnas.2016650117 (2020).

206 Zust, R. *et al.* Coronavirus non-structural protein 1 is a major pathogenicity factor: implications for the rational design of coronavirus vaccines. *PLoS Pathog* **3**, e109, doi:10.1371/journal.ppat.0030109 (2007).

207 Volk, A. *et al.* Coronavirus Endoribonuclease and Deubiquitinating Interferon Antagonists Differentially Modulate the Host Response during Replication in Macrophages. *J Virol* **94**, doi:10.1128/JVI.00178-20 (2020).

208 Kindler, E. *et al.* Early endonuclease-mediated evasion of RNA sensing ensures efficient coronavirus replication. *PLoS Pathog* **13**, e1006195, doi:10.1371/journal.ppat.1006195 (2017).

209 Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME Suite. *Nucleic Acids Res* **43**, W39-49, doi:10.1093/nar/gkv416 (2015).

210 Will, S., Joshi, T., Hofacker, I. L., Stadler, P. F. & Backofen, R. LocARNA-P: Accurate boundary prediction and improved detection of structural RNAs. *Rna* **18**, 900-914, doi:10.1261/rna.029041.111 (2012).

211 Raden, M. *et al.* Freiburg RNA tools: a central online resource for RNA-focused research and teaching. *Nucleic Acids Research* **46**, W25-W29, doi:10.1093/nar/gky329 (2018).

212 Hofacker, I. L. Vienna RNA secondary structure server. *Nucleic Acids Res* **31**, 3429-3431, doi:10.1093/nar/gkg599 (2003).

213 Tavares, R. C. A., Mahadeshwar, G., Wan, H., Huston, N. C. & Pyle, A. M. The global and local distribution of RNA structure throughout the SARS-CoV-2 genome. *J Virol* **95**, doi:10.1128/JVI.02190-20 (2021).

214 Andrews, R. J. *et al.* A map of the SARS-CoV-2 RNA structurome. *NAR Genom Bioinform* **3**, lqab043, doi:10.1093/nargab/lqab043 (2021).

215 Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R. & Hofacker, I. L. The Vienna RNA websuite. *Nucleic Acids Res* **36**, W70-74, doi:10.1093/nar/gkn188 (2008).

216 Huston, N. C. *et al.* Comprehensive in vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. *Mol Cell* **81**, 584-598 e585, doi:10.1016/j.molcel.2020.12.041 (2021).

217 Choi, R. *et al.* High-throughput screening of the ReFRAME, Pandemic Box, and COVID Box drug repurposing libraries against SARS-CoV-2 nsp15 endoribonuclease to identify small-molecule inhibitors of viral activity. *PLoS One* **16**, e0250019, doi:10.1371/journal.pone.0250019 (2021).

218 Tengs, T., Kristoffersen, A. B., Bachvaroff, T. R. & Jonassen, C. M. A mobile genetic element with unknown function found in distantly related viruses. *Virol J* **10**, 132, doi:10.1186/1743-422X-10-132 (2013).

219 Scharpf, M. *et al.* Antitermination in bacteriophage lambda. The structure of the N36 peptide-boxB RNA complex. *Eur J Biochem* **267**, 2397-2408, doi:10.1046/j.1432-1327.2000.01251.x (2000).

220 Legault, P., Li, J., Mogridge, J., Kay, L. E. & Greenblatt, J. NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell* **93**, 289-299, doi:10.1016/s0092-8674(00)81579-2 (1998).

221 Patel, D. K., Gebbie, M. P. & Lee, V. T. Assessing RNA interactions with proteins by DRaCALA. *Methods Enzymol* **549**, 489-512, doi:10.1016/B978-0-12-801122-5.00021-0 (2014).

222 Orr, M. W. & Lee, V. T. Differential Radial Capillary Action of Ligand Assay (DRaCALA) for High-Throughput Detection of Protein-Metabolite Interactions in Bacteria. *Methods Mol Biol* **1535**, 25-41, doi:10.1007/978-1-4939-6673-8_3 (2017).

223 Bhardwaj, K. *et al.* Structural and functional analyses of the severe acute respiratory syndrome coronavirus endoribonuclease Nsp15. *J Biol Chem* **283**, 3655-3664, doi:10.1074/jbc.M708375200 (2008).

224 Ricagno, S. *et al.* 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, doi:10.1073/pnas.0601708103 (2006).

225 Joseph, J. S. *et al.* Crystal structure of a monomeric form of severe acute respiratory syndrome coronavirus endonuclease nsp15 suggests a role for hexamerization as an allosteric switch. *J Virol* **81**, 6700-6708, doi:10.1128/JVI.02817-06 (2007).

226 Zhao, J. *et al.* Coronavirus Endoribonuclease Ensures Efficient Viral Replication and Prevents Protein Kinase R Activation. *J Virol*, doi:10.1128/JVI.02103-20 (2020).

227 Deng, X. *et al.* Inactivating Three Interferon Antagonists Attenuates Pathogenesis of an Enteric Coronavirus. *J Virol* **94**, doi:10.1128/JVI.00565-20 (2020).

228 Liu, X. *et al.* Porcine deltacoronavirus nsp15 antagonizes interferon-beta production independently of its endoribonuclease activity. *Mol Immunol* **114**, 100-107, doi:10.1016/j.molimm.2019.07.003 (2019).

229 Shokri, S., Mahmoudvand, S., Taherkhani, R. & Farshadpour, F. Modulation of the immune response by Middle East respiratory syndrome coronavirus. *J Cell Physiol* **234**, 2143-2151, doi:10.1002/jcp.27155 (2019).

230 Lei, Y. *et al.* MAVS-mediated apoptosis and its inhibition by viral proteins. *PLoS One* **4**, e5466, doi:10.1371/journal.pone.0005466 (2009).

231 Bhardwaj, K., Liu, P., Leibowitz, J. L. & Kao, C. C. The coronavirus endoribonuclease Nsp15 interacts with retinoblastoma tumor suppressor protein. *J Virol* **86**, 4294-4304, doi:10.1128/JVI.07012-11 (2012).

232 Athmer, J. *et al.* Selective Packaging in Murine Coronavirus Promotes Virulence by Limiting Type I Interferon Responses. *mBio* **9**, doi:10.1128/mBio.00272-18 (2018).

233 Stefl, R. & Allain, F. H. A novel RNA pentaloop fold involved in targeting ADAR2. *RNA* **11**, 592-597, doi:10.1261/rna.7276805 (2005).

234 Vicens, Q. & Kieft, J. S. Thoughts on how to think (and talk) about RNA structure. *Proc Natl Acad Sci U S A* **119**, e2112677119, doi:10.1073/pnas.2112677119 (2022).

235 Godoy, A. S. *et al.* Allosteric regulation and crystallographic fragment screening of SARS-CoV-2 NSP15 endoribonuclease. *Nucleic Acids Res*, doi:10.1093/nar/gkad314 (2023).

236 Frazier, M. N., Riccio, A. A., Wilson, I. M., Copeland, W. C. & Stanley, R. E. Recent insights into the structure and function of coronavirus ribonucleases. *FEBS Open Bio* **12**, 1567-1583, doi:10.1002/2211-5463.13414 (2022).

237 Godoy, A. S. *et al.* Allosteric regulation and crystallographic fragment screening of SARS-CoV-2 NSP15 endoribonuclease. *Nucleic Acids Res* **51**, 5255-5270, doi:10.1093/nar/gkad314 (2023).

238 Gaglia, M. M., Rycroft, C. H. & Glaunsinger, B. A. Transcriptome-Wide Cleavage Site Mapping on Cellular mRNAs Reveals Features Underlying Sequence-Specific Cleavage by the Viral Ribonuclease SOX. *PLoS Pathog* **11**, e1005305, doi:10.1371/journal.ppat.1005305 (2015).

239 Athmer, J. *et al.* In Situ Tagged nsp15 Reveals Interactions with Coronavirus Replication/Transcription Complex-Associated Proteins. *mBio* **8**, doi:10.1128/mBio.02320-16 (2017).

240 Reina, J. & Iglesias, C. [Nirmatrelvir plus ritonavir (Paxlovid) a potent SARS-CoV-2 3CLpro protease inhibitor combination]. *Rev Esp Quimioter* **35**, 236-240, doi:10.37201/req/002.2022 (2022).

241 Avis, J. M., Conn, G. L. & Walker, S. C. Cis-acting ribozymes for the production of RNA in vitro transcripts with defined 5' and 3' ends. *Methods Mol Biol* **941**, 83-98, doi:10.1007/978-1-62703-113-4_7 (2012).