Characterization and Inhibition of Viral Nuclease Enzymes

by

Brendan Geoffrey John Todd

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

Doctor of Philosophy

in

Virology

Department of Medical Microbiology and Immunology
University of Alberta

© Brendan Geoffrey John Todd, 2021

Abstract

Viral infection accounts for significant morbidity, mortality, and economic loss worldwide. At no time in modern history has this been more apparent than it is now given the current global pandemic. This problem is compounded by the emergence of viral strains resistant to currently available therapeutics. Therefore, the identification and characterization of antiviral compounds with novel targets and mechanisms of action is urgently required. One such underexplored target that has recently shown promise is viral nuclease enzymes. These enzymes are required for the replication of several viruses including herpesviruses, influenza viruses, SARS-CoV-2, HIV, and others.

In this thesis, I expressed and purified nuclease enzymes from three clinically relevant human viruses with the goal of developing biochemical systems to allow us to understand the mechanism of action of antiviral drugs. To this end, I expressed and purified the terminase complex of cytomegalovirus, the RNA dependent RNA polymerase heterotrimer of influenza B (FluB-ht), and the proofreading exonuclease complex of SARS-CoV-2. These enzymes were chosen as they had been shown to either be the target of or involved in resistance to recently approved FDA approved antiviral drugs. For example, terminase is the target of letermovir, FluB-ht is the target of baloxavir marboxil, and the SARS-CoV-2 proofreading exonuclease may be involved in resistance against nucleoside analogue inhibitors of SARS-CoV-2 such as remdesivir.

First, I successfully expressed and purified the cytomegalovirus (CMV) terminase complex and developed assays to monitor its ATPase, nuclease, and DNA binding activities. As this complex is known to be the target of letermovir, it was our goal to elucidate the mechanism of action

underlying this new FDA approved drug. The assays that I have developed will facilitate this process. Furthermore, by truncating our expression construct, I was able to show that a small protein lacking enzymatic activity, UL51, is necessary for terminase expression.

Secondly, with regards to the influenza B polymerase heterotrimer (FluB-ht), I developed assays to monitor the endonuclease and transcriptional activities of the complex. These assays revealed that the active form of the FDA approved influenza endonuclease (PA) inhibitor baloxavir marboxil, baloxavir acid (BXA), is a slow tight binding competitive inhibitor. Furthermore, I was able to show that BXA is a selective inhibitor of the endonuclease as it displays a greater than 50-fold selectivity for the endonuclease activity.

Finally, with respect to the SARS-CoV-2 exonuclease, I successfully purified the active complex and showed that this complex can excise incorporated nucleotide analogues with analogues that represent mismatched base-pairs being more readily excised. Furthermore, I was able to show that under the right circumstances, the excision of the FDA approved SARS-CoV-2 inhibitor remdesivir is possible. This revealed a possible avenue for the development of resistance against nucleoside analogue inhibitors of SARS-CoV-2 and produced an assay for the excision of nucleoside analogue inhibitors *in vitro*.

Preface

Dr. Hsiao-Wei Liu assisted with the expression and purification of several CMV terminase preparations used in chapter 3.

Portions of chapter four have been previously published as Todd B., Tchesnokov E.P., and Götte M. “The active form of the influenza cap-snatching endonuclease inhibitor baloxavir marboxil is a tight binding inhibitor” Journal of Biological Chemistry, vol. 296, 2021. I performed all the experimental procedures present in the manuscript. Dr. E.P. Tchesnokov and Dr. M. Götte assisted me with experimental design and the writing of the manuscript.

Dr. E.P. Tchesnokov assisted with the design and execution of several experimental procedures shown in chapter 5. These contributions are outlined in the chapter.

Emma Woolner and Dr. Dana Kocincova assisted with the size exclusion purification of the SARS-CoV-2 exonuclease complex nsp10/nsp14. Dana Kocincova also generously provided recombinant SARS-CoV-2 RNA depenent RNA polymerase complexes for use in the experiments detailed in chapter 5.

Dedication

I dedicate this work to my family and friends. Without your tireless support the completion of this PhD thesis would not have been possible. Thank you for being there for me.

Acknowledgements

This work could not have been brought to completion without the support, mentorship, and guidance of my supervisor, Dr. Matthias Götte. I would also like to thank the members of my supervisory committee, Dr. Smiley and Dr. Glover, for their insights and guidance. The support and guidance of my fellow lab members was also instrumental to the successful completion of this task. Specifically, I would like to thank my fellow lab members, past and present, Egor Tchesnokov, Dana Kocinkova, Emma Woolner, Calvin Gordon, Hsiao-Wei Liu, Michael Menni, Kieran Maheden, Marriane Ngure, Parisa Raesimakiani, and others. The discussions we had about science and life were essential to my development as a scientist and a human being. Thank you for your support and guidance over the years.

Additionally, I would like to thank all my friends in the Medical Microbiology and Immunology department community. Specifically, Kevin Joannou, Julia May, Farah Elawar, and all the others who made the department such a pleasant place to work and learn over the years. Also, Kevin, thank you for starting up a departmental Dungeons & Dragons™ group! I had so much fun playing tabletop games with you guys! Thank you for being such awesome people!

Finally, I would like to thank my family and friends for their support as I pursued my PhD. Specifically, I would like to thank my close friends Thomas McConnell, Gabriel Caciula, and Hsiao-Wei Liu. Science, like life, can often feel like a roller-coaster ride of emotions. One day the experiments are working as expected and you are generating publishable data and the next day you are back at the drawing board troubleshooting. Thank you for being there for me. Without you I never could have completed this degree program.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

bp – basepair

nt - nucleotide

HHV5 – human herpesvirus 5 (Cytomegalovirus)

CMV – cytomegalovirus

FluB – Influenza B

SARS-CoV – severe acute respiratory syndrome coronavirus

MERS-CoV – middle east respiratory syndrome coronavirus

SARS-CoV-2 – severe acute respiratory syndrome coronavirus two

T3 – cytomegalovirus terminase complex UL51/UL56/UL89

ExoN - exonuclease

nt – nucleotide

wt – wildtype

FluB-ht – influenza B RNA dependent RNA polymerase heterotrimer (PA/PB1/PB2)

BXA – baloxavir acid

BXM – baloxavir marvoxil

RDV – remdesivir

ATP – adenosine triphosphate

UTP – uridine triphosphate

TP – Triphosphate

NSP – nonstructural protein

UL – unique long

PA – polymerase Acidic

PB – Polymerase Basic

RNA – ribonucleic Acid

DNA – deoxyribonucleic acid

RdRp – RNA dependent RNA polymerase

RdDp – RNA dependent DNA polymerase

DdRp – DNA dependent RNA polymerase

DdDp – DNA dependent DNA polymerase

EDTA – ethyenediaminetetraacetic acid

TBE - Tris-HCL, Boric Acid, EDTA

PAGE – polyacrylamide gel electrophoresis

SDS – sodium dodecyl sulphate

PCR – polymerase chain reaction

Cy5 – Cyanine 5 dye

NiNTA – nickel nitrilotriacetic acid

Strep – streptavidin

LC-MS/MS – Liquid Chromatography with tandem mass spectroscopy

Chapter 1: Introduction

1.1 The Public Health Burden of Viral Disease

Viral infection has been and continues to be responsible for significant morbidity, mortality, and economic loss worldwide. The severe public health threat posed by viruses has at no time in modern history been as apparent as it is now amid the devastation wreaked by the current SARS-CoV-2 pandemic which, at the time of writing, has infected 160 million people around the globe leading to 3.3 million deaths. However, this pales in comparison to the 1918 influenza pandemic and for that reason any discussion of pandemics must begin with the Spanish influenza pandemic of 1918 which has been referred to as "the mother of all pandemics" (1). This grim moniker is well deserved given that the 1918 Spanish influenza virus infected approximately one third of the world's population resulting in 50 million deaths (1). However, despite this, we still do not understand the factors underlying the unusual pathogenicity of the influenza strain(s) responsible for the 1918 pandemic (1). However, it has been hypothesized that the pathogenicity of these strains may have resulted from the fact that the virus crossed the species barrier from birds to humans (1-3). This is because recent zoonoses, that is viruses that have recently crossed the species barrier to infect humans, are often highly pathogenic because they are poorly adapted to replicating in a human host (4). Given this, viral zoonoses represent a major public health threat worldwide.

The significant public health threat posed by viral zoonoses is illustrated by the emergence of several viral diseases from wild animals in recent times including human immunodeficiency virus (HIV), ebola virus, the middle east respiratory syndrome coronavirus (MERS), severe acute respiratory syndrome coronavirus (SARS), and, the causative agent of COVID-19, severe acute

respiratory virus coronavirus two (SARS-CoV-2). Specifically, HIV originated in chimpanzees while Ebola and the MERS, SARS, and SARS-CoV-2 coronaviruses are hypothesized to have a bat origin (5-7). However, while all these viruses are important in their own right, HIV and SARS-CoV-2 are arguably the most relevant today given the profound, recent, and continuing public health impacts of the HIV and COVID-19 pandemics.

In the early 1980s, increasing numbers of young gay and bisexual men perished due to opportunistic infections and rare cancers (5). This was determined to be the result of an immunodeficiency which would later be described in the medical literature as acquired immunodeficiency syndrome (8). Soon after, the retrovirus (HIV-1) was isolated and determined to be responsible for this disease via its effects on CD4+ T-cells (8-11). This virus would spread around the world leading to a pandemic that, to date, has resulted in the mortality of over 40 million people (12). This is partly because, before the mid-1990s, few antiviral treatments were available and treatment of HIV infection largely consisted of prophylaxis for opportunistic infections and supportive care (13). Due to a lack of effective treatments, HIV infection was largely considered a “death sentence” as patients would almost invariably progress to develop acquired immune deficiency syndrome (AIDS) and succumb to opportunistic infections (14). However, with the advent of small molecule inhibitors targeting viral enzymes, i.e., reverse transcriptase and protease inhibitors, it became possible to inhibit viral replication which would in turn prevent patients from developing AIDS (13,14). Thus, HIV infection went from being a fatal disease to a treatable, chronic condition (12,15). However, despite these advances much still remains to be done to work towards an eventual cure as approximately 700,000 people succumbed to complications resulting from AIDS in 2019 (16).

Furthermore, the significant public health burden imposed by emerging viral disease is highlighted by the continuing COVID-19 pandemic. While several treatments and effective vaccines have emerged since the pandemic began, vaccine hesitancy, and the potential of the virus to mutate and acquire resistance to currently available therapeutics mean that the public health and economic impacts of COVID-19 may continue for years to come (17-19).

Thus, biochemical systems that allow for the characterization of viral enzymes and the elucidation of the mechanisms surrounding antiviral action and resistance are urgently required to combat the significant worldwide public health burden imposed by viral infection. This is addressed in this thesis where we develop such systems to study antivirals targeting nuclease enzymes from several clinically relevant human viruses including influenza, SARS-CoV-2, and cytomegalovirus.

1.2 What is a nuclease?

Nucleases are, as the name suggests, enzymes that cleave the phosphodiester backbone bonds of nucleic acid polymers, i.e. DNA or RNA (20). These enzymes are generally classified with regards to the type of cleavage they perform as either endo or exonucleases (20). The difference being that exonucleases cut at the end of the DNA or RNA polymer while endonucleases cut at an internal position (20). These enzymes are integral to many biological processes including but not limited to nucleic acid replication and metabolism, DNA repair, transcription, and cell signaling (20-22). Furthermore, nucleases are involved in the replication of many viruses including herpesviruses, influenza viruses, coronaviruses, retroviruses, and others (21,23-27). Given the significant burden of viral infections on human health, in this thesis, we chose to study the function and inhibition of viral nucleases.

1.3 Nuclease Biochemistry

So far, all nucleases that have been described can be assigned to one of three classes based on the metal ions involved in their biochemical mechanism of action (20). These are two-metal ion-dependent nucleases, one-metal ion-dependent nucleases, and metal ion independent nucleases (20).

Metal ion-dependent nucleases are those nucleases that exhibit their function by coordinating divalent metal ions. Specifically, these endonucleases contain negatively charged resides such as aspartic or glutamic acid. These residues co-ordinate (either one or two) divalent metal ions (usually Mg^{2+} or Mn^{2+}) which generally polarize a water molecule (20). This polarized water molecule then acts as a nucleophile and attacks the phosphate backbone of the nucleic acid polymer facilitating cleavage. Metal ion independent nucleases are similar except they use the sidechain of a polar residue such as serine, threonine, or histidine as a nucleophile (20).

Of all of these, two metal ion-dependent endonucleases are the most commonly observed and they facilitate the widest range of biological functions (20). It is perhaps unsurprising then that the nucleases studied in this thesis fall into this category (25,27-29).

1.4 Role of Nucleases in the biology of clinically relevant mammalian viruses

Nucleases are involved in the replication of several mammalian viruses, as previously noted, and their inhibition is a promising avenue for drug design as evidenced by the recent FDA-approved antivirals targeting these enzymes (30,31). As these enzymes are now validated drug targets, nucleases from some clinically relevant viruses and antiviral compounds targeting these enzymes will be discussed in detail.

1.4.1 HIV, RNAse H, Integrase, and Inhibitors

Since human immunodeficiency virus (HIV), the virus responsible for acquired immune deficiency syndrome (AIDS) first emerged and became widely recognized in the 1980s significant progress has been made with regards to the management of HIV infection (32). Specifically, with the advent of antiretroviral therapy HIV patients receiving treatment now have a normal life expectancy (33,34). This development was borne out of years of research on compounds that inhibit several enzymes involved in the life cycle of HIV including protease, integrase, and reverse transcriptase (RT) (35). Both integrase and the RNAse H activity of reverse transcriptase (RNase H) have similar active sites, a two-metal ion catalytic mechanism, and have been studied in detail (26,36-38).

Reverse transcriptase (RT) is a multifunctional enzyme that is essential to the lifecycle of HIV (39). RT encodes an RNA-dependent RNA polymerase activity (RdRp), RNA-dependent DNA polymerase activity (RdDp), DNA-dependent DNA polymerase activity (DdDp), and an RNase activity (RNAse H) (26,39). Specifically, the polymerase activities of RT are responsible for converting the RNA genome of HIV into double-stranded DNA (dsDNA) so that it can be integrated into the genome of the host cell by integrase (39). RNAse H is responsible for degrading the RNA portions of RNA/DNA intermediates that arise during this process (39). Thus, inhibition of RNAse H or integrase prevents HIV replication and this has been exploited for the development of inhibitors (26,40-43).

Most inhibitors function by binding to and chelating metal ions at the active site of RNAse H or Integrase thereby preventing the interaction of the enzyme with its nucleic acid substrate

(26,36,41-44). This research has led to the FDA approval of several integrase inhibitors for the management of HIV infection, while RNase H inhibitors are still investigational (26,40). Nevertheless, these nucleases are a validated target for the development of antivirals and the similarity of the two-metal ion active sties of HIV RNAse H and integrase to other viral nucleases is significant (25,31). For example, the FDA-approved integrase inhibitor raltegravir was used as a scaffold for the development of the now FDA-approved influenza cap snatching endonuclease inhibitor baloxavir marboxil (25). Indeed, the similarity of the active sites of HIV RNAse H/Integrase and the nuclease actives sites of herpesviral terminases, and the influenza cap snatching endonuclease provided a theoretical basis for much of the research detailed in this thesis.

1.4.2 Herpesviral DNA elements (Pac Elements) involved in genome packaging.

To date, the majority of antivirals against herpesviruses work by inhibition of the viral DNA polymerase (45). This prevents viral genome replication, however many other aspects of the viral lifecycle can be targeted by antiviral drugs. For example, to successfully replicate a virus must enter a susceptible cell, replicate its genome, package these newly produced genomes into virions, and finally release these from the cell (Figure 1.1). One aspect of the herpesviral life cycle that has recently been targeted for antiviral drug development is genome packaging (30). Therefore, the genetic elements involved in herpesviral genome packaging will be discussed here.

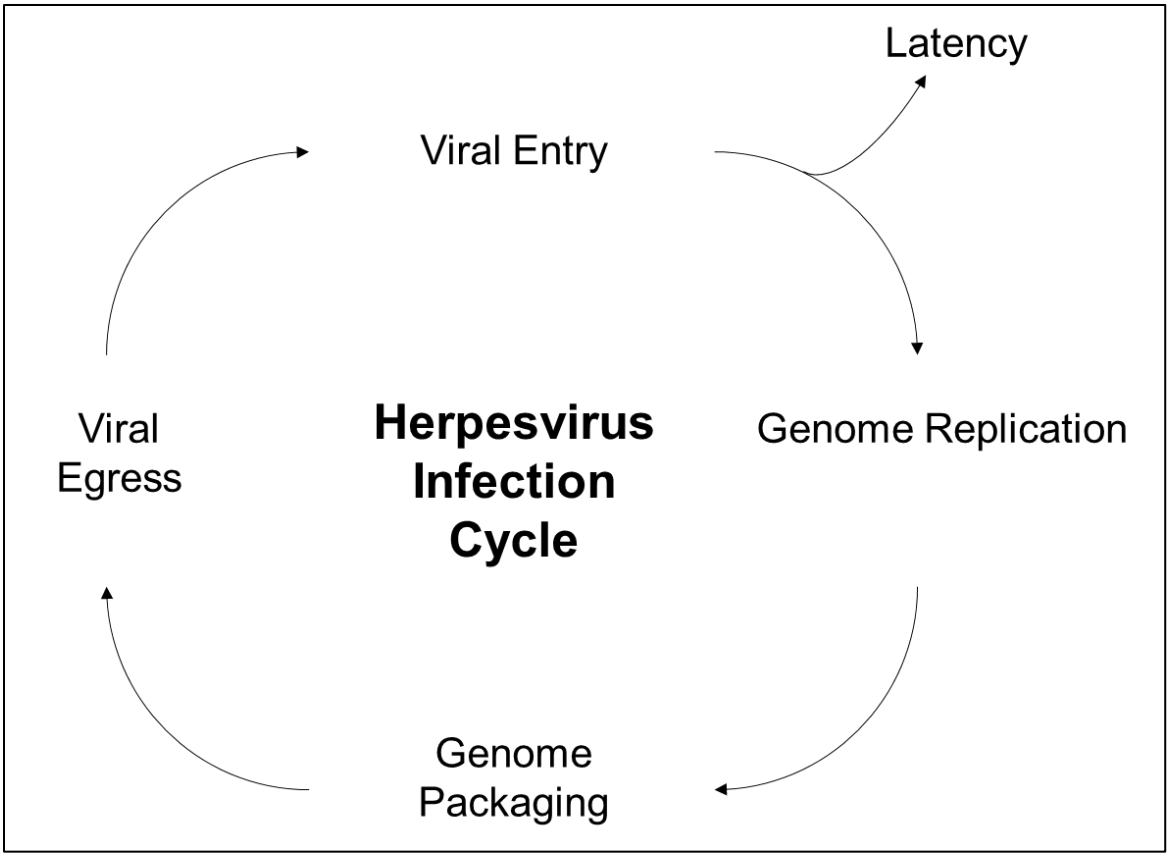


Figure 1.1 Schematic detailing the lifecycle stages of a typical herpesvirus.

The genomes of the prototypical herpesvirus herpes simplex virus 1 (HSV-1) and other herpesviruses consist of covalently linked long and short segments of dsDNA referred to as the unique long (U_L) and unique short (U_s) regions. In general, the U_L and U_s regions are flanked by highly repetitive repeats referred to as the a, b, and c sequences, respectively. Of these sequences, the a sequence is thought to be crucial for genome maturation and packaging (46). This is because these sequences are found at both ends of the genome and they contain conserved cis-acting sequences referred to as *Pac1* and *Pac2*. *Pac1* consists of a 3-7 base pair poly A region that is flanked by poly C tracts while *pac2* consists of 5-10 base pair conserved A rich region (47,48). These sequences were first identified by Deiss *et al.* in 1986 and shown by subsequent studies to be required for viral cleavage and packaging (47,49-60). These sequences are found at opposing

ends of the genome, separated by approximately 60 base pairs, and are thought to be recognized by the viral packaging machinery such that a cut is generated between the two motifs (Figure 1.2). (60) This is important because herpesvirus replication produces high molecular weight concatamers, consisting of head to tail linked genomes, that must be resolved and packaged to generate infectious virions (Figure 1.2 and 1.3) (61). This process is discussed in further detail in section 1.4.3. Due to the conservation and location of the a sequences, they were hypothesized to be important for recognition by the viral DNA packaging enzymes, here referred to as the viral terminase (62).

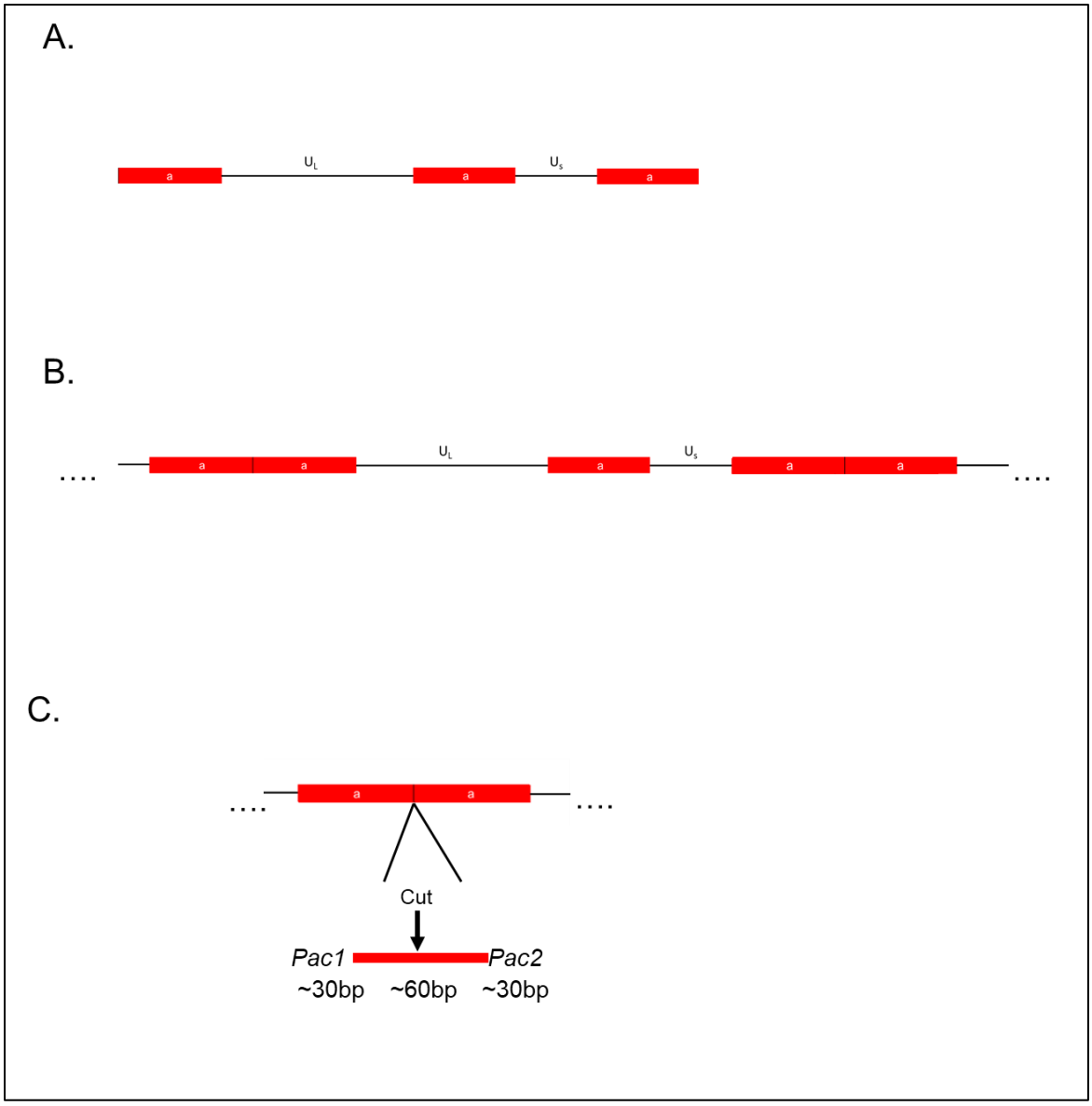


Figure 1.2 Position of packaging motifs (a) in a prototypical herpesviral genome or genome concatamer. A. Position of the a sequence (red boxes) relative to the unique long U_L and unique short U_s genome sequences. B. Orientation of the a sequences in a typical herpesviral concatemer. C. position of the Pac 1 and Pac 2 motifs relative to the position at which endonucleolytic cleavage and concatemer resolution occurs.

1.4.3 Herpesvirus Terminase Proteins

Herpesviral replication produces high molecular weight concatemers of several genomes linked head to tail (Figure 1.2 and Figure 1.3) (63,64).

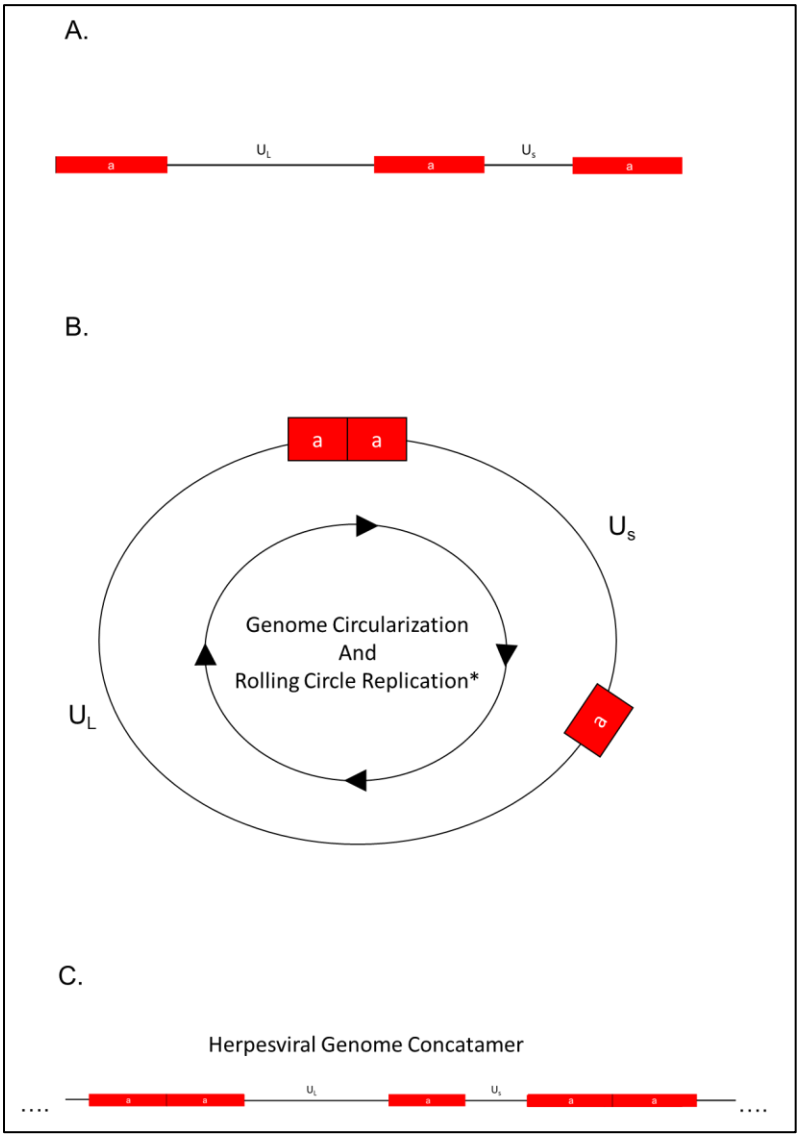


Figure 1.3 Schematic detailing herpesviral genome structure and the genome elements involved in viral genome packaging. A. Schematic of a herpesviral genome with the sequences involved in packaging (a sequences) indicated with red boxes. B. Schematic diagram of a

circularized herpesvirus genome which serves as a template for the generation of head to tail linked concatamers of genomic DNA. (*) indicates that this replication mechanism of herpesviruses is still controversial. C. Schematic diagram of a herpesviral genome concatemer.

This was historically considered to be the result of rolling circle replication, however, this is controversial as data exists which suggest an alternative recombination based mechanism (65). Nevertheless, regardless of the nature of the replication mechanism, these concatamers are produced and they must be packaged into viral capsids to produce infectious virions (66,67). This packaging is performed by an enzyme referred to as the viral terminase (61). In CMV, terminase is a heterotrimer consisting of three proteins UL51, UL56, and UL89 (61,67). Models for genome packaging in CMV and other herpesviruses were first derived based on analogy to mechanisms utilized by dsDNA bacteriophages (68,69). This is logical given the evolutionary relatedness of herpesviruses and tailed dsDNA bacteriophages(69,70). Specifically, two main mechanisms, headful and sequence specific packaging have been described in dsDNA bacteriophages (71,72). Headful packaging, utilized by phages such as T4, is a biophysical process by which a capsid is packed to completion, at which point the nuclease is activated to cleave the concatemer (72). This is in contrast to sequence specific packaging mechanisms, such as those described for phage T3 and phage T7, where conserved sequences found within the viral concatemer determine where packaging begins and where cleavage occurs (71). This later mechanism is reminiscent of those which have been described for herpesviruses.

In herpesviruses, the viral terminase binds to conserved sequences at the terminus of the genome referred to as pac elements (73). These sequences are thought to drive the initiation of and completion of viral genome packaging (54,62,74). Specifically, the prevailing model of herpesviral

genome packaging suggests that the terminase binds to pac elements at the end of the genome and then, uses energy derived from ATP hydrolysis to translocate a single genome through the portal vertex, which serves as a channel for genome packaging, and into the viral capsid (61) (Figure 1.4).

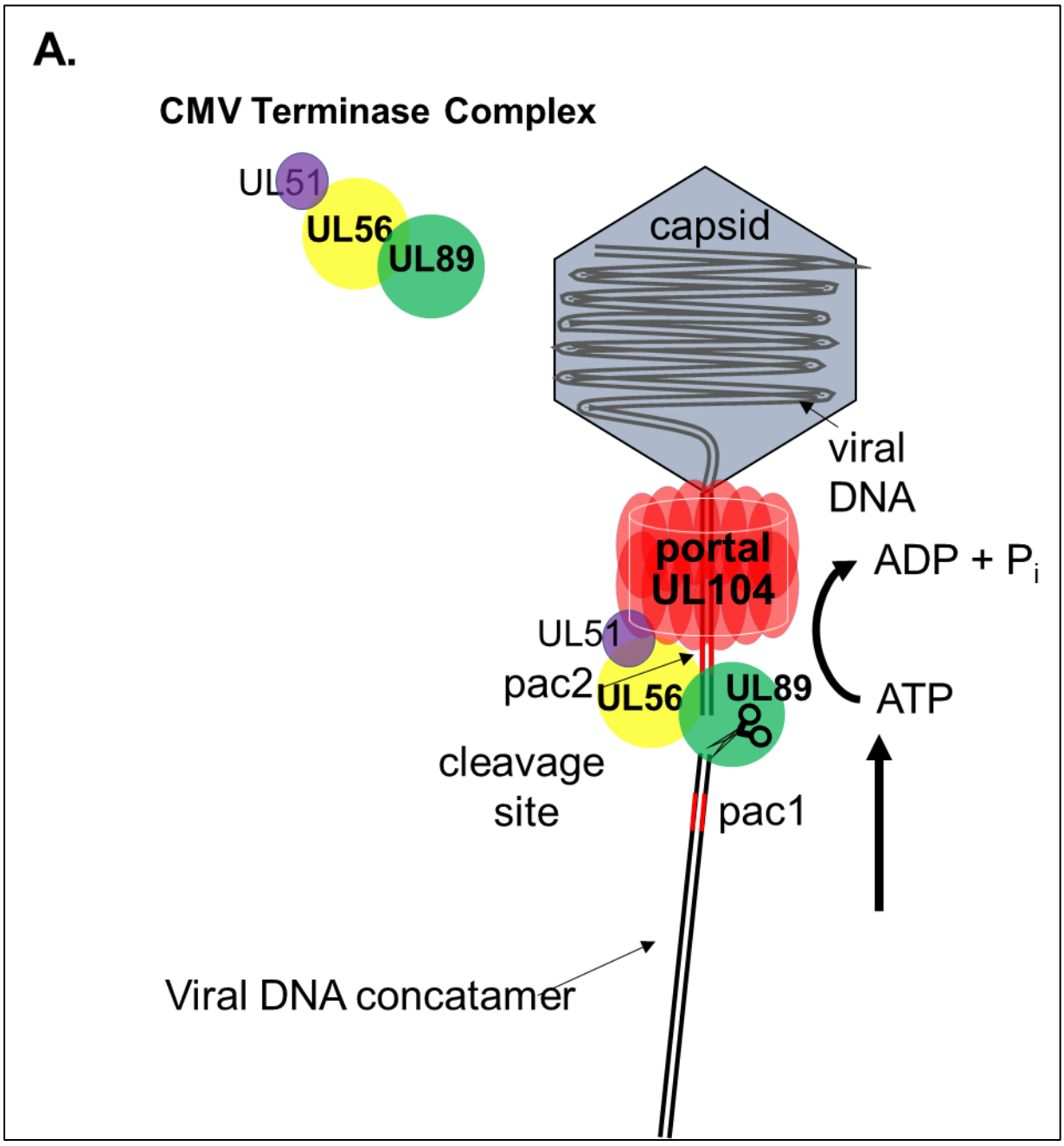


Figure 1.4. The CMV Terminase complex UL51/UL56/UL89 and viral genome packaging. A. Schematic diagram detailing both the composition of the CMV terminase complex and its role in viral genome packaging. The portal vertex is depicted as a dodecameric complex based on analogy to phage T4 (75).

Once a single genome is packaged, the terminase then cleaves the concatemer releasing a single genome into the capsid (61). Currently, the precise biochemical mechanisms responsible for this packaging are unclear. However, it has been suggested that interactions between the conserved pac sequences described above, the terminase, and the portal vertex proteins are crucial for packaging and this has been supported by recent structural work (76).

Previous work aimed at a biochemical characterization of the CMV and herpesviral terminases in general have been primarily limited to cell culture based approaches and have been largely focused on HSV-1 (77). Specifically, it has been shown that mutation or inhibition of terminase proteins results in the accumulation of uncleaved concatemeric DNA and empty capsids within the nucleus (30,77). Furthermore, the components of the HSV-1 terminase (UL15, UL28, and UL33) were shown to form a complex based on the fact that these components copurified both from infected cells and when expressed endogenously using a baculovirus based expression system (78,79).

The components of the CMV terminase complex, UL51, UL56, and Ul89, were identified based on homology to their counterparts in HSV-1 UL15, UL28, and UL33 (61,80). Similarly, these components were demonstrated to form a complex based on the fact that they can be copurified when expressed endogenously using a baculovirus based system (67,80). However, the biochemistry of the CMV terminase complex remains understudied (61). What little work that has been done with the CMV terminase has utilized cell lysates, truncated isolated domains, or more

recently full length UL89 outside of the context of the larger terminase complex (25,81-84). As a result, the CMV terminase complex remains poorly characterized and the nature of both the endonucleolytic cuts and ATPase activity produced by the CMV terminase remain to be described adequately. Furthermore, these nuclease assays here utilized agarose gels and plasmid DNA (25,84). While linearization of a plasmid containing CMV a sequences was demonstrated, the low resolution of these gels and the use of truncated proteins was not sufficient to characterize the precise biochemical nature of the endonucleolytic cutting performed by CMV (81,84,85). Additionally, the natural substrate of the CMV terminase is linear dsDNA, which suggests that results obtained from plasmid based assays, showing non-specific cleavage and degradation, may not adequately capture the biology of the CMV terminase complex (25).

This is further complicated by the fact that some authors have reported nuclease activity in both UL89 and UL56 using data obtained from impure protein preparations and without the necessary active site mutants (81-83,85). Indeed, in these studies the nature of the nuclease active sites is not discussed. Furthermore, UL56 lacks domains that are commonly associated with nuclease and ATPase activities. Thus, these data are inconclusive and significant ambiguity exists with regards to the location of the HCMV terminase nuclease and ATPase active site(s).

However, the nuclease active site of UL89 has been crystalized, and its fold resembles the nuclease active sites found in phage terminases (25,69,84). Therefore, we hypothesized that any endonucleolytic activity observed would be due to the presence of UL89.

Literature on the ATPase activities of the complex are similarly convoluted with ATPase activities being proposed for both UL89 and UL56 but only demonstrated biochemically (via polyethyleneimine based TLC assays) for UL56 (82,83,86). However, these studies worked with

truncated proteins, impure protein preparations, and in many cases lacked active site mutants complicating data analysis.

The lack of biochemical assays utilizing full length terminase represents a major barrier to antiviral development as no in-vitro biochemical system for the identification and characterization of inhibitors of the CMV terminase complex, such as the FDA approved CMV terminase inhibitor letermovir, exists. To address this, we set out to express, purify, and develop assays to characterize enzymatic activities of the full length CMV terminase complex.

1.4.4 Influenza PA Cap Snatching Endonuclease

The influenza RNA-dependent RNA polymerase (RdRp) is a heterotrimer consisting of three proteins PA, PB1, and PB2 (87,88). PA contains the "Cap-snatching" endonuclease and is required for transcription (89). Specifically, influenza lacks the machinery necessary to generate its own capped mRNA molecules (90). Therefore, it utilizes PA to hijack cellular mRNAs which are then used as primers for the generation of influenza transcripts (87). Thus, inhibition of PA prevents transcription and viral replication (91).

1.4.5 Coronavirus Proofreading Exonuclease

The RNA dependent RNA polymerases (RdRp) of coronaviruses are error prone and this is a general feature of RdRp enzymes found in RNA viruses (92). To counteract this and improve replication fidelity, coronaviruses encode a proofreading exonuclease referred to as nonstructural protein 14 (nsp14) (23,24,93-95). Indeed, it has been hypothesized that increased replication fidelity afforded by the presence of the proofreading nuclease may account for the fact that coronaviruses, and other members of the nidovirus family, have larger genomes than other RNA

viruses (96). However, the biochemical mechanisms by which proofreading occurs in these viruses is still poorly understood. Therefore, to develop a basis from which to understand coronavirus proofreading we looked at what is known about DNA polymerases with proofreading functions as these enzymes are more well studied.

1.4.6 Proofreading in DNA polymerases

DNA polymerases are involved in cellular replication and are responsible for high fidelity DNA replication (97). These polymerases possess a low error rate and the ability to correct the insertion of misincorporated bases using a 3'-5' proofreading activity (97,98). This proofreading function increases the fidelity of DNA replication by anywhere between 10 and 1000 times (97). However, the factors that trigger a polymerase to shift from a mode favoring nucleic acid extension to a mode favoring excision are still poorly understood. Recently, it has been suggested that the misincorporation of nucleotides can trigger changes in the translocation state of the polymerase leading to activation of the proofreading exonuclease (99). Specifically, this model suggested that increased exonuclease activity is observed under conditions favoring the pre-translocation state of the polymerase (99). As models for viral proofreading are lacking, we used this as a starting point for our work examining the activity of the SARS-CoV-2 proofreading exonuclease nsp14.

1.4.7 Mismatch excision and "proofreading" in coronaviruses.

To date, the excision of mismatched nucleosides and nucleoside analogues has only been directly demonstrated using recombinant SARS coronavirus nsp14 (23). nsp14 is a bifunctional protein containing both exonuclease and methyltransferase activities (28). The methyltransferase activity is involved in RNA capping which is necessary for viral transcription and evasion of the host's

innate immune defense while the exonuclease functions to increase replication fidelity (24,28,93,100).

Nsp14 has been shown to cleave both single and double-stranded RNA with a 3'-5' directionality (28,101). However, a second protein, nonstructural protein 10 (nsp10) is required for efficient activity (28). Based on structural studies, nsp10 forms a 1:1 complex with nsp14 and stabilizes the exonuclease active site of nsp14 (28). Thus, in this work, nsp14 was expressed in tandem with NSP10 and purified the resulting complexes using both affinity and size exclusion chromatography.

1.5 Why target nucleases?

As described above, many viral nucleases have highly similar active sites and mechanisms of action. This is made evident by the fact that similar inhibitors that chelate divalent metal ions present in the active site have been developed to target nucleases from several distinct viruses including HIV RNAse H, Integrase, herpesviral terminase proteins, and the influenza PA cap snatching endonuclease (25,31,36,102). Thus, nucleases are a validated drug target as the potential exists not only to develop highly potent inhibitors, but to develop inhibitors with a broad spectrum of action due to the structural similarity of these active sites. This similarity is highlighted in the attached figure. (Figure 1.5)

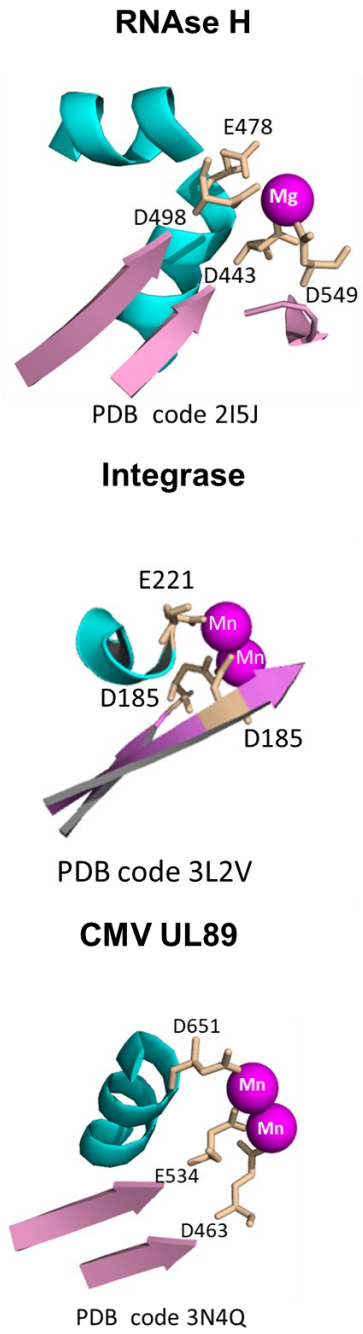


Figure 1.5 Similarity of the nuclease active sites between Integrase, RNAse H, and CMV-UL89. The nuclease active sites of HIV-1 RNAse H, SFV Integrase, and CMV UL89 with metal

ion coordinating residues are shown. Beta sheets are represented by pink arrows. Alpha helices are shown in blue. Manganese ions are shown as magenta-colored balls.

1.6 CMV, Influenza, and SARS-CoV-2

As noted previously, CMV, influenza, and SARS-CoV-2 are responsible for significant morbidity and mortality worldwide necessitating the development of novel antivirals. All of these viruses encode nuclease enzymes with related active sites providing a good basis for a drug development program. For this reason, we chose to study these viruses and both the clinical relevance of these viruses, and available inhibitors are discussed below.

1.6.1 Global burden of CMV

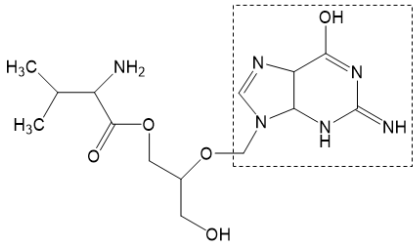
Herpesviruses are double-stranded DNA viruses that cause disease in both humans and a wide range of animal hosts (103). Following primary infection, these viruses generally establish a lifelong latent infection in lymphoid tissue or neurons (104). Infection by these agents produces a wide range of syndromes ranging from skin legions to meningitis (104). While initial infection generally results in mild disease in an immunocompetent host, if the immune status of the host changes these viruses can reactivate from latency leading to disease recrudescence (104). Among the herpesviruses, infection with human cytomegalovirus is particularly problematic and is responsible for considerable morbidity and mortality worldwide (105). Although, exposure to CMV generally leads to an asymptomatic latent infection in immunocompetent hosts, CMV causes severe disease in the immunocompromised, transplant recipients, and in newborns (105-107). This results in significant disease given that the virus is ubiquitous (108). For example, seroprevalence is ~50% in the United States and in many developing countries this approaches 100% (108,109).

This is further complicated by the fact that many of the currently used antivirals require IV administration and are plagued by the emergence of resistance, nephrotoxicity, and potential teratogenicity (105). Thus, the development of novel antiviral agents against CMV is urgently required and we hope that the work presented in this thesis will contribute to this effort.

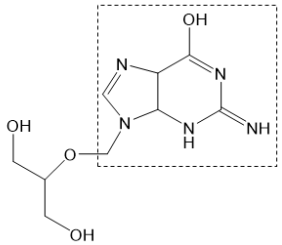
1.6.2 Commonly used anti-CMV Antivirals.

Management of CMV infection commonly relies on the use of drugs that inhibit the viral DNA-dependent DNA polymerase (105). These include the nucleoside analogues ganciclovir, valganciclovir, and the pyrophosphate analogue foscarnet (105). (Figure 1.6) Letermovir has a novel mechanism of action and was shown to target the CMV terminase complex (30). Specifically, mutations providing resistance to letermovir arose in the terminase complex, and treatment with letermovir was shown to produce empty capsids (30). Resistance conferring mutations in the terminase complex have also arisen in the clinic (110). These data suggested that letermovir targets the CMV terminase complex preventing viral genome packaging. However, the biochemistry of this inhibition has yet to be established as no enzymological assay has been established for the study of CMV terminase function.

Guanosine Analogues

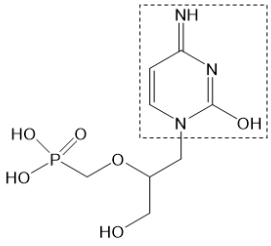


Valganciclovir



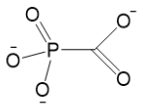
Ganciclovir

Cytosine Analogue



Cidofovir

Pyrophosphate Analogue



Foscarnet

Figure 1.6. Clinically relevant CMV inhibitors targeting the viral polymerase. The structures of commonly used CMV polymerase inhibitors are detailed. The base portion of the nucleoside analogues is outlined with a thatched line.

These antivirals, while commonly used, are associated with the development of drug resistance and have significant side effects as described above (105). This has necessitated the development of antivirals with novel mechanisms of action. One of these, letermovir, has been recently approved by the FDA for the prophylaxis of CMV disease in solid organ transplant recipients (30).

1.6.3 Global Burden of Influenza

Influenza is a severe, contagious respiratory illness caused by infection with the influenza virus (111). Influenza viruses are segmented negative-sense RNA viruses that together comprise six of the eight members of the family *Orthomyxoviridae*. Four genera of influenza, Influenza A, B, C, and D have been identified two of which, A and B, cause significant morbidity and mortality worldwide (112). For instance, the WHO reports that influenza A and B together account for up to 5 million severe illnesses and 650,000 deaths annually (113). This accounts for a total economic burden of $87 billion per year in the United States alone (112). This is despite the availability of effective vaccines against both influenza A and B (114).

1.6.4 Baloxavir and other antivirals targeting the Influenza RNA-dependent RNA polymerase (RdRp).

Antivirals represent an important tool for the treatment and management of influenza infection. Traditionally, Influenza virus infection has been treated using M2 ion channel and neuraminidase inhibitors (115). Unfortunately, in recent seasons >99% of circulating influenza strains have been resistant to M2 ion channel inhibitors limiting treatment options (116). This has necessitated the development of antivirals with novel mechanisms of action. To this end, inhibitors of the influenza RNA-dependent RNA polymerase (RdRp) (PA/PB1/PB2) have been developed.

The RNA-dependent RNA polymerase of influenza is a multifunctional heterotrimer encoding cap dependent endonuclease, polymerase, and 7-methylguanylate cap binding activities (117). These activities make the viral RdRp an attractive drug target as it is essential to viral transcription and genome replication.

Influenza performs transcription via a well described process known as "cap snatching"(87-89,117-121). Briefly, PB2 binds host mRNAs bringing them into proximity to PA which cleaves them producing short, capped primers that are subsequently extended by PB1 to produce viral transcripts. Each aspect of this process has currently been targeted for drug development and compounds targeting each component of the polymerase are currently available.

For example, the first in class PA endonuclease inhibitor baloxavir marboxil (BXM) was approved by the FDA in December of 2018 for the treatment of uncomplicated influenza in adults. A single dose of BXM results in the alleviation of influenza symptoms approximately 23-28 hours faster than placebo (122). This is likely because baloxavir is highly potent and capable of inhibiting influenza viruses in the low nM range (31). BXM is a prodrug, the active form of which, baloxavir acid (BXA), binds to divalent metal ions in the PA endonuclease active site preventing catalysis (123). Baloxavir is active against all known genera of influenza including viruses resistant to neuraminidase inhibitors (31,124). This, in addition to the fact that BXM can be administered orally as a single dose make BXM highly attractive for clinical use (122). However, soon after the discovery of BXM, resistant strains emerged. The majority of described resistance mutations are substitutions in the I38 position, with I38T/M/F resulting in a 2-50-fold increase in EC_{50} with I38T having the largest effect (122,123). It has been postulated that this reduced susceptibility is due to the fact that I38T may reduce van der Waals contacts with the inhibitor (123) . It has been reported that these mutations have a significant adverse effect on viral fitness so it is unknown how

prevalent they may become should the use of BXM become widespread (123). However, the development of widespread resistance appears possible as in clinical trials the emergence and transmission of viruses containing BXM resistance mutations has been documented (122,125).

The RNA synthesis activity of the PB1 subunit of the polymerase has been targeted by the use of the nucleotide analogue favipiravir which is approved in Japan for the treatment of pandemic influenza (126). Favipiravir has been evaluated in clinical trials and while the exact mechanism of action requires clarification, it is thought to function through the induction of lethal mutagenesis (127). Specifically, it is thought that incorporation of favipiravir leads to an increase in G to A and C to T transversion mutations (127). However, resistance to favipiravir has been reported. Specifically, the PB1 K229R mutation located within the conserved F motif of the RdRp has been shown to confer resistance to favipiravir at the cost of viral fitness (128). However, this loss of fitness is restored by the PA P653L mutation (128).

Several unpublished clinical studies on the efficacy of favipiravir for the treatment of uncomplicated influenza have been performed with somewhat variable results (129). For example, studies have reported anywhere between a 6 and 14 hour decrease in time to influenza symptom resolution versus placebo (129). This moderate effect when combined with the fact that teratogenicity has been observed in multiple animal models at doses similar to those used in treatment limit the utility of favipiravir in the treatment of uncomplicated influenza (129). Despite these limitations, favipiravir still possesses a high degree of utility for the treatment of pandemic influenza.

The PB2, or cap binding subunit of the influenza polymerase, has been targeted by pimodivir (130). Pimodivir binds to the cap binding domain of PB2 at nanomolar concentrations preventing

the binding of capped host mRNAs. This, in turn, prevents "cap-snatching" and viral transcription (131). However, pimodivir is only active against Influenza A viruses likely owing to significant amino acid sequence divergence in the cap binding domain of PB2 between different genera of influenza (131). Additionally, resistance to pimodivir has been reported with the PB2 Q306H, S324I, S324N, S324R, F404Y, and N510T mutations conferring a greater than 10-fold decrease in susceptibility (132). Furthermore, resistant strains of influenza containing PB2 substitutions have also emerged in clinical trials examining the efficacy of pimodivir (129,130). The most common of these is the PB2 M431I variant which has been shown to confer a 57-fold reduction in pimodivir susceptibility in cell culture (129,130). However, interestingly, in patients treated with a combination of a neuraminidase inhibitor (Tamiflu) and pimodivir the emergence of resistant strains was greatly reduced (129,133).

Taken together, the Influenza RdRp represents a promising, validated target for the development of antiviral drugs.

1.6.5 Global Burden of SARS-CoV-2

Coronaviruses are single-stranded positive sense RNA viruses that are responsible for disease in humans and animals (134,135). Members of this family are known to cause a variety of pathologies ranging from mild cold-like illness to severe respiratory syndromes such as those resulting from infection with SARS-CoV (severe acute respiratory syndrome coronavirus), MERS-CoV (middle east respiratory syndrome coronavirus), and SARS-CoV-2 (severe acute respiratory syndrome two coronavirus), the causative agent of COVID-19 (135,136). In late 2019, SARS-CoV-2 emerged in Wuhan, China (135,137,138). Current evidence suggests that the virus was of zoonotic origin and that it crossed the species barrier from bats into humans (139). The virus then spread around the

world leading the WHO to declare a global pandemic (138,140). To date, SARS-CoV-2 has infected 160 million people leading to approximately 3.3 million deaths worldwide. This is because SARS-CoV-2 is highly transmissible and can cause severe respiratory disease in vulnerable populations, i.e., the elderly and those with preexisting conditions (138). Thus, there is an urgent need to develop antivirals against SARS-CoV-2.

1.6.6 Remdesivir and SARS-CoV-2

To date, the only small molecule inhibitor approved by the FDA for the treatment of infection with SARS-CoV-2 is remdesivir. Remdesivir is a nucleoside analogue inhibitor of the SARS-CoV-2 polymerase nsp12 (134,141). Remdesivir is an ATP analogue that was first developed by Gilead Pharmaceuticals for the treatment of Ebola virus infection (142) and was shown to function via inhibition of the viral RNA-dependent RNA polymerase (134,142,143). Before the emergence of SARS-CoV-2 remdesivir had also been shown to be active against coronaviruses in a laboratory setting (144). Thus, when SARS-CoV-2 emerged there was interest in testing remdesivir against the virus and it was ultimately shown to inhibit SARS-CoV-2 replication via inhibition of the viral polymerase nsp12 (134,141,145). This inhibition of the viral polymerase occurs by a twofold mechanism (141,145). Firstly incorporation of remdesivir was shown to result in chain termination at a position three nucleotides downstream of the point of incorporation (141). This is referred to as “delayed chain termination”(141). This inhibition can be overcome at high nucleotide concentrations leading to the incorporation of remdesivir into the viral genome (145). However, the incorporation of nucleotides across from RDV in the template strand leads to polymerase stalling and this represents a second mechanism by which remdesivir functions to inhibit SARS-CoV-2 polymerase (145).

In clinical trials, treatment with remdesivir was shown to reduce time to recovery by 31% compared to placebo and based on this data the compound was approved by the FDA (134). While many antiviral therapies have shown promise, at the time of writing, remdesivir is the only FDA-approved small-molecule inhibitor for the treatment of COVID-19. Therefore, there is great interest in characterizing mechanisms of resistance to remdesivir.

One possible resistance mechanism is presented by the presence of the proofreading nuclease nsp14. As nsp14 has previously been shown to reduce the effectiveness of nucleotide analogue antivirals in coronaviruses related to SARS-CoV-2 we hypothesized that this could also occur in SARS-CoV-2 and we devised biochemical assays to study this process (23,144).

1.7 Research Project Aims

1.7.1 Chapter 3: Human Cytomegalovirus Terminase

The functions of the human cytomegalovirus terminase are understudied and as of writing, there are no established systems that would allow for the study of terminase function and inhibition. This is despite the recent FDA approval of the antiviral letermovir which targets the terminase. Furthermore, sufficient evidence does not exist to allow for the unambiguous assignment of nuclease and ATPase activities to each of the terminase components. To this end, we aimed to express full-length CMV terminase and develop biochemical systems for the characterization of the endonuclease, ATPase, and nucleic acid binding activities of the complex with the goal of providing a system for the elucidation of mechanisms of antiviral action and resistance.

By understanding mechanisms surrounding antiviral action and resistance we hope to both inform the development of novel antiviral compounds targeting this enzyme complex and provide a system to facilitate the discovery of these compounds *in vitro*.

1.7.2 Chapter 4: Influenza B PA Endonuclease

The influenza RNA-dependent RNA polymerase (RdRp) is a heterotrimer that consists of three subunits, PA, PB1, and PB2. As previously described, PA is the target of the recently approved antiviral drug baloxavir. Studies exploring the mechanism of action of baloxavir and its exceptionally high (low nM IC_{50}) potency were lacking. Therefore, we expressed and purified the RdRp and developed a cap-snatching endonuclease assay to examine mechanisms of baloxavir function and resistance. Here we show that baloxavir functions as a tight slow binding competitive inhibitor and show that a commonly observed resistance mutation diminishes tight binding. Furthermore, our work has led to the establishment of a biochemical system for the identification of novel inhibitors of the influenza "cap snatching" endonuclease and provided a framework by which tight binding inhibitors can be analyzed.

1.7.3 Chapter 5: SARS CoV-2 Exonuclease

Coronaviruses encode a proofreading exonuclease that increases replication fidelity (24,93,95). However, a consequence of encoding an exonuclease is that the excision of nucleoside analogue antiviral drugs is possible (23). This has been demonstrated for related coronaviruses but not SARS-CoV-2 (23).

Thus, there are several fundamental and applied reasons to study the SARS CoV-2 exonuclease including our need to better understand proofreading in the context of coronaviruses as well as the

potential to better understand mechanisms of resistance to approved nucleoside analogue inhibitors such as remdesivir. To this end we developed assays to examine the excision of incorporated nucleotide analogue inhibitors by nsp14. Furthermore, these assays are useful in that they allow for the identification of nucleoside analogues that are resistant to excision by nsp14. These analogues, provided they are incorporated efficiently by the viral RdRp, should be potent inhibitors of SARS-CoV-2 replication.

Additionally, we established a system that can be used for the development and refinement of inhibitors of the SARS-CoV-2 ExoN. As the ExoN has been shown to be critical for the replication of SARS-CoV-2 it is possible that this system could be used for the identification of novel inhibitors of SARS-CoV-2 replication that target nsp14. Such inhibitors could also, in theory, improve the efficacy of nucleoside analogue inhibitors that are not effective because they are efficiently excised by nsp14.

Chapter 2: Materials and Methods

2.1 Methods utilized in chapter 3: Characterization of the CMV terminase complex UL51/Ul56/Ul89.

2.1.1 Protein Expression and Purification

The constructs for the expression of WT, Endonuclease mutant, and ATPase mutant HHV5 terminase were generated by DNA synthesis and purchased from Genscript (Piscataway, NJ, USA). All other constructs were generated via polymerase chain reaction (PCR). Recombinant CMV Terminase proteins (UL89/UL56/UL51) and all variants were expressed using an insect baculovirus-based expression system according to established protocols (146). In short, the pFastBac-1 plasmid (Invitrogen, Burlington, Ontario, Canada) containing codon-optimized

synthetic DNA sequences encoding HCMV UL89, Ul56, and UL51 derived from HCMV strain AD169 were used for protein expression in insect cells. Purification was accomplished by NiNTA affinity chromatography making use of an N-Terminal 8-histidine tag on UL56. Specifically, the insect cell pellet containing the protein of interest was mixed with one volume of cell lysis buffer. To this, one Roche Ultra Complete EDTA Free Protease Inhibitor tablet was added. The resulting solution was then vortexed until the pellet was completely resuspended in lysis buffer. The resulting solution was then sonicated to ensure complete cell lysis. Following sonication, the resulting solution was centrifuged for 30 minutes at 30000 x g using a JA-25.5 rotor in a Beckman Avanti J-20 floor centrifuge. The resulting supernatant was transferred to a column containing 500 µL of Thermofisher His-Pur™ NiNTA resin in equilibration buffer. The supernatant and resin were then allowed to incubate with rotation at 4 °C for one hour to allow time for the terminase complex to bind to the resin. The resin is then washed with 60 column volumes of wash buffer to remove nonspecifically bound proteins (Table 2.1). The terminase complex is then eluted in 500 µL fractions using elution buffer (Table 2.1). This is continued until eluted protein is no longer detected via Bradford (i.e., the Bradford reagent no longer changes color). The fractions are then pooled and concentrated using Amicon Ultra centrifugal filters. The composition of the finalized protein preparations was assessed by SDS PAGE followed by staining with Coomassie brilliant blue G-250. The identity of the purified proteins was confirmed using LC-MS/MS (appendices 1).

2.1.2 Nucleic Acids

All DNA oligos used in chapter 3 were purchased from Integrated DNA Technologies (Coralville, IA, USA).

Table 2.1 Buffers used in the NiNTA purification of the HCMV terminase complex UL51/UL56/UL89.

Buffer Name	Buffer Composition
Cell Lysis Buffer	100mM Tris-HCl pH 7.5 200mM NaCl 0.2% Tween 20 10% glycerol 8mM TCEP
Wash Buffer	50 mM Tris-HCl pH 7.5 200 mM NaCl 25 mM Imidazole 10% Glycerol 4mM TCEP
Column Equilibration Buffer	50 mM Tris-HCl pH 7.5 200mM NaCl 4mM TCEP 0.1% Tween-20 10% glycerol
Elution Buffer	50 mM Tris-HCl pH 7.5 200 mM NaCl 200 mM Imidazole 10% Glycerol 8 mM TCEP 0.01 % Tween-20

2.1.3 Fluorescent Labeling of DNA oligos

Oligos were ordered containing a 5' amino group attached to a six-carbon linker. These oligos were then reacted with Cy5-NHS ester purchased from Fisher Scientific (Edmonton, AB, Canada) under standard conditions to generate 5' Cy5 labeled oligos. Excess dye and remaining unlabeled oligo were removed via PAGE purification.

2.1.4 5' Radiolabeling of DNA oligonucleotides

Radiolabeling was accomplished using T4-polynucleotide (T4-PNK) kinase according to standard protocols. T4-PNK was purchased from Thermofisher scientific (Edmonton, AB, Canada), and [α-32-P] GTP was purchased from Perkin Elmer (Boston, MA, USA).

2.1.5 Terminase endonuclease assay

1 μM HHV5 terminase, 25 mM NaCl, 30 mM Tris-HLC pH 8, and 100 nM of fluorescent or radiolabeled oligo was mixed and allowed to equilibrate for 5 minutes at 37 °C. Reactions were then initiated with 1 mM of divalent metal ions ($MgCl_2$ or $MnCl_2$) and quenched at the specified timepoints with 1 volume of 0.5 M EDTA/formamide solution.

The resulting products were then run on a 20% TBE-PAGE gel and visualized using a GE-Typhoon® variable mode imager.

2.1.6 Terminase ATPase assay

1 uM of HHV5 T3 30 mM tris pH 8.0, 25 mM NaCl, and 200 nM of [α-32-P] GTP were mixed and allowed to equilibrate for 5 minutes at 37 °C. The reactions were then initiated via the addition of 1 mM $MnCl_2$. After the desired amount of time had elapsed the reactions were quenched via the addition of 1 volume of 0.5M EDTA. The resulting products were then spotted on polyethylene imine TLC plates and resolved. Specifically, 0.5 M LiCl, 0.5 M Formic Acid was used as running buffer and the buffer was allowed to run ¾ of the way up the plate. The migration patterns of the reaction products were then visualized by autoradiography.

2.1.7 Terminase DNA binding Assay

Various concentrations of WT or Variant HHV5 terminase complex were mixed with 30 mM tris HCL, 25 mM NaCl, and 100 nM of Cy5 labeled 24 nucleotide (nt) DNA substrate. The sequence of this DNA substrate is as follows: Forward strand: 5' Cy5-GCGACGGCGAATAAAAGCGACGTG3'. Reverse strand: 5' Cy5-GCGACGGCGAATAAAAGCGACGTG 3'. Following incubation at room temperature for 5 minutes the fluorescence anisotropy of the substrate was measured using a Molecular Devices SpectraMax™ M5 fluorescent plate reader.

2.2 Methods utilized in chapter 4: Characterization and Inhibition of the Influenza PA endonuclease.

2.2.1 Enzymes and nucleic acids

Wild-Type (WT) and variant FluB-ht were expressed and purified as described previously (Strain: B/Memphis/13/03) (27,142). All FluB-ht variants were generated by DNA synthesis and purchased from Genscript (Piscataway, NJ, USA). The NCBI accession numbers of the source protein sequences are as follows: PA AAU94844, PB1 AAU94857, and PB2 AAU94870. The 5'-triphosphorylated RNAs (5'-pppRNA) used to generate the 5' m^7G capped oligos were designed according to Reich *et al*., 2020 (27). 5'-pppRNAs of up to 20 nucleotides (nts) were used to generate capped oligos and purchased from Chemgenes (Wilmington, MA, USA) while a 39-nt RNA without triphosphorylation at the 5' end corresponding to the influenza B promoter consensus sequence (vRNA) was purchased from Dharmacon (Lafayette, CO, USA) (27). The sequence of the RNA used to generate the longest (20nt) capped oligo is 5' ppp-AAUCUAUAAUAGCAUUAUCC 3'. The 14-nt, 12-nt and 10-nt oligos are shorter versions of this 20-nt oligo that have been truncated at the 3' end. Their sequences are as follows: 14-nt oligo, 5' ppp-AAUCUAUAAUAGCA 3', 12-nt oligo, 5' ppp-AAUCUAUAAUAG 3', 10-nt oligo 5' ppp-AAUCUAUAAU 3'. The capped 20-nt oligo was used as a substrate for the endonuclease reactions, while an equimolar mix of capped 20-nt, 14-nt,12-nt, and 10-nt, oligos were used as molecular weight markers. The sequence of the vRNA is as follows: 5'AGUAGUAACAAGAGGGUAUUGUAUACCUCUGCUUCUGCU 3'. Capping and radiolabeling of RNA oligos were performed using the New England Biolabs Vaccinia Capping System (Fisher Scientific, Edmonton, Alberta, Canada). Briefly, [α-^{32}P] GTP (Perkin Elmer,

Boston, MA, USA), RNA markers, and vaccinia capping enzyme were mixed, and the reactions were allowed to proceed for 30 minutes at 37 °C. Reactions were then incubated for 10 minutes at 95 °C to inactivate the capping enzyme. Capped, radiolabeled RNA oligos were then further purified using GE healthcare microspin G-25 columns (Chicago, IL, USA) followed by phenol-chloroform extraction.

2.2.2 Baloxavir acid (BXA)

BXA was purchased from MedChemExpress (Monmouth Junction, NJ, USA) and resuspended at a concentration of 10 mM in 100% DMSO.

2.2.3 FluB-ht PA endonuclease assay

The influenza nuclease assay was performed as described previously by Reich et al., 2014 and the concentrations of WT and variant FluB-ht were modified (27). Briefly, 55 nM of WT or variant FluB-ht was incubated at 30 °C with 1.7 μM vRNA in 30 mM Tris-HCl pH 7.5, 25 mM NaCl, and 5 mM $MgCl_2$ (buffer A). 25 mM NaCl was chosen based on a NaCl optimization experiment (Fig. 4.11). The reactions were then initiated with 100 nM of 20-nt substrate. At the indicated time points, aliquots of the reaction master mix were quenched with formamide and 25 mM EDTA. Reactions were heat-inactivated at 95 C° for 10 minutes and resolved on a 20% denaturing polyacrylamide gel. The reaction products were then visualized by phosphorimaging. Product fractions were determined using QuantityOne software (Biorad). Product fraction refers to the ratio of the signal produced by the 5'-capped 11- and 12-nt products to the sum of these products plus the remaining substrate. The 14-nt product was not quantified as its contribution to the overall

signal was negligible. The resulting data were plotted using GraphPad Prism 8 (Graphpad, San Diego, California). (44).

2.2.4 IC_{50} determination for inhibition of the FluB-ht PA endonuclease by BXA

For IC_{50} determinations, various concentrations of FluB-ht WT or FluB-ht PA I38T were incubated with 1.7 μM vRNA and increasing concentrations of BXA at 30 °C in buffer A and 5% DMSO for five minutes. Reactions were then initiated with the 20-nt capped RNA substrate. Reactions were quenched at 12 minutes post initiation and the products resolved and visualized as described above. The resulting data was quantified as described for the endonuclease assay, normalized based on activity in the absence of inhibitor, graphed using Prism 8 (Graphpad, San Diego, CA, USA), and IC50 values determined. Percent error was determined as follows: % error = (Standard Deviation / Mean IC50) *100%

2.2.5 Time Dependent Inhibition Assay

55 nM of FluB-ht WT or FluB-ht PA I38T was incubated with 75 nM BXA, 5% DMSO, and 1.7 μM vRNA in buffer A at 30 °C for 3, 6, 9, 12, 15, 18, 21, 24, or 27 minutes. Following this, the reactions were initiated via the addition of 20-nt capped, radiolabeled, RNA substrate and allowed to proceed for 8 minutes. At this time, the reactions were quenched, resolved, and visualized as described above.

2.2.6 Order of addition assays

55 nM FluB-ht WT or FluB-ht PA I38T and 100 nM capped 20-nt substrate were mixed as specified with and without 75 nM BXA in buffer A with 5% DMSO and incubated at 30 °C for

one hour to allow for complex formation. The reactions were then initiated and allowed to proceed for 1, 2, 4, 8, 16, 32, or 120 minutes. The products were then separated via 20% polyacrylamide denaturing gel electrophoresis (PAGE) in 1x TBE buffer and visualized by phosphorimaging.

2.2.7 Transcription Inhibition Assays

In brief, a 20-nt capped radiolabeled substrate was incubated with FluB-ht and $MgCl_2$ in the absence of nucleotides to generate a 12-nt capped primer. Or, lieu of this step , FluB-ht was incubated with synthetic 12-nt capped primer. Following this, where specified, BXA was added to the reaction mixture and incubated for 10 minutes. Nucleotides were then added, and the reaction quenched at various timepoints. The products of the reaction where then resolved via 20% PAGE and autoradiography.

2.3 Methods utilized in chapter 5: characterization of the SARS-CoV-2 proofreading Exonuclease complex nsp14/nsp10.

2.3.1 Nucleic acids and chemicals

All RNA oligos used in the study were 5'-phosphorylated and purchased from Horizon Discovery Group (Lafayette, CO, USA). RDV-TP was provided by Gilead Sciences (Foster City, CA). NTPs were purchased from Fisher Scientific (Edmonton, AB, Canada). Favipiravir TP was purchased from Toronto Research Chemicals (North York, Ontario, Canada).

2.3.2 Protein Expression and Purification

All protein expression constructs used in the study were generated by DNA synthesis and purchased from Genscript (Piscataway, NJ, USA). Recombinant SARS-CoV-2 RdRp

(nsp7/nsp8/nsp12) was expressed and purified as previously described (141). This work was carried out by Emma Woolner and Dana Kocincova. SARS-CoV-2 proofreading exonuclease complexes (nsp10/nsp14) were expressed using an insect baculovirus based expression system according to established protocols (146). In short, the pFastBac-1 plasmid (Invitrogen, Burlington, Ontario, Canada) containing codon-optimized synthetic DNA sequences encoding NSP-5, 10, and 14 derived from the SARS-CoV-2 1ab sequence (NCBI: QHD43415.1) was used for protein expression in insect cells. Purification was accomplished by NiNTA affinity chromatography making use of a N-Terminal 8-histidine tag on NSP 14. NiNTA affinity chromatography was followed by gel filtration chromatography on a GE Superdex^{TM} 200 increase 10/300 GL column to yield final purified protein (Figure 1). Composition of the finalized protein preparations was assessed by SDS PAGE followed by staining with Coomassie brilliant blue G-250. The identity of the purified proteins was confirmed using LC MS/MS (appendices 1).

2.3.3 Sars-Cov-2 exonuclease assay

All assays except for the assay utilizing templated remdesivir were performed as follows: Model primer/template, 140 nM of sars-CoV-2 RdRp, 0.1 μM [α-^{32}P] GTP, 0.1 μM of NTPs, 200 μM primer, 1 uM template, and 5mM $MgCl_2$ were mixed and allowed to incubate for 30 minutes at 30 $^{\circ}$C to allow the RdRp complex to generate the desired dsRNA product. These products were then mixed with the SARS-CoV-2 ExoN complex nsp14/nsp10 and allowed to incubate for 15 minutes in the case of an enzyme titration or, in the case of a time course assay, varying lengths of time before being quenched with formamide and 25 mM EDTA. The resulting pattern(s) of ExoN activity were then resolved via TBE-PAGE electrophoresis and autoradiography. The results were quantified by taking the ratio of the product of interest which was either a 6-nt product or a 12-nt

product to the total signal in the lane and multiplying by 100%. The assay utilizing templated RDV required 1 μM of RNA template and 1 uM of Sars-CoV-2 RdRp. Lower concentrations failed to produce adequate signal, so these conditions were used. Templates containing RDV were generated and generously provided by Dr. E.P. Tchesnokov (145).

Chapter 3: Characterization of the CMV Terminase Complex UL51/Ul56/Ul89

3.1 Introduction

Cytomegalovirus (CMV) is a double-stranded DNA (dsDNA) virus and a member of the Herpesviridae (147). Infection with CMV generally leads to an asymptomatic latent infection in immunocompetent individuals (147). However, when the immune status of the host is compromised or modified such as in pregnancy or as a result of immunosuppressive treatments, such as those given to solid organ transplant recipients, CMV infection can lead to significant pathology. Given that the seroprevalence of CMV generally ranges from 60-90% worldwide depending on the country sampled, this results is significant morbidity and mortality (147). For example, CMV infection is a leading cause of congenital hearing loss, neurodevelopmental abnormalities, and graft loss among solid organ transplant recipients (147-149). Currently available antivirals, including foscarnet, cidofovir, and ganciclovir, have limited clinical utility due to concerns about renal toxicity, teratogenicity, and the emergence of drug resistance (147,148,150,151). To address this problem, the development of antivirals with novel mechanisms

of action and targets is required. One such target is the CMV terminase complex (T3). This complex is a tripartite protein complex consisting of UL51, UL56, and UL89 (66,67). The terminase complex possesses both NTPase and endonuclease activities which are required for the packaging of the viral genome into the capsid (30). Thus, inhibiting its activity blocks an essential step in viral replication and there has been significant interest in targeting the complex (30). Furthermore, as the herpesvirus large terminase (UL56 in CMV) is highly conserved between all herpesviruses it is hypothesized that the emergence of resistance mutations in terminase will be unlikely (30). This is because, mutations in highly conserved proteins generally result in a loss of viral fitness.

Recently, one such antiviral, letermovir, gained FDA approval (30,151). This was the first FDA approved anti-CVM compound that targets something other than the CMV DNA dependent DNA polymerase (DdDp) (30,151). Specifically, letermovir treatment leads to the production of empty capsids and southern blot analysis showed that it inhibited the endonucleolytic cleavage of concatemeric CMV genomes (30). Serially passaging CVM in the presence of letermovir produced resistant virus with amino acid substitutions in UL56 (30). This suggested that the CMV terminase complex may be the target of letermovir. However, mechanistic studies on the biochemistry and inhibition of the CMV terminase complex are lacking. To address this, in this chapter we expressed and purified the CMV terminase complex and developed assays to monitor its nuclease, ATPase, and double stranded DNA (dsDNA) binding activities.

3.2 Results

3.2.1 Expression and Purification of the CMV Terminase complex

To express the CMV terminase complex we made use of an insect cell (SF9 cell) based protein expression system that has been utilized by our group and others to produce large protein complexes (141,142,146,152-154). This allows one to bypass the solubility issues and size limitations encountered in other commonly used protein expression systems such as those utilizing *E. coli (146).* Furthermore, the use of a eukaryotic protein expression system allows for the maintenance of post-translational modifications, thus results obtained from proteins produced in this way are likely more physiologically relevant (146,155).

We designed a construct that would express terminase as a polyprotein in conjunction with Tobacco Etch Virus (TEV) protease. The construct was designed such that UL51, UL56, UL89, and TEV protease would be expressed as a long polyprotein with a TEV site between each of the individual components (Figure 3.1 A). The idea being that this would facilitate auto-cleavage of the polyprotein into the individual components. These components would then associate with one another and we would pull down the complex using affinity chromatography (Figure 3.1A). To this end, an 8 x histidine tag was attached to the N-terminus of UL56 and the complex was purified using Ni-NTA affinity chromatography (Figure 3.1B). A streptavidin tag was also included but was not used as we were unable to successfully purify this complex using the tag.

Following purification of the complex, we performed size exclusion chromatography to determine the relative size of the complex (Figure 3.1C). In the presence of 4 mM tris(2-carboxyethyl) phosphine TCEP (a reducing agent) we observed a single major peak corresponding to an ~240

KDA complex (Figure 3.1C). Conversely in absence of TCEP multiple higher order peaks were observed (Figure 3.1C). As TCEP reduces disulphide bonds these larger peaks likely represent aggregation due to cysteine crosslinking. This represents an approximately 1:1:1 complex of UL56:UL89 and Ul51 and suggested that we had successfully purified the terminase complex.

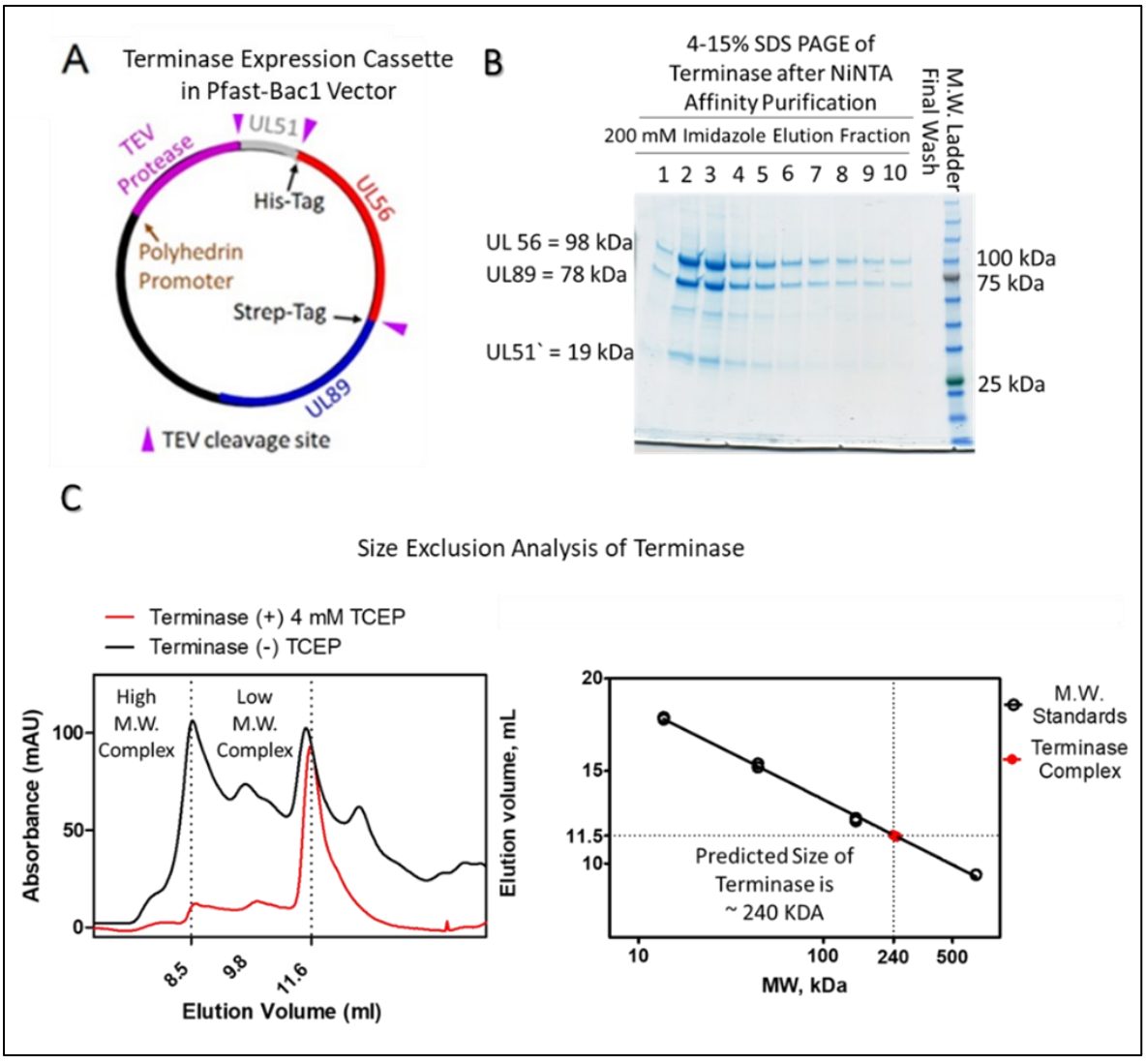


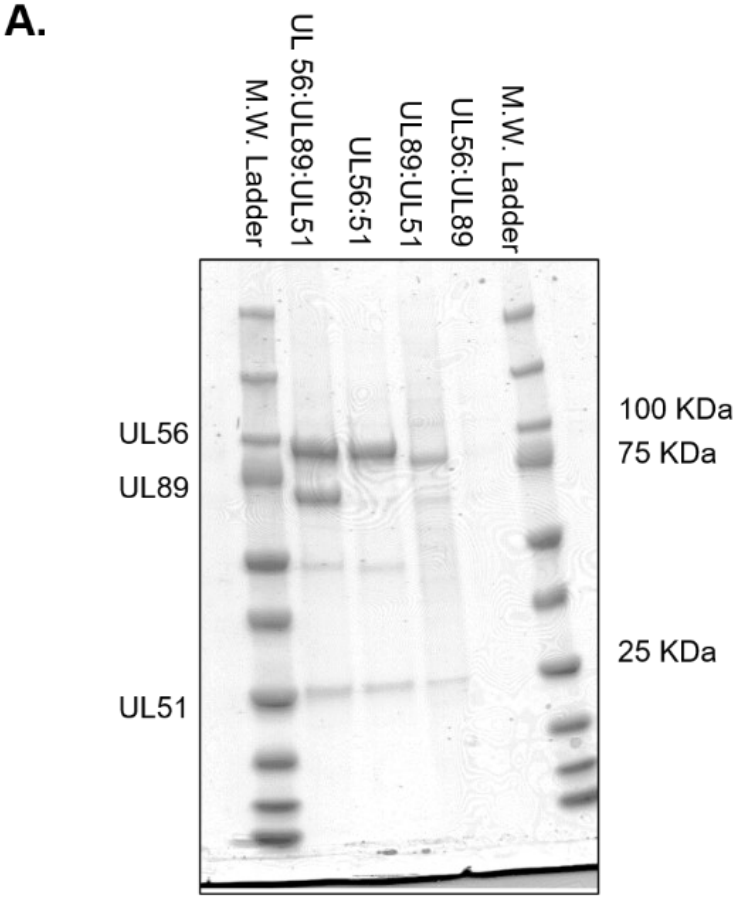
Figure 3.1 Expression, NiNTA purification, and size exclusion chromatography (SEC) analysis of the human cytomegalovirus terminase complex UL51/UL56/UL89 (T3). (A) Schematic diagram of the T3 expression cassette. T3 is expressed as a polyprotein containing TEV

protease and cleavage sites between each of the constituents. The polyprotein then self-cleaves and can be purified via affinity chromatography. (B) 4-15% SDS page illustrating the results of a NiNTA affinity purification. (C) Size-Exclusion chromatography analysis of the purified T3 complex. Dr. Hsiao-Wei Liu assisted in performing the SEC analysis.

3.2.2 Expression Requirements of the Terminase Complex

To examine the expression requirements of the terminase complex I used the polymerase chain reaction (PCR) to truncate the plasmid shown in figure 3.1A to generate constructs containing UL51+UL56, UL51 + UL89, and UL56 + UL89. For constructs lacking UL56 an 8x poly-histidine tag was attached to the C-terminus of UL51 for use in purification. UL56 was successfully expressed in the absence of UL89 and the presence of UL51. Conversely, constructs containing UL89 + UL51 or UL56 and UL89 but lacking UL51 failed to express. These data suggest that UL51 is necessary for the expression of the CMV terminase complex and that the presence of UL56 and UL51 is required for the expression of UL89 (Figure 3.2). This is in line with published data suggesting that UL51 is required for the optimal expression of the terminase complex and UL56 specifically (67).

Furthermore, previous biochemical and structural studies of the CMV terminase have been limited by difficulties in producing adequate quantities of the proteins comprising the CMV terminase complex(84). Thus, by defining optimal expression conditions for terminase these studies may facilitate future biochemical and structural studies involving the CMV terminase complex.



B.

Construct	Expression
UL56:89:51	Yes
UL:56:51	Yes
51+89	No
56+89	No

Figure 3.2 CMV Terminase expression requirements. (A) 4-15% SDS page showing the results of the expression and NiNTA purification of several T3 variant constructs. (B) UL51 and UL56 are required for the optimal expression of the T3 complex.

3.2.3 Nuclease Activity of the Terminase Complex (UL51/UL56/UL89)

The terminase complex of CMV is known to possess a nuclease activity that is involved in DNA packaging (25). Previous studies have shown that a c-terminal fragment of UL89 contains an endonuclease active site that is active in the presence of divalent metal ions (25). To study the nuclease activity of the terminase complex we designed an assay that utilizes small labeled (either radiolabeled or fluorescently labeled as indicated) substrates and TBE page to monitor the nuclease activity of the complex.

To do this, we chose to work with a system utilizing small fluorescently or radio-labeled substrates (Figure 3.3). Although, one could argue that it would be more biologically relevant to work with large substrates that resemble the packaging sequences found in a herpesviral concatemer, these substrates were chosen because they are inexpensive and easily amenable to the development of biochemical assays. Furthermore, previous work utilizing a truncated peptide containing the nuclease active site of ul89 cleaved plasmid DNA without any sequence specificity. Given that we observed what appeared to be a sequence specific endonucleolytic cleavage, we chose to continue working with this substrate and related variants throughout the work presented in this thesis (Figure 3.3 and Figure 3.5).

In the presence of $MnCl_2$ we observed that our 24 nucleotide (nt) substrate was cut endonucleolytically at a position 19-nt from the 5' label (Figure 3.3). Analysis of the reaction kinetics revealed that this cut is produced early on, and therefore represents an initial endonucleolytic cut (Figure 3.3). Conversely, when the same substrate is labeled on the 3' end a 5-nt product is produced corresponding to cleavage at the same position (Figure 3.3).

Following this initial endonucleolytic cut, additional cuts are generated (Figure 3.3 A). These cuts are likely exonucleolytic in nature as the pattern of laddering on the gel suggests that this cutting is most likely occurring as a result of cleavage at the 3' end of the nucleic acid strand. However, to unambiguously verify that these subsequence cuts are exonucleolytic it would be necessary to order a 19-nt substrate where the 3' end is protected from degradation. For example, this could be via ordering a substrate lacking a 3'OH group or one with a bulky modification attached to the 3' position. As many 3'-5' exonucleases, have been shown to require the 3'OH group for activity, these substrates would likely remain uncut by the terminase (101). This would confirm that these subsequent cuts are exonucleolytic. Conversely, if these cuts were still observed these data would suggest that there are multiple endonucleolytic sites and the one at position 19 is favored.

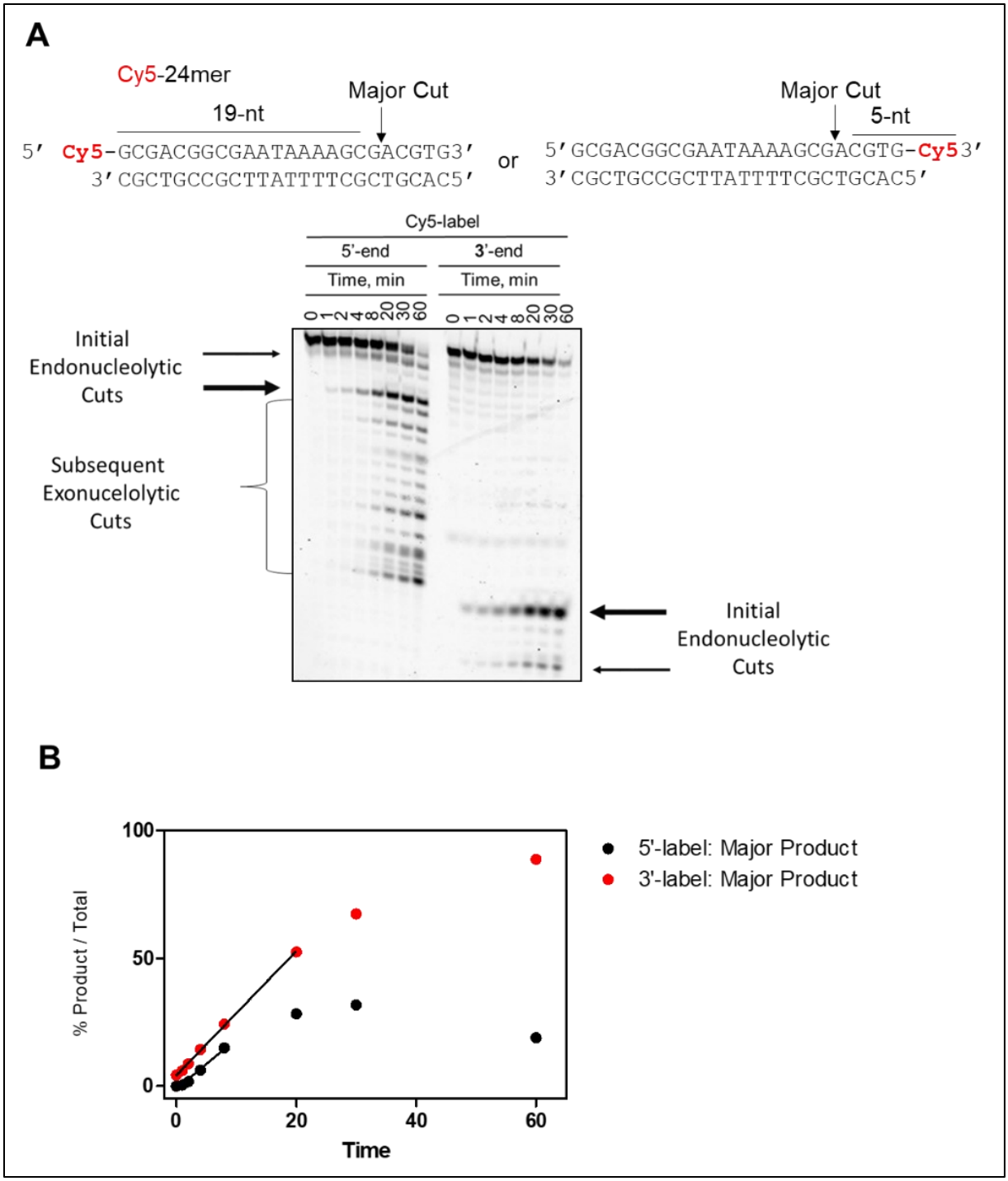


Figure 3.3 Nuclease activity of the CMV terminase complex UL51/UL56/UL89. (A) PAGE showing migration patterns of the products of terminase nuclease activity. A 24 nucleotide (nt) substrate is cleaved to yield both a 19-nt and 5-nt product. Both are visible depending on whether

the 5’ (left) or 3’ (right) end of the substrate is labeled. (B) Graphical representation of the results presented in A. Linear portions of the reaction curve are shown to illustrate that the slope for these two reactions is similar suggesting that these two products are the result of the same cleavage event.

3.2.4 Mutation or deletion of UL89 fails to eliminate the nuclease activity of the CMV terminase complex.

To determine that the observed activity of the CMV terminase complex is due to the protein complex of interest and not a copurified contaminant we generated a construct where the three residues involved in metal ion co-ordination by UL89, D463, E534, and D651 were mutated to alanine (25) (Figure 3.4B). Following NiNTA purification, endonuclease activity similar to that observed for wt enzyme was observed suggesting that either the observed activity is not due to UL89 or the activity is due to a copurified contaminant (Figure 3.4B). Furthermore, cleavage of the substrate was observed for both wild-type and the endonuclease mutant enzyme in the presence of either $MnCl_2$ or $MgCl_2$.

In line with this observation, constructs containing UL51 and UL56 but lacking UL89 produced a cleavage pattern almost identical to those observed in the presence of UL89 suggesting that the observed endonuclease activities do not depend on UL89 (Figure 3.4). This was a highly puzzling observation as UL56 lacks any known ATPase or nuclease domains and structural studies to date have suggested that the nuclease domain is present in UL89 based on both homology to phage large terminase enzymes and biochemical data. Furthermore, the structure of the HSV-1 homologue of UL56, UL28, lacks any folds that are known to be associated with nuclease or ATPase activities (76).

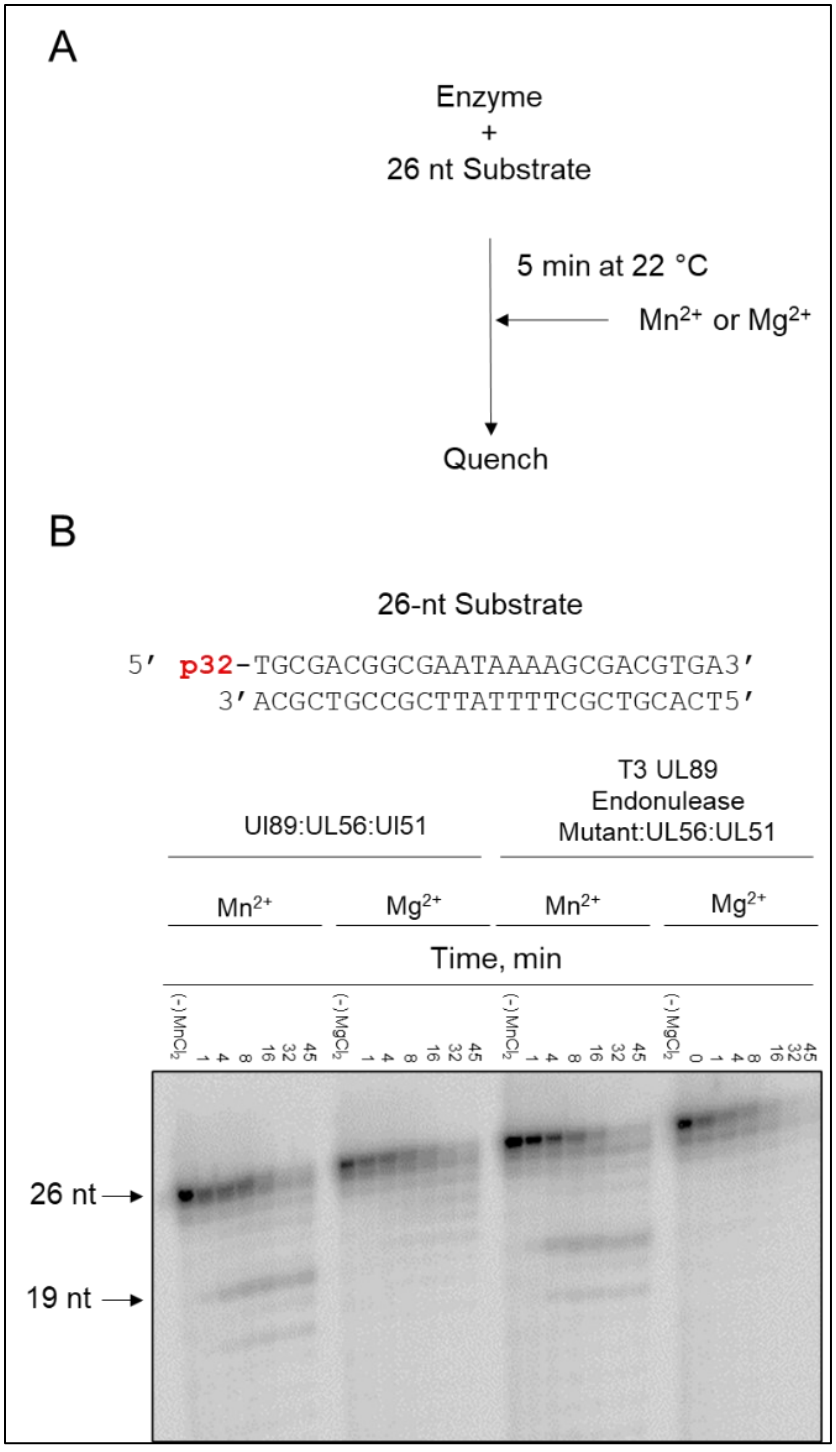


Figure 3.4 Nuclease activity of the wild type (wt) and endonuclease mutant CMV terminase complex UL51/UL56/UL89. (A) Schematic detailing the reaction setup. (B) PAGE showing the

migration patterns of the products of the wt and UL89 endonuclease mutant terminase preparations in the presence of either 1 mM $MnCl_2$ or $MgCl_2$.

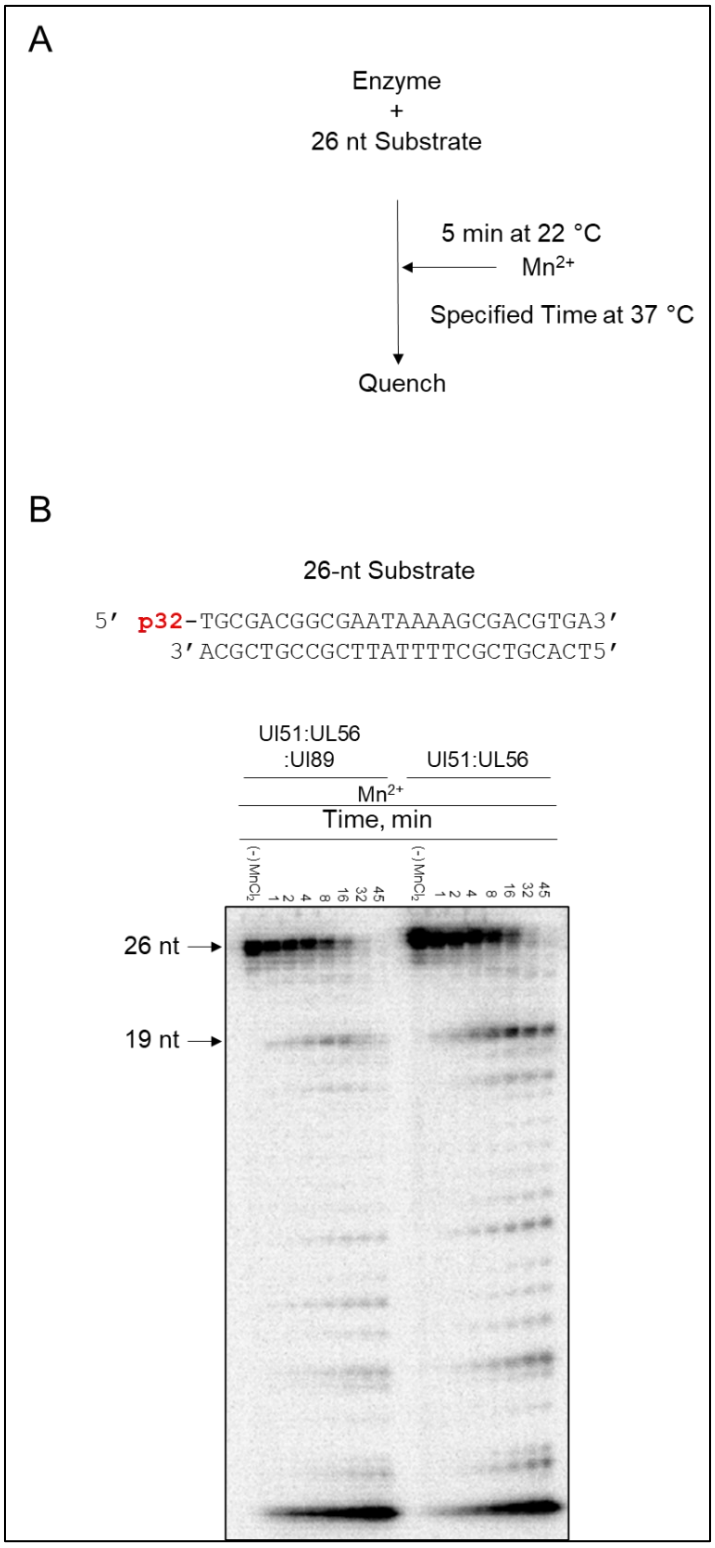


Figure 3.5. Nuclease activity of UL51/UL56/UL89 versus a binary complex of UL51 and UL56. A. Schematic detailing the reaction setup. B. TBE PAGE showing the migration pattern of the products of the endonuclease activity of both constructs.

3.2.5 Varying substrate length does not modify endonucleolytic cut position suggesting that the observed cutting is sequence specific.

To assay whether the observed endonucleolytic cut is sequence specific we varied the length of the substrate and examined the resulting cleavage pattern (Figure 3.6). In the presence of a 24-nt, 26-nt, or 29-nt substrate an initial 19nt endonucleolytic cut is observed (Figure 3.6). This suggests that the observed endonucleolytic cut is likely sequence specific rather than length specific. To confirm this, it would be important to both mutate the nucleic acid residues surrounding the cut site as well as to order constructs where the distance of the cut site from the 5' end of the substrate is increased. If modification of the cleavage site results in loss of the cut and increasing the distance of the cut site from the 5' end of the construct has no effect, we could conclude that the observed cuts are indeed sequence specific.

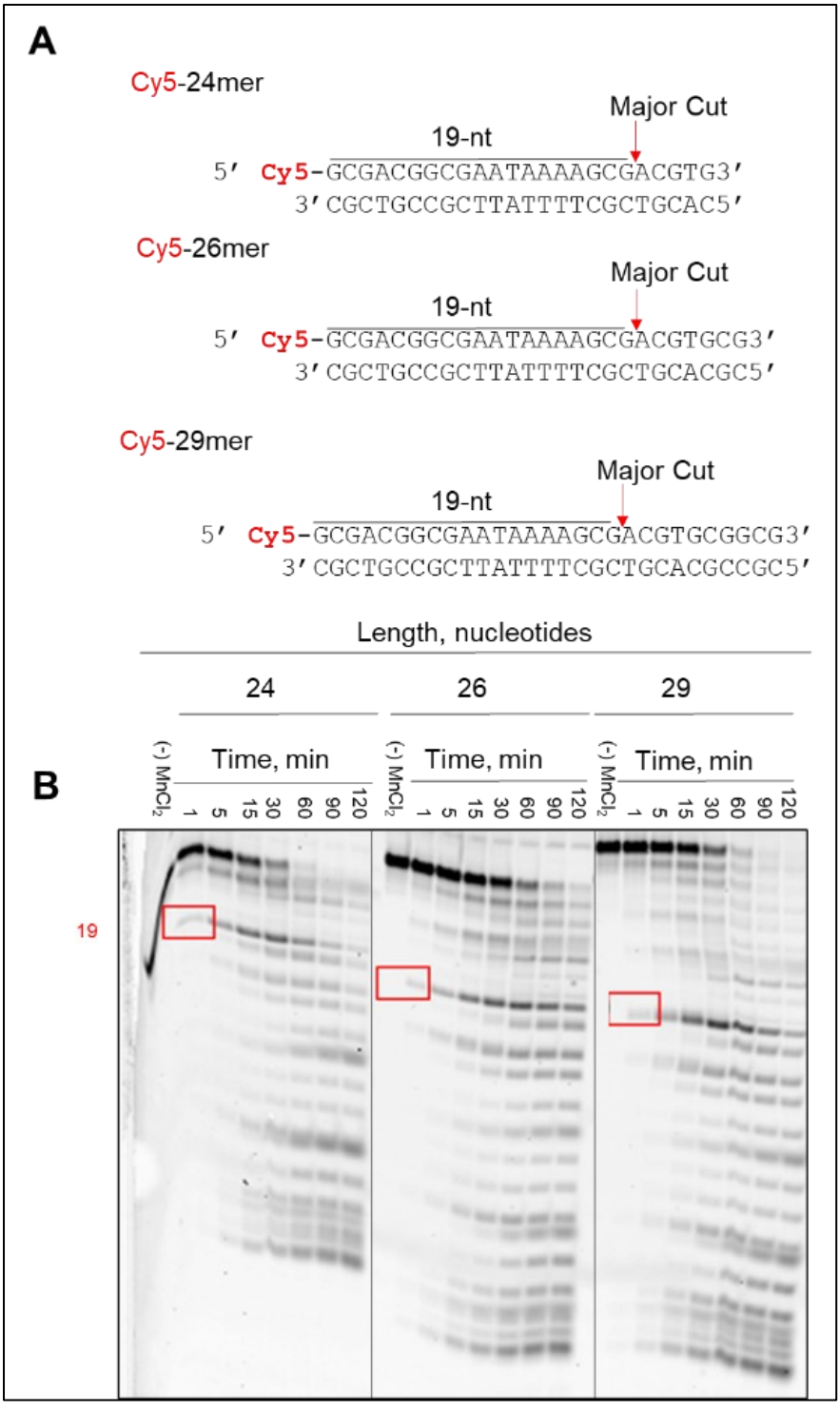


Figure 3.6 T3 Produces a s endonucleolytic cleavage product that appears to be sequence specific. (A) Schematic diagram of the substrates used in the assay. (B) PAGE showing the

production of what appears to be a sequence specific endonucleolytic 19-nt cleavage product (red box) by T3.

3.2.6 Titration of ATP abrogates the endonuclease activity of the CMV terminase complex.

In addition to possessing a nuclease activity, the CMV terminase is known to possess an ATPase activity that provides the necessary energy for viral genome packaging (25,82). Therefore, I hypothesized that ATP might affect the ability of the terminase complex to cleave DNA. I observed that the titration of both ATP and ADP inhibited the nuclease activity of the complex (Figure 3.7). Titrating ATP or ADP resulted in a 50% reduction of DNA cleavage (IC_{50}) at concentrations of 2 mM and 4 mM, respectively. This is in line with observations that have been made in phage showing that the ATPase domains of the large terminase, i.e. the subunit homologous to UL89, are involved in DNA binding and required for nuclease activity (156). Specifically, these data suggest that saturating the ATPase domains of the CMV terminase complex may prevent the binding of DNA substrates to the terminase complex. Furthermore, these data lend support to a possible model, such as that proposed for phages, where ATP binding modulates the various conformational states of the CMV terminase complex (157). Additionally, the doubling of the IC_{50} observed when ADP is titrated instead of ATP suggests that the negatively charged phosphate moiety of ATP may be important for the binding of this substrate to the terminase complex (Figure 3.7).

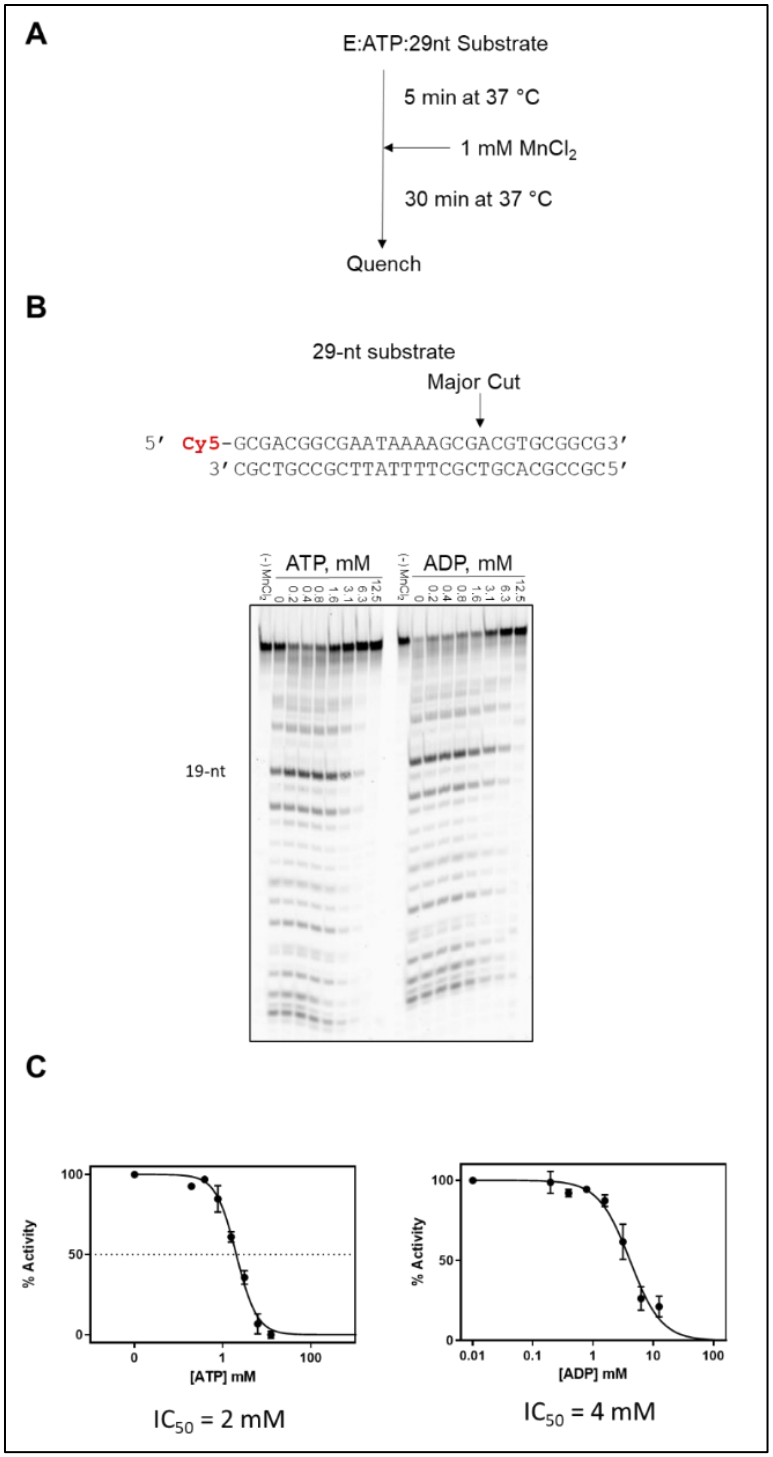


Figure 3.7 Inhibition of T3 nuclease activity by ATP and ADP. (A) Schematic diagram detailing the reaction. (B) PAGE showing the effects of ATP titration on t3 nuclease activity. (C) graphical representation of the data presented in panel B. n=2

3.2.7 The purified terminase complex possesses ATPase activity.

The terminase complex is expected to contain an ATPase activity that provides the required energy for genome packaging in herpesviruses such as CMV (158). To assay for this activity we made use of a polyethyleneimine (PEI) based thin layer chromatrography assay (TLC). In these assays terminase is mixed with ATP and reactions started via the addition of 1mM $MnCl_2$.. The reactions are then quenched, spotted, and run on PEI-TLC plates. This allows for the reaction components to be separated on the basis of size and detected via autoradiography. $MnCl_2$ dependent ATPase activity was observed for the purified complex (Figure 3.8B). ATP was titrated and the Km and V_{max} of the terminase was determined to be 123 μM and 0.03 μM sec^{-1} respectively. No activity was observed in the absence of protein suggesting that the activity observed is due to the protein preparation.

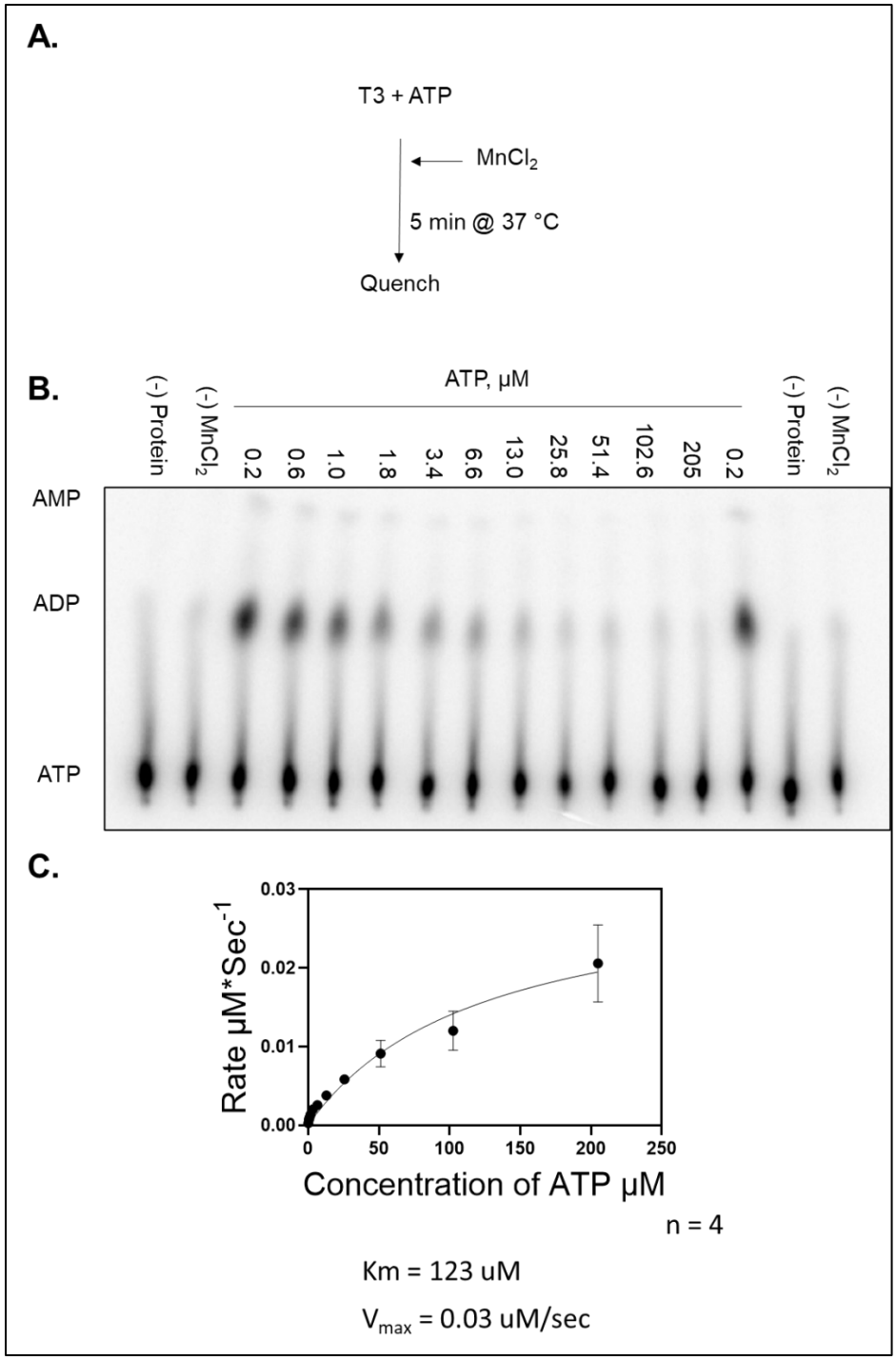


Figure 3.8 ATPase activity of the terminase complex UL51/UL56/UL89. (A) Schematic diagram of the ATPase reaction setup**.** (B) Polyethylene imine cellulose TLC assay showing the

conversion of ATP into ADP by T3. (C) Effects of ATP titration on the rate of ATP hydrolysis by T3 and determination of Km and Vmax. N=4.

3.2.8 Both deletion of UL89 and mutation of the putative UL56 walker A box does not abrogate ATPase activity.

Both UL89 and UL56 have been suggested to contain ATPase activity (82,86). However, to date only UL56 has been shown to have ATPase activity in *invitro* biochemical assays (83). Therefore, to study this activity, we deleted UL89, however this did not produce any difference in the ATPase activity observed (Figure 3.9A and B). This suggested that the ATPase activity of the complex may reside in UL56 as has been proposed by others (82). However, this is highly unusual as structural studies on the HSV-1 homologue of UL56, UL28, revealed that this protein lacks folds that are associated with ATPase activity (76).

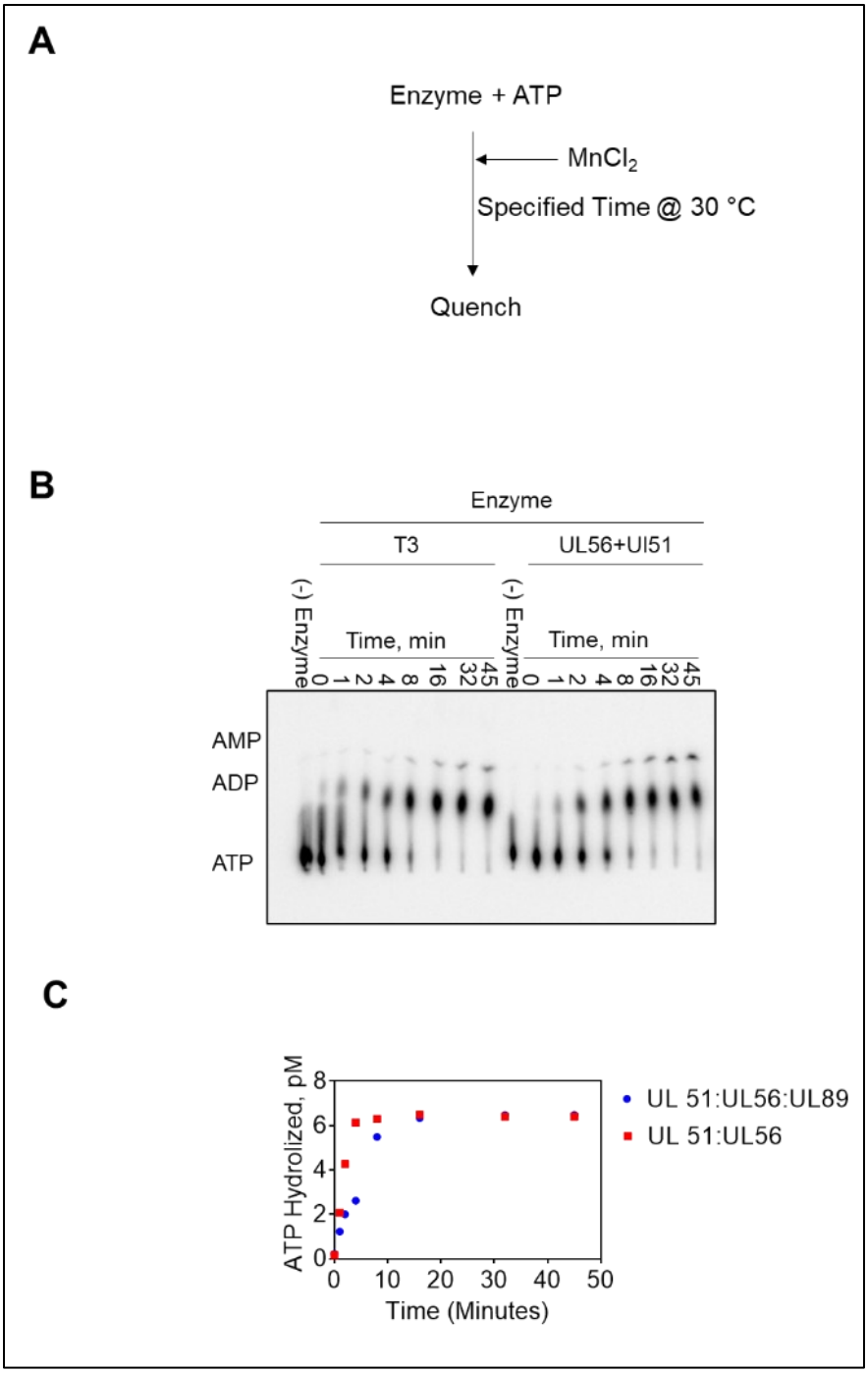


Figure 3.9 Comparison of the ATPase activity of the Tripartite CMV terminase coupled UL51:UL56: UL89 with a binary complex of UL51 and UL56. (A) Schematic diagram of the reaction setup. (B) PEI-TLC assay comparing the hydrolysis of ATP over time by either T3 or a binary complex consisting of UL56+UL51. (C) Graphical representation of the data in panel B.

Therefore, we suspected that the ATPase activity might reside in UL56 and generated a terminase mutant containing two mutations in the proposed walker A box of UL56 that has been proposed by Scholz *et al*. (83). Specifically, we generated a UL56 variant containing a G714I, K715N double amino acid substitution. Mutation of these residues resulted in a >75% loss of ATPase activity in other studies making use of truncated GST tagged UL56 (83). The conserved lysine, if it resembles those found in typical walker A boxes, is critical for phosphate binding and its mutation to isoleucine should result in a loss of activity. However, in our hands this was not observed (Figure 3.10). This suggests that the observed ATPase activity of the terminase complex is likely due to either a copurified contaminant or a non-traditional ATPase motif that has yet to be identified.

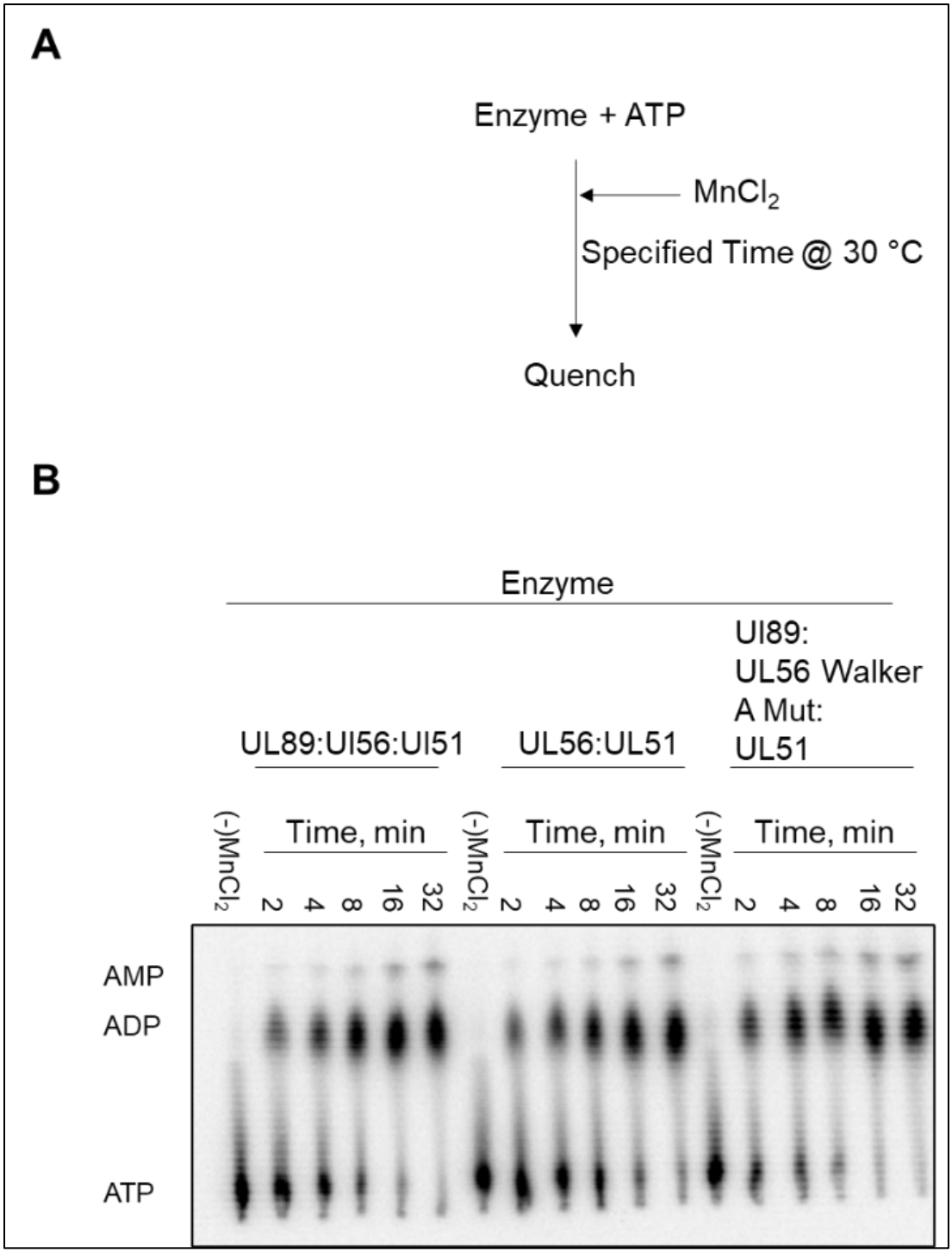


Figure 3.10 Mutating the walker A motif of UL56 does not reduce ATPase activity. (A) Schematic diagram of the reaction setup. (B) PEI-TLC showing the hydrolysis of ATP by

UL89:UL56: UL51, UL56:UL51, and a terminase construct where the walker A motif of UL56 has been mutated.

3.2.9 Terminase and a binary complex of UL51 and UL56 possess DNA binding activity and titration of ATP abrogates DNA binding by terminase.

In order to package DNA into a capsid the CMV terminase complex must bind DNA. Therefore, we developed fluorescence polarization (or anisotropy) based assays to measure the binding of the terminase complex to DNA substrates. In these assays binding of the DNA substrate by terminase reduces the tumbling of the substrate in solution increasing the production of polarized light. This is also referred to as fluorescence anisotropy (Figure 3.11). We titrated either a binary complex of UL51 and UL56 or the tripartite terminase complex (UL51/UL56/Ul89) and observed an increase in fluorescence anisotropy as a function of protein concentration until a plateau, representing binding saturation, was reached. Interestingly, titration of ATP or ADP resulted in a decrease in fluorescence anisotropy for the terminase complex. No such decrease was observed for a binary complex lacking UL89. This suggests that the ATP binding motif present in UL89 may be integral to the binding of DNA substrates by the terminase complex as has been suggested for highly related phage terminase proteins (156). Furthermore, this is supportive of a model akin to that described for phages where ATP binding may regulate DNA substrate binding by terminase(157).

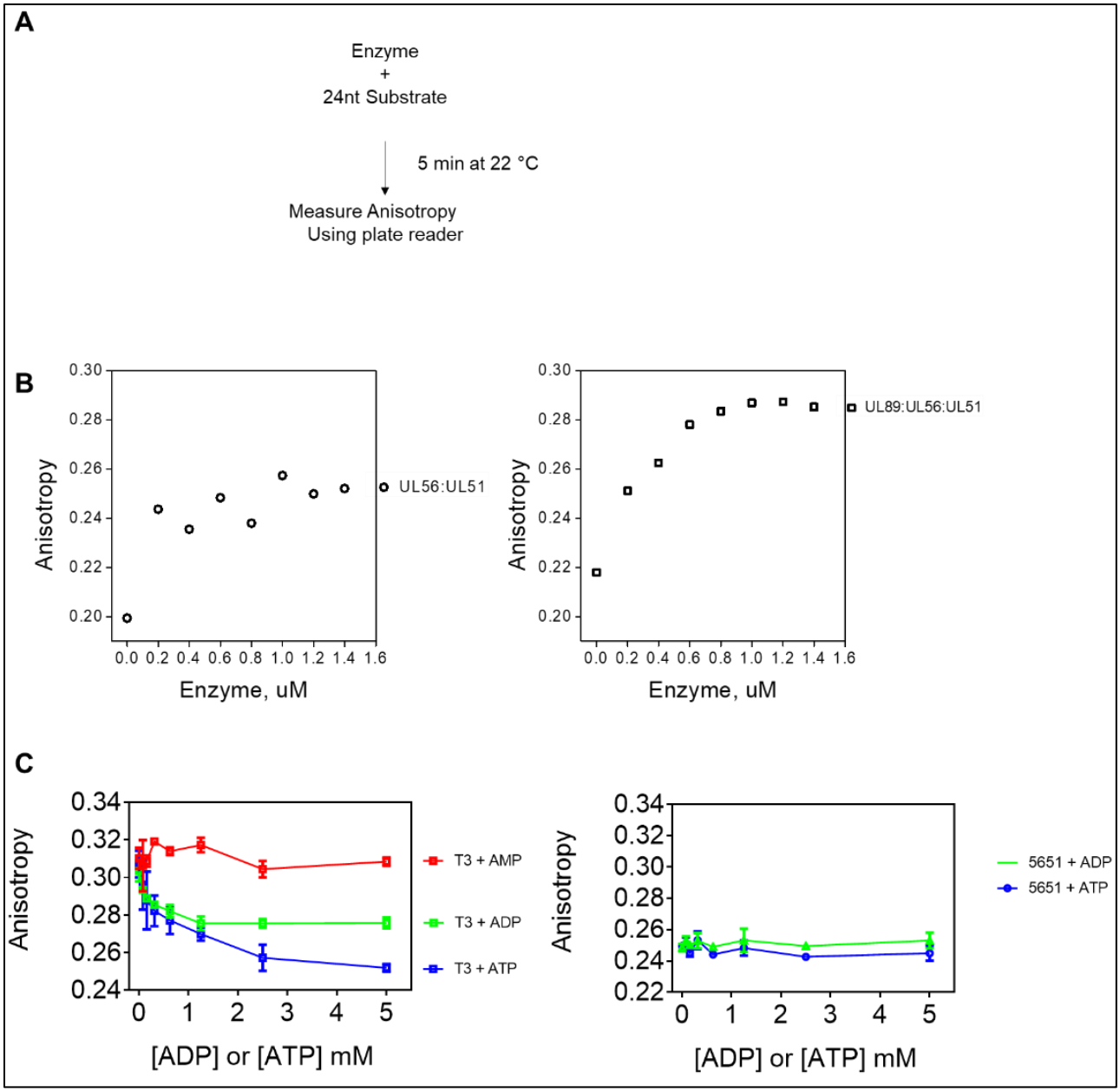


Figure 3.11 DNA binding activity of the tripartite terminase complex (T3) UL51:UL56:Ul89 and a binary complex of UL56 and UL51. (A) Schematic detailing the reaction setup. (B) Effect of titrating enzyme on the production of polarized light by the fluorescent 24nt substrate. (N=1) (C) ADP and ATP inhibit DNA binding by the terminase complex. (N=3)

3.3 Discussion

In this chapter, we investigated the nuclease, ATPase, and DNA binding activities of the human cytomegalovirus terminase complex UL51/UL56/UL89. I, in collaboration with my colleagues Dr. Liu and Dr. E Tchesnokov, successfully expressed and purified the CMV terminase complex and developed assays to measure its nuclease, ATPase, and DNA binding activities. In addition to this, we were able to show that the presence of UL51 is required for the expression of the complex in insect cells.

It is generally accepted that the nuclease activity of the complex resides in the c-terminal portion of Ul89 while the ATPase may reside in both UL56 and UL89 (25,82,83,86). However, ATPase activity has only been functionally demonstrated in UL56 (82,83). The ATPase and nuclease activities of the complex have never been investigated using full length CMV terminase complex and the lack of these assays poses a major obstacle to characterizing and developing novel anti-CMV compounds that target this complex. To this end, we expressed and purified the CMV terminase complex and developed assays to measure its ATPase, nuclease, and DNA binding activities.

The purified terminase complex was found to possess ATPase, Nuclease, and DNA binding activities. However, mutation or deletion of the requisite domains failed to abrogate the Nuclease or ATPase activities of the complex. Therefore, we were unable to exclude the possibility that the observed activities are due to a copurified contaminant. Alternatively, the observed activities could point to the presence of non-canonical ATPase and nuclease domains within the complex. Future work must address these issues. If this is done, the developed ATPase, nuclease, and NTPase assays could be of great use in the characterization and refinement of antiviral compounds such as letermovir that target the CMV terminase complex.

If future work can remove this contaminant or a novel motif conferring ATPase or nuclease activities is discovered within one or more of the proteins comprising the terminase complex then the assays developed here would be of significant use in the characterization of the mechanisms of action and mechanisms of resistance to antivirals, such as letermovir, which target the CMV terminase complex.

One way to increase the purity of the terminase complex and remove contaminating activities would be to incorporate additional steps, such as ionic exchange chromatography, into the purification of the complex. Following these additional steps, the resulting fractions could be screened for nuclease or ATPase activity. If the mutants then lost activity it would be possible to draw conclusions with regards to which subunit(s) possess the nuclease and ATPase activities of the CMV terminase.

An additional possibility is that the nuclease of the complex is not active on the substrate we tested and we therefore only pickup contaminating activities. To address this, we could work with substrates making use of both *Pac* motifs and an intervening sequence. This sequence would be more biologically relevant and it is possible these sequences may be required for the activity. Therefore, if upon further purification the detected nuclease activity is lost, it would make sense to switch to this larger substrate to see if an endonuclease activity is recovered. If this occurs, this would provide evidence that the pac sequences are required for the nuclease activity of the terminase.

However, it is also possible that the nuclease and ATPase domains identified in UL89 and UL56 respectively are not responsible for the activities of the complex. If these motifs are required for viral cleavage and packaging, they will be essential to viral replication. CMV reverse genetics

systems have been described (159). Constructs could be generated containing amino acid substitutions in the proposed ATPase and nuclease motifs of both UL89 and UL56. If infectious clones are recovered this would indicate that these amino acids are not required for CMV genome packaging and, as such, they can be ruled out.

In addition to this approach, we could turn to structural biology for answers. Specifically, a structure of the HSV-1 terminase in complex was recently solved using a cryo-electron microscopy based approach (76). The purity of the protein preparations used in this publication appear to be similar to those which we obtained for the CMV terminase (Figure 3.1). Therefore, it is highly likely that a similar approach could be used to obtain structures of the CMV terminase complex. Specifically, the purified complex could be soaked with either nucleic acid substrate, ATP, and metal ions and these structures solved. These structures would provide us with information on which amino acid residues and components of the complex are involved in binding either ATP or nucleic acid substrate or co-ordinating metal ions. This information would in turn, due to the fact that structure and function strongly correlated, provide us with information that would allow us to generate targeted amino acid substitutions and unambiguously assign the enzymatic activities of the complex.

We also showed that the titration of ATP abrogates the nuclease and DNA binding activities of the terminase complex. Therefore, if we could exclude contaminating nuclease activities this biochemical data could provide insight into modes of DNA packaging by the terminase. Specifically, this data would allow for the refinement of models related to DNA packaging and terminase. For example, phage terminases that are highly related to the CMV terminase grip DNA using a walker A like motif (156). If such a motif exists in the CMV terminase this would explain why under high concentrations of ATP DNA binding and endonuclease activity are lost. This

would lend support to a model for CMV DNA packaging where ATP binding modulates the conformational state of the terminase complex (157).

Furthermore, the systems we developed, particularly the endonuclease, ATPase, and DNA binding assays are easily adaptable to the screening of small molecule inhibitors. Thus, if contaminating activities were excluded these assays could be used for the discovery, characterization, and refinement of novel inhibitors of the CMV terminase complex. The major advantage of this system, over those previously described, is that this system makes use of full-length protein and therefore is more biologically relevant potentially allowing for the detection of inhibitors that would be missed by other approaches.

Chapter 4: Characterization and Inhibition of the Influenza PA endonuclease.

4.1 Introduction

Influenza virus infection is a source of significant morbidity and mortality worldwide (111,160). Recent estimates suggest that Influenza infection is responsible for 300,000 - 650,000 deaths annually (160). Hence, there is a concern that the co-circulation of influenza, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and other respiratory viruses could exceed healthcare capacity in several settings (161). Furthermore, the utility of anti-influenza therapies is often diminished due to the emergence of resistant strains (162). For example, amantadine, which antagonizes the influenza A proton channel is no longer recommended as a treatment due to widespread resistance (162). Alternatively, neuraminidase inhibitors such as oseltamivir can be used for the treatment of influenza A and B (162); however, the potential clinical benefits are still

debated (162). Recently, baloxavir marboxil (BXM), a first-in-class antiviral targeting the viral polymerase gained FDA approval (91). Intracellular hydrolysis of the prodrug yields baloxavir acid (BXA), which is the active form. BXM treatment, much like oseltamivir, results in a ~20-25% reduction in time until symptom resolution and the clinical utility of both drugs is compromised by the emergence of resistant viruses (163-165). However, BXM can be given as a single dose while oseltamivir is given twice daily over five days (163,164). A recent study also demonstrated that a single dose of BXM showed post-exposure prophylactic efficacy in reducing household transmissions from 13.6% in the placebo group to 1.9% in the BXM group (166). Based on this data, the FDA expanded the approval to post-exposure prevention.

Progress has also been made in elucidating the mechanism of inhibition of viral replication by BXA. The viral replication complex is a heterotrimer (ht), possessing a 7-methylguanylate (m^7G) cap binding domain (PB2) that allows binding of cellular mRNAs (90,167,168), an endonuclease domain (PA) that cleaves the bound mRNA to generate primers for viral transcription in a process referred to as “cap-snatching”(27,89,90,168-170), and the PB1 domain which contains the RNA dependent RNA polymerase (RdRp) active site that is required for RNA synthesis during replication and transcription (27,90,117,171). PA recruits two divalent metal ions to its active site (169,172) and BXA was designed to bind to these catalytic metal ions (172). A similar approach was utilized for the development of inhibitors of the human immunodeficiency virus type 1 (HIV-1) integrase (173,174) and ribonuclease H (RNase H) inhibitors (26,175). While several HIV-1 integrase inhibitors are approved for clinical use, HIV-1 RNase H inhibitors remain investigational. A common motif for these metal binders is three strategically positioned heteroatoms that form an “anchor domain” (176). Crystal structures of the truncated PA subunit with and without BXA have shown that the inhibitor binds to the two metal ions with three oxygen

atoms and a hydrophobic area contributing to this binding (172). This area is generally referred to as the “specificity domain”, which provides selectivity for a specific target (Fig. 4.1).

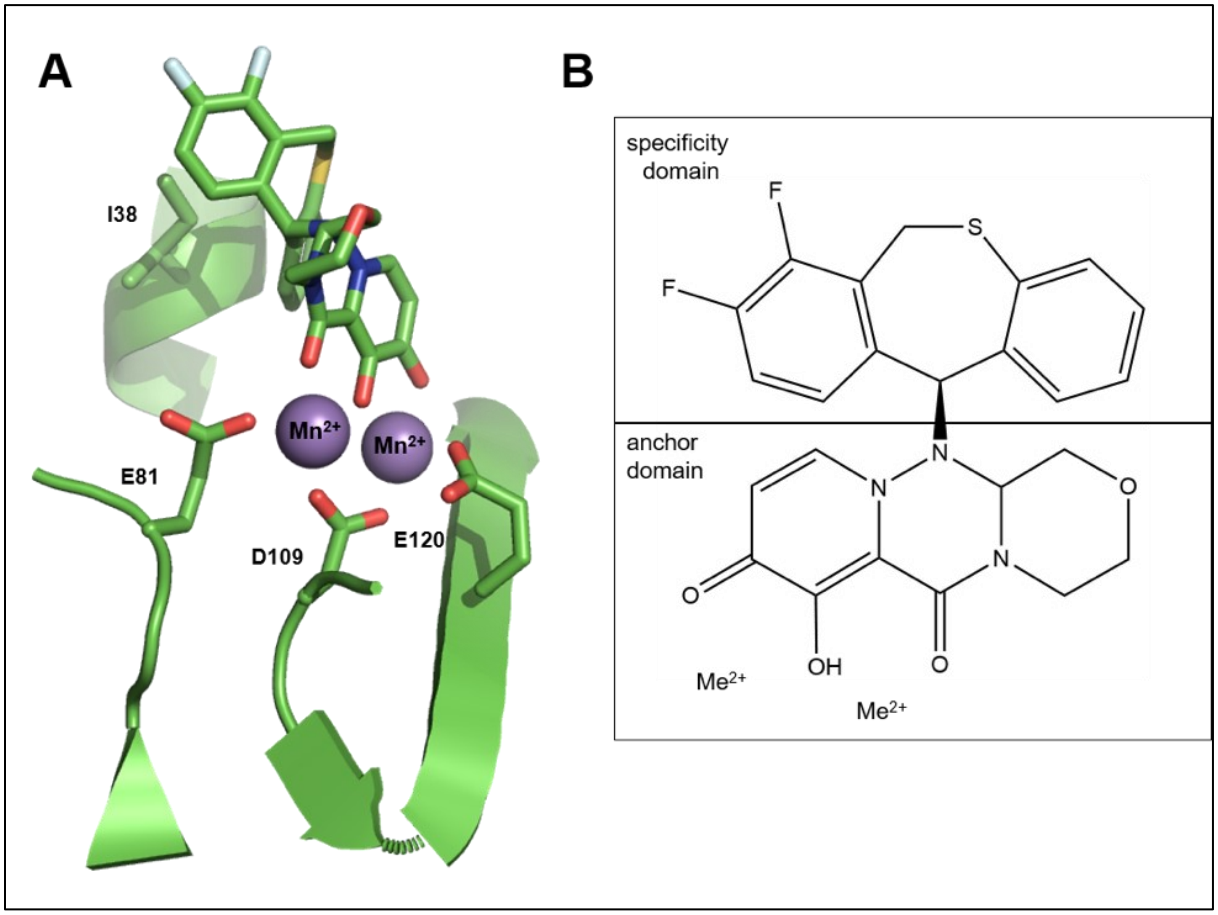


Figure 4.1 Baloxavir Acid (BXA) chelates divalent metal ions at the active site of the Influenza B cap-snatching endonuclease subunit (PA). (A) Model depicting baloxavir acid (BXA, top) binding to the Influenza B PA active site (PDB:6FS8) (21). Selected residues are depicted as sticks. E81, D109, and E120 coordinate metal ions (purple balls) which bind to the anchor domain of baloxavir. I38 forms part of the hydrophobic interface that interacts with the specificity domain of BXA. (B) Schematic diagram of baloxavir acid (BXA) highlighting anchor and specificity domains. Me^{2+} represents either Mg^{2+} or Mn^{2+}.

Mutations conferring resistance to BXM have been observed clinically (163-166,172). In a recent trial, I38T/M/F substitutions appeared in 9.7% of patients receiving BXM (163). The impact of these substitutions on BXM susceptibility has been explored using cell culture systems. Mutant viruses are associated with up to 50-fold increases in EC_{50} values with I38T showing the most significant effect (172). The I38T mutation reduces Van Der Waals interactions between the hydrophobic area of BXA and its binding pocket (172,177). Here we studied BXA-mediated inhibition of recombinant influenza B (FluB) heterotrimeric PA/PB1/PB2 RNA dependent RNA polymerase (RdRp), referred to hereafter as FluB-ht. The biochemical data demonstrate that BXA can be classified as a tight binding inhibitor and that the I38T mutation diminishes these properties.

4.2 Results

4.2.1 Cap-snatching endonuclease activity of FluB-ht wild-type (WT) and FluB-ht PA I38T

The baculovirus expression system has been successfully used to express the RdRp complex of several segmented negative-sense RNA viruses including human influenza B (27,117,142,154). Here we used this approach to generate WT FluB-ht, a variant associated with BXM resistance (FluB-ht PA I38T), and a variant containing a catalytically inactive endonuclease (FluB-ht PA QNQ) where metal co-ordinating residues were mutated (E81Q, D109N, E120Q) (29,172). The three complexes were purified to near homogeneity (Fig. 2A). Due to the similar molecular weight of the three domains comprising FluB-ht (PA, PB1, and PB2), the complex is seen as a single band on SDS PAGE (Figure 4.2A). Diluting the sample helps to visualize three individual bands (142). In this study, the presence of all three peptides was confirmed using LC-MS/MS (appendices 1).

To evaluate the PA mediated endonuclease activity of FluB-ht, we used an experimental design similar to that which has been previously described (27,117,178). The substrate for the endonuclease reaction is a radiolabeled capped 20-nt long RNA. It has been shown that efficient cleavage of the substrate requires the presence of a second RNA molecule, referred to as vRNA, which resembles the bound viral RNA (27,178-182). Time-course experiments with each of the three enzymes are shown in Fig. 4.2 C. The cap-snatching reaction is initiated in the presence of Mg^{2+} which yields a 12-nt major product with cleavage after 5'G. Minor products represent a shorter 11-nt and a longer 14-nt RNA both with cleavage after 5'A. The three products are distinct from RNA background hydrolysis seen in absence of enzyme. These control reactions show a faint hydrolysis product that migrates slightly faster than the major 12-nt product of the nuclease reaction. Differences in migration patterns are expected as nuclease products contain a 3'-hydroxyl group, while hydrolysis yields a negatively charged 3'-phosphate group. The FluB-ht PA I38T shows a very similar pattern as seen with the WT FluB-ht (Fig. 4.2C); however, a quantitative comparison of WT and PA I38T FluB-ht reveals a reduction of the overall activity of the PA I38T FluB-ht as has been noted by others (Fig 4.2D) (172). Finally, the active site mutant, FluB-ht PA QNQ, shows a complete loss of endonuclease activity (Fig. 4.2 C, D). This data rules out the presence of contaminating endonucleases, while the hot spots for hydrolysis are seen with all three enzyme preparations and controls in the absence of enzyme.

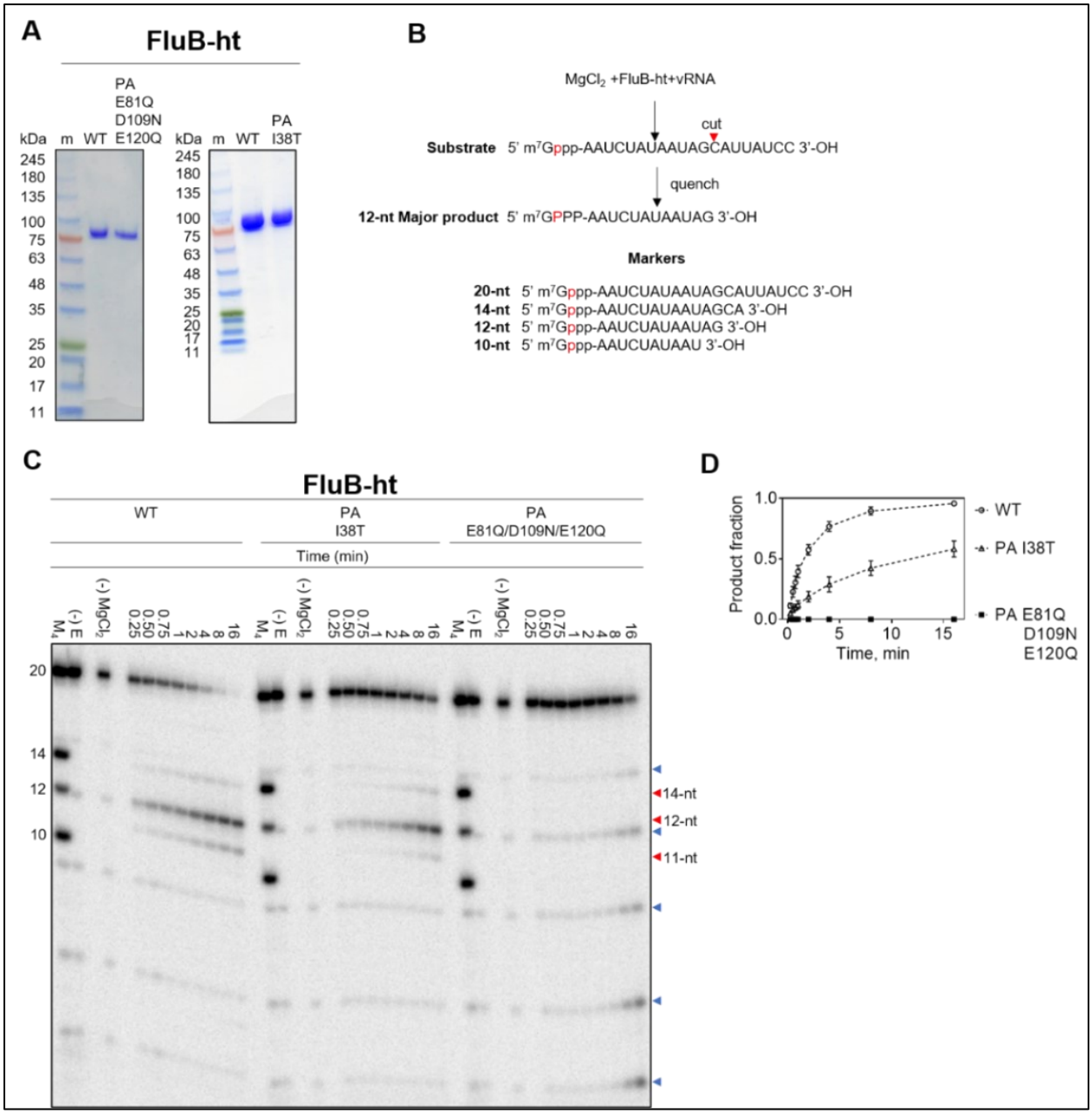


Figure 4.2 Purification and biochemical characterization of the Influenza B polymerase heterotrimer (FluB-ht). (A) SDS-PAGE migration patterns of the purified WT FluB-ht, baloxavir resistance mutant (PA I38T), or endonuclease-deficient mutant (PA E81Q/D109N/E120Q). Enzyme preparations were stained with Coomassie brilliant blue R-250. The size of the molecular weight markers (m) is indicated in kDa to the left of the gel. The band above the 75 kDa marker

corresponds to full length FluB-ht containing PA, PB1, and PB2 subunits. The identity of each of the subunits was confirmed via LC-MS/MS (appendices 1). (B) Schematic representation of the nuclease reaction with the position of the radiolabel highlighted in red. The sequences of the 20, 14, 12, and 10-nt markers are highlighted (Lane M_4). **(**C**)** Endonuclease activity of WT, PA I38T, or endonuclease-deficient (PA E81Q/D109N/E120Q) FluB-ht on PAGE. M_4 indicates the migration patterns of an equimolar mixture of 5'- m7G capped 20-nt, 14-nt, 12-nt, and 10-nt oligos here utilized as molecular weight markers (-) E indicates the migration pattern of 5'capped 20-nt substrate in the absence of enzyme. (-) $MgCl_2$ indicates the results of the reaction after 16 min in the absence of Mg^{2+}. Blue arrows indicate hydrolysis products while red arrows indicate the position of bona fide endonucleolytic cleavages. (D) Graphical representation of data shown in A. Data points are connected using a dotted line to illustrate the progress of the reaction. Product fraction refers to the ratio of the signal produced by the 5' capped 11- and 12-nt products to the sum of these products plus the remaining substrate. The 14-nt product was not quantified as its contribution to the overall signal was negligible. Error bars represent the standard deviation of at least three independent experiments. All experiments were performed under the following final conditions: 55 nM FluB-ht, 100 nM 20-nt substrate, 30 mM Tris-pH 7.5, 25 mM NaCl, 1.7 μM vRNA, and 5 mM $MgCl_2$.

4.2.2 Inhibition of FluB-ht WT and FluB-ht PA I38T endonuclease by BXA

To quantify inhibition of WT and PA I38T FluB-ht by BXA we initially determined IC_{50} values of 112 nM and 374 nM for the WT and I38T variants, respectively (Fig. 4.3A). This represents a ~3-fold decrease in inhibition with the mutant enzyme. However, this IC_{50} value for WT FluB-ht is an order of magnitude higher than previously reported (91). The concentrations of enzyme used in this earlier study was not indicated, which makes it difficult to compare the data (91). The IC_{50}

value of 112 nM measured in our experiments is similar to the enzyme concentration of 165 nM used in our assay. IC_{50} values for tight binding inhibitors are expected to increase with enzyme concentration (183). Specifically, plots of IC_{50} as a function of enzyme concentration correspond to the linear function $IC_{50} = ½ [E] + K_i^{app}$ (183).

To test this hypothesis, IC_{50} values were determined for both WT and PA I38T FluB-ht at various concentrations of enzyme (Fig. 3B). Plots of IC_{50} as a function of enzyme concentration for both WT and PA I38T FluB-ht yielded linear plots with slopes of ~0.5 (Fig. 4.3B), suggesting tight binding in both cases. However, the K_i^{app} of BXA for PA I38T of 217 nM is ~18 fold increased with respect to WT FluB-ht (12 nM) (Fig. 4.3C). The lower this value, the higher the probability that tight binding is occurring. The ideal tight-binding inhibitor would have an extrapolated K_i^{app} of zero resulting in the IC_{50} being equal to exactly half the enzyme concentration. Furthermore, this 18-fold difference in K_i^{app} is significantly larger than the initially observed 3-fold difference in IC_{50} that we obtained utilizing a single FluB-ht concentration (165 nM) illustrating the utility of this approach for the biochemical characterization of BXA and tight binding inhibitors in general. The observed IC_{50} values were independent within a narrow range of substrate concentrations (Fig. 4.4), which shows that the enzyme concentration is the only relevant variable that affects the determination of IC_{50} values.

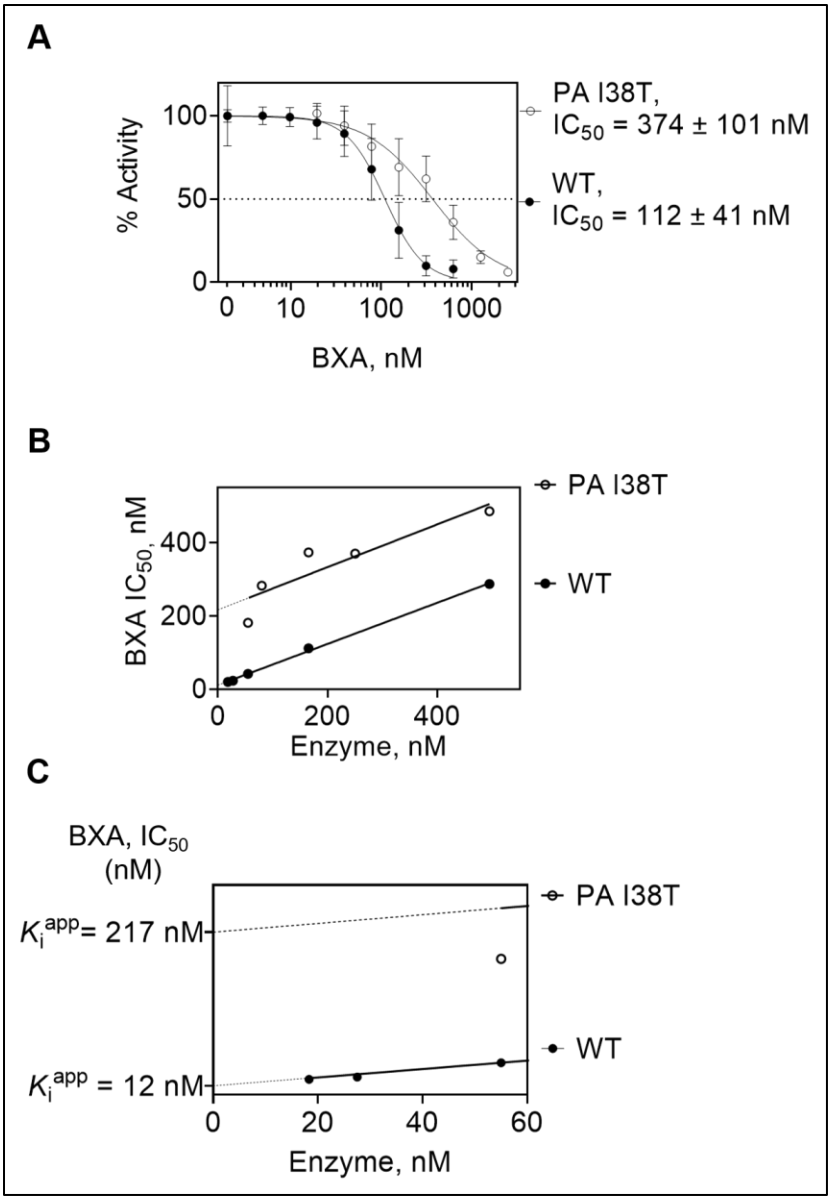


Figure 4.3 Differential inhibition of WT and PA I38T FluB-ht by BXA. (A) 165 nM of WT or PA I38T was mixed with increasing concentrations of BXA and inhibition quantified. The dotted line represents the half maximal inhibitory concentration (IC_{50}) of BXA. Error bars represent the standard deviation of at least three independent experiments. (B) Dependence of the IC_{50} value on FluB-ht concentration. Plots of IC_{50} vs enzyme concentration for both WT and PA I38T yield a linear relationship with a slope of approximately 0.5 (WT = 0.56, PA I38T = 0.58.)

(C) Close up of the Y-axis from panel B emphasizing the apparent inhibition constant (Ki^{app}) of BXA for WT and PA I38T FluB-ht. The dotted line indicates the threshold below which robust IC_{50} determination was not possible due to insufficient endonuclease activity. The percent error for WT IC_{50} measurements was 27% while percent error for PA I38T was 36%. All experiments were performed under the following final conditions: 100 nM 20-nt substrate, 30 mM Tris-pH 7.5, 5% DMSO, 25 mM NaCl, BXA as specified, FluB-ht as indicated, 1.7 µM vRNA, and 5 mM $MgCl_2$.

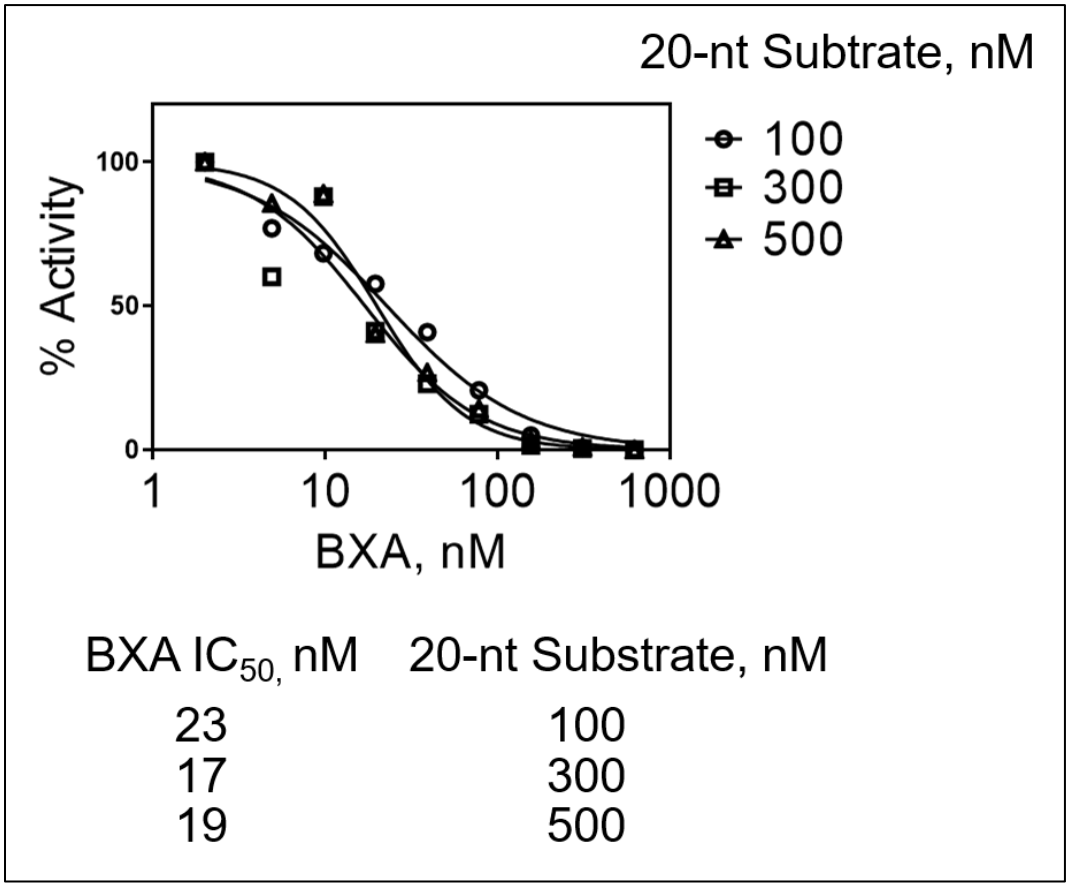


Figure 4.4 IC50 of baloxavir acid (BXA) as a function of substrate concentration. The IC_{50} of baloxavir acid was measured in the presence of 100, 300, and 500 nM 20-nt radiolabeled capped

substrate, 55 nM FluB-ht WT, 30 mM Tris-HCL pH 7.5, 25 mM NaCl, 5 mM $MgCl_2$, and 1.7 μM vRNA.

4.2.3 Effects of order-of-addition on inhibition of FluB-ht WT and FluB-ht PA I38T by BXA.

Previous structural studies have shown that BXA binds to the active site of the isolated PA endonuclease domain (172). However, these complexes lack the RNA substrate. Structures of FluB-ht with bound inhibitor and RNA are also not available. For cleavage to occur, the RNA must traverse the active site in close proximity to bound metal ions. Hence, RNA binding can prevent or reduce inhibitor binding. To test this hypothesis, we performed time-course assays where the order of addition of inhibitor and substrate was varied. The endonuclease reaction in the absence of inhibitor for both WT and PA I38T FluB-ht was largely unaffected by the order of addition of FluB-ht (E), capped 20-nt Substrate (S), and Mg^{2+} (Fig. 4.5A). In the presence of 75 nM BXA (I), inhibition was only observed under conditions where an E: I:Mg^{2+} complex was pre-formed (Fig. 4.5B). Conversely, under the same conditions, no inhibition of the PA I38T FluB-ht variant was observed (Fig. 4.5B). Taken together, these data suggest that BXA binds to the PA endonuclease active site and prevents, at least locally, binding of RNA substrate. Nuclease activity was negligible over longer periods of time, suggesting slow dissociation of the inhibitor.

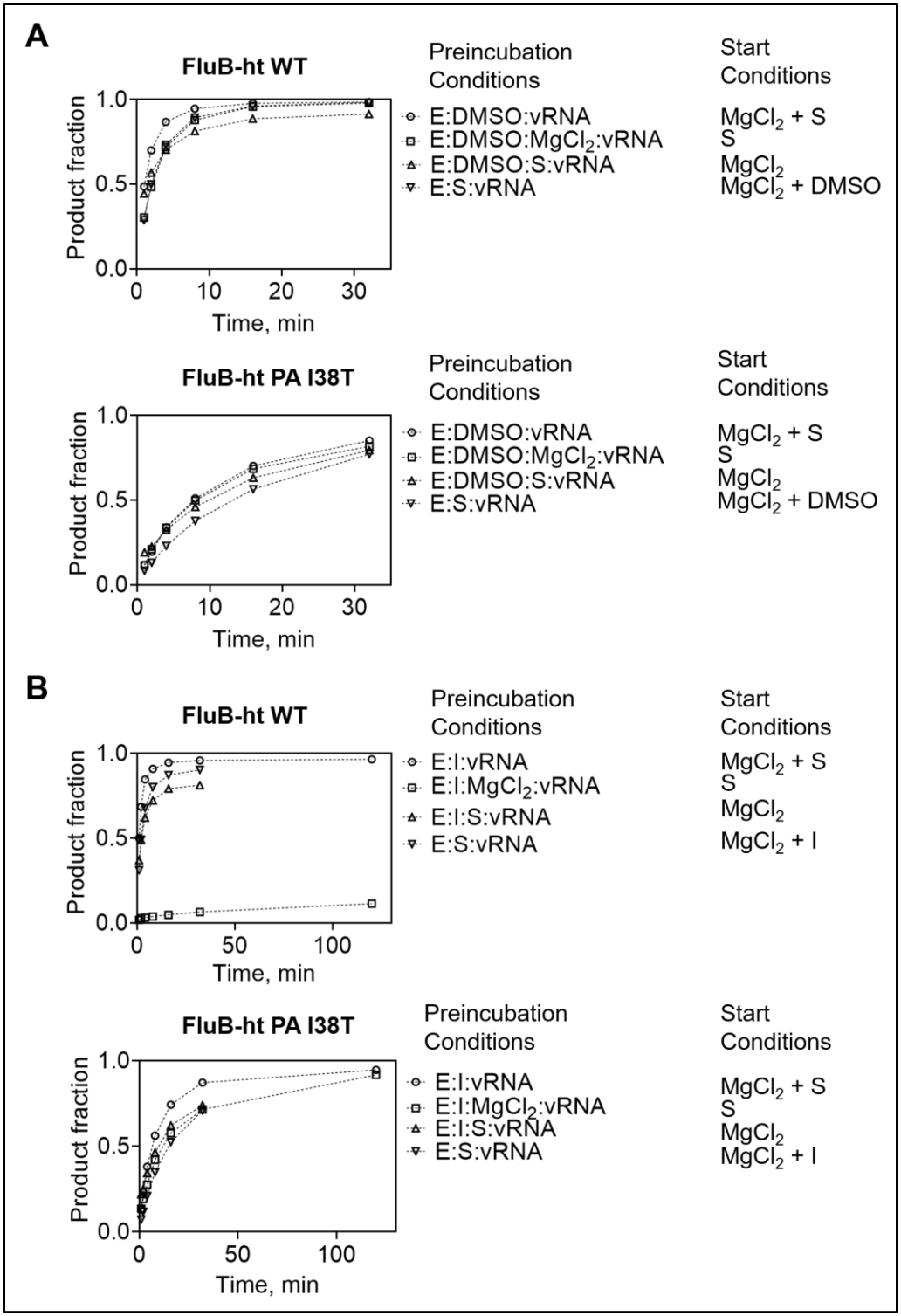


Figure 4.5 Effects of the order-of-addition of FluB-ht (E), substrate (S), BXA (I), and MgCl2 on the endonuclease activity of WT and PA I38T FluB-ht. (A) Order of addition of reaction

components in the absence of BXA has negligible effects on reaction progress for both WT (top) and PA I38T (bottom) FluB-ht. (B) Preincubation of E, I, and $MgCl_2$ is required for inhibition of WT (top, ▫) endonuclease activity. Inhibition of PA I38T (bottom) is not observed regardless of the order of addition of the reaction components. Data represent at least three independent experiments. Error bars were omitted for clarity. All experiments were performed under the following final conditions: 100 nM 20-nt substrate, 30 mM Tris-pH 7.5, 5% DMSO, 25 mM NaCl, 75 nM BXA, 55 nM FluB-ht, 1.7 μM vRNA, and 5 mM $MgCl_2$.

4.2.4 Effect of BXA preincubation on inhibition of FluB-ht WT and FluB-ht PA I38T.

To this end, WT or PA I38T FluB-ht, and BXA were preincubated for a fixed period of time with 75 nM BXA to allow formation of an E:I:Mg^{2+} complex before initiating the reaction with substrate. To assess the association kinetics of the inhibitor, we added the capped RNA substrate at different times of E:I:Mg^{2+} pre-incubation (Fig. 4.6A). For WT FluB-ht increasing inhibition was observed as a function of preincubation length with maximal inhibition being observed after ~ 30 minutes of preincubation (Fig. 4.6B). These data suggest that the association of BXA with PA is slow as is often the case for inhibitors that exhibit tight binding properties (183). Conversely, preincubation length had only negligible effects on the activity of the PA I38T FluB-ht (Fig. 4.6C). Furthermore, an effect is only observed when 600 nM BXA (8 fold more than for WT) is used (Fig. 4.7). The difference observed between WT and PA I38T FluB-ht provides additional evidence to show that the association of BXA with FluB-ht is impaired by the PA I38T substitution.

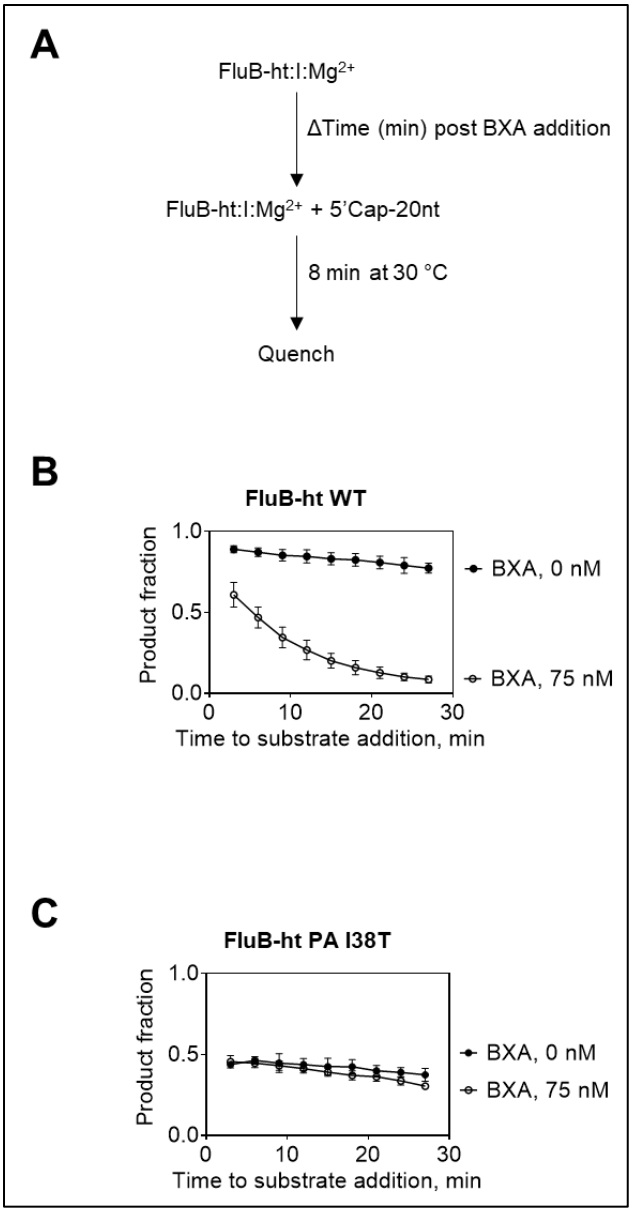


Figure 4.6 Time-dependent inhibition of the PA endonuclease activity of FluB-ht by BXA (I). (A) Schematic representation of the experimental setup. (B) Inhibition of WT FluB-ht by BXA increases with the duration of BXA preincubation (C) Inhibition of PA I38T FluB-ht by BXA is time-independent. Error bars represent the standard deviation of at least three independent experiments. All experiments were performed under the following final conditions: 100 nM 20-nt

substrate, 30 mM Tris-pH 7.5, 5% DMSO, 25 mM NaCl, BXA as specified, 55 nM FluB-ht, 1.7 µM vRNA, and 5 mM $MgCl_2$.

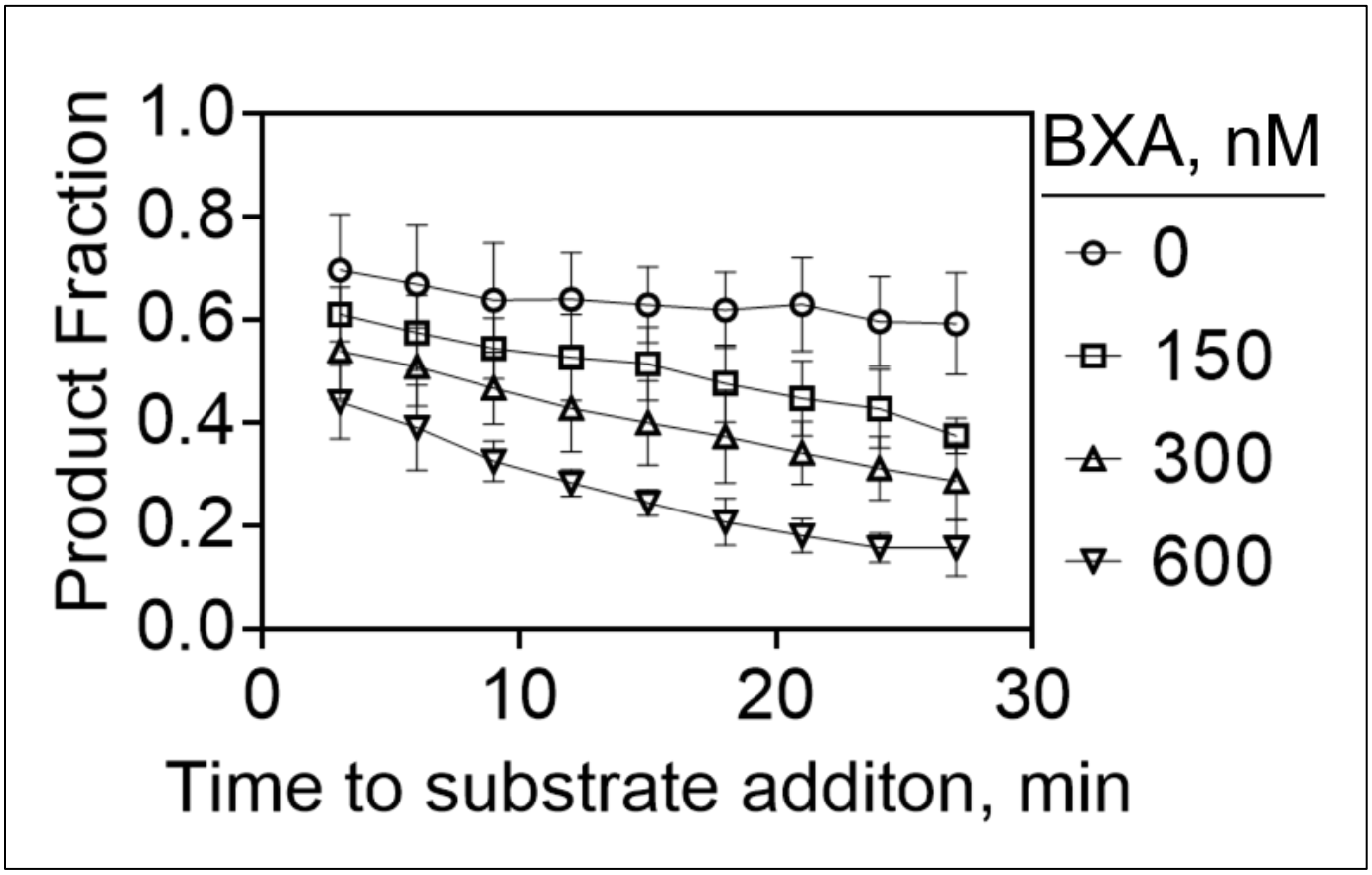


Figure 4.7 Time-dependent inhibition of FluB-ht I38T variant under increasing concentrations of baloxavir acid (BXA). Error bars represent the standard deviation of three independent experiments. Final conditions: 30 mM Tris-HCL pH 7.5, 25 mM NaCl, 5 mM $MgCl_2$, 55 nM FluB-ht PA I38T, 1.7 µM vRNA, 5% DMSO, BXA as indicated and 100 nM 20-nt radiolabeled capped substrate.

4.2.5 Effect of mutations used in the study on the transcriptional activity of FluB-ht.

We assessed the transcriptional activity of FluB-ht WT, PA I38T, and PA E81Q, D109N, E120Q and found that all three preparations were able to elongate a 12-nt capped radiolabeled primer using the 39-nt synthetic vRNA as a template (Figure 4.8). All constructs displayed a similar level of activity suggesting that these mutations in the endonuclease domain do not compromise the ability of the enzyme to perform transcriptional elongation (Figure 4.8 and 4.9).

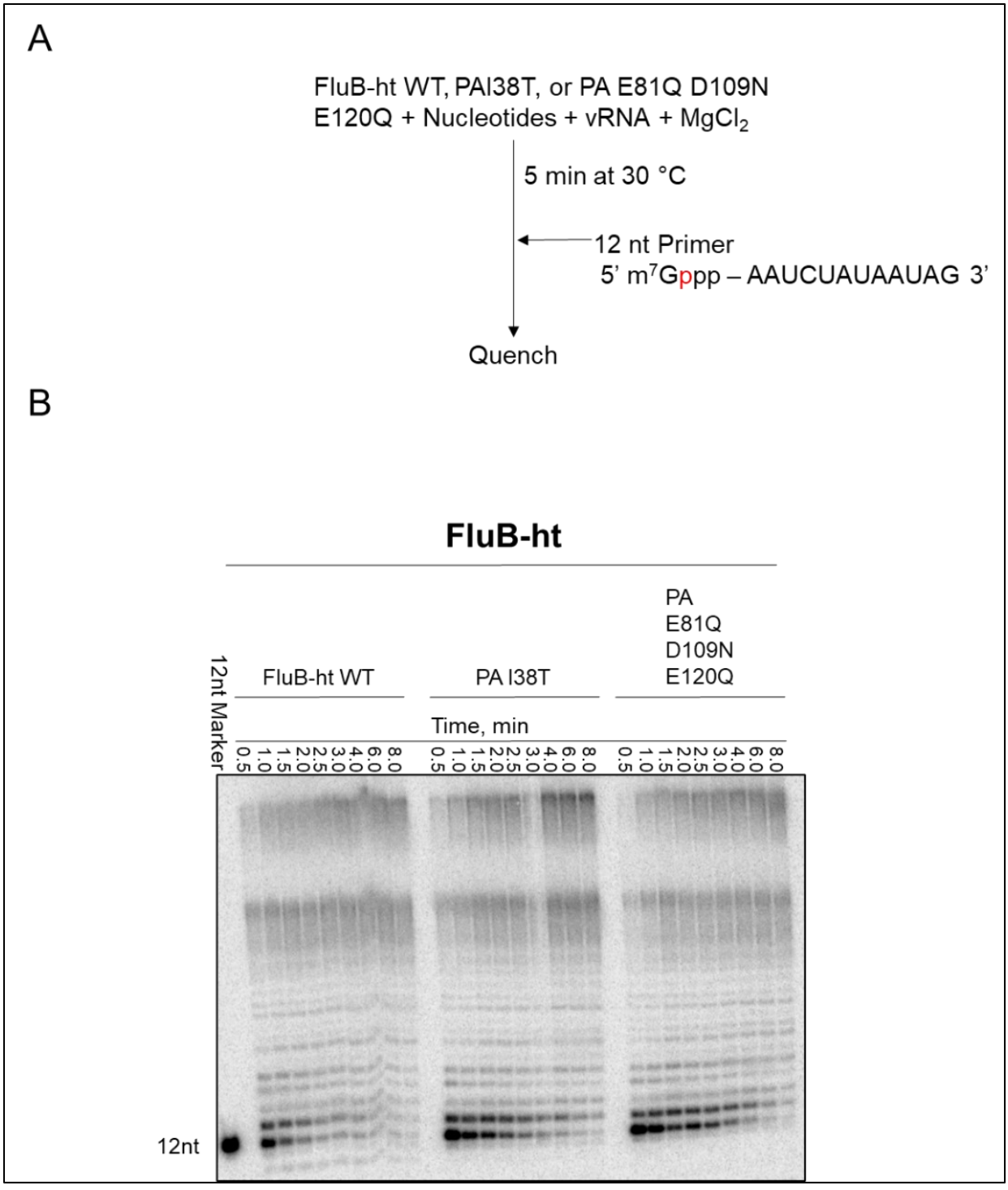


Figure 4.8 RNA synthesis activity of the FluB-ht constructs. The RNA elongation activity of the protein is not abrogated by the PA I38T or E81Q D109N E120Q mutations. (A) Schematic diagram of the assay setup. Red denotes the presence of a radiolabel. (B) PAGE showing the

migration pattern of the products of RNA synthesis by FluB-ht WT, PA I38T, and PA E81Q D109N E120Q.

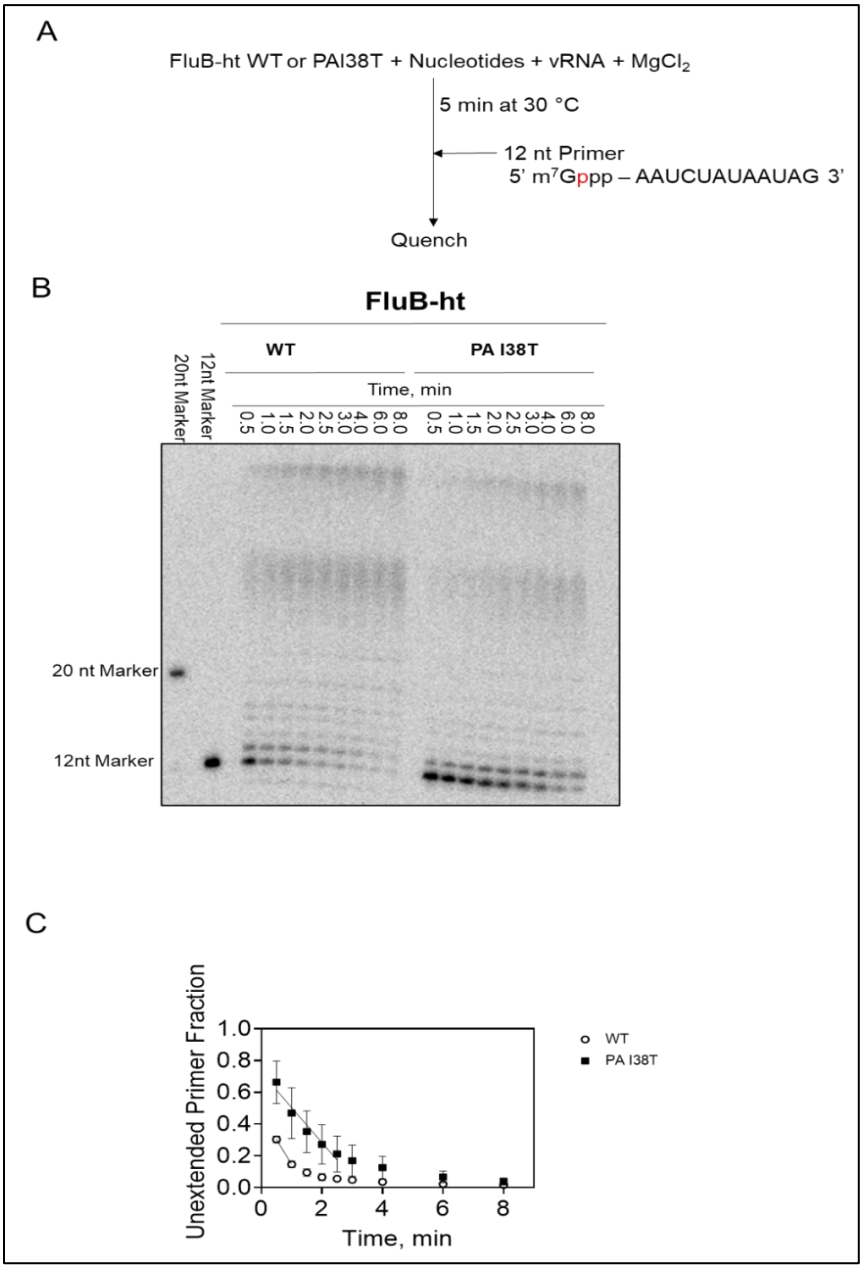


Figure 4.9 RNA elongation activity of FluB-ht WT and FluB-ht PA I38T for a capped 12-nt RNA primer. (A) Schematic diagram of the reaction setup. (B) PAGE showing the migration

pattern of the products of RNA synthesis by WT and PA I38T FluB-ht. (C) Quantification of the results presented in B. The line indicates the linear portion of the reaction. Rates of 12-nt primer extension are as follows: FluB-ht WT = 31 nMol/min, FluB-ht PA I38T = 22 nMol/min. Data represent the average of three independent experiments. (n=3)

4.2.6 Effect of BXA on transcriptional elongation by FluB-ht.

At a concentration of 3 uM BXA does not appreciably inhibit the transcriptional elongation activity of FluB-ht. In contrast, the IC_{50} observed for the endonuclease at this concentration of protein is (~25 nM) (Figure 4.3). This provides strong evidence to show that BXA is a selective inhibitor of the PA endonuclease and that the inhibition of influenza transcription observed in other studies is due to inhibition of cap snatching rather than inhibition of the RdRp activity residing in PB1 (91).

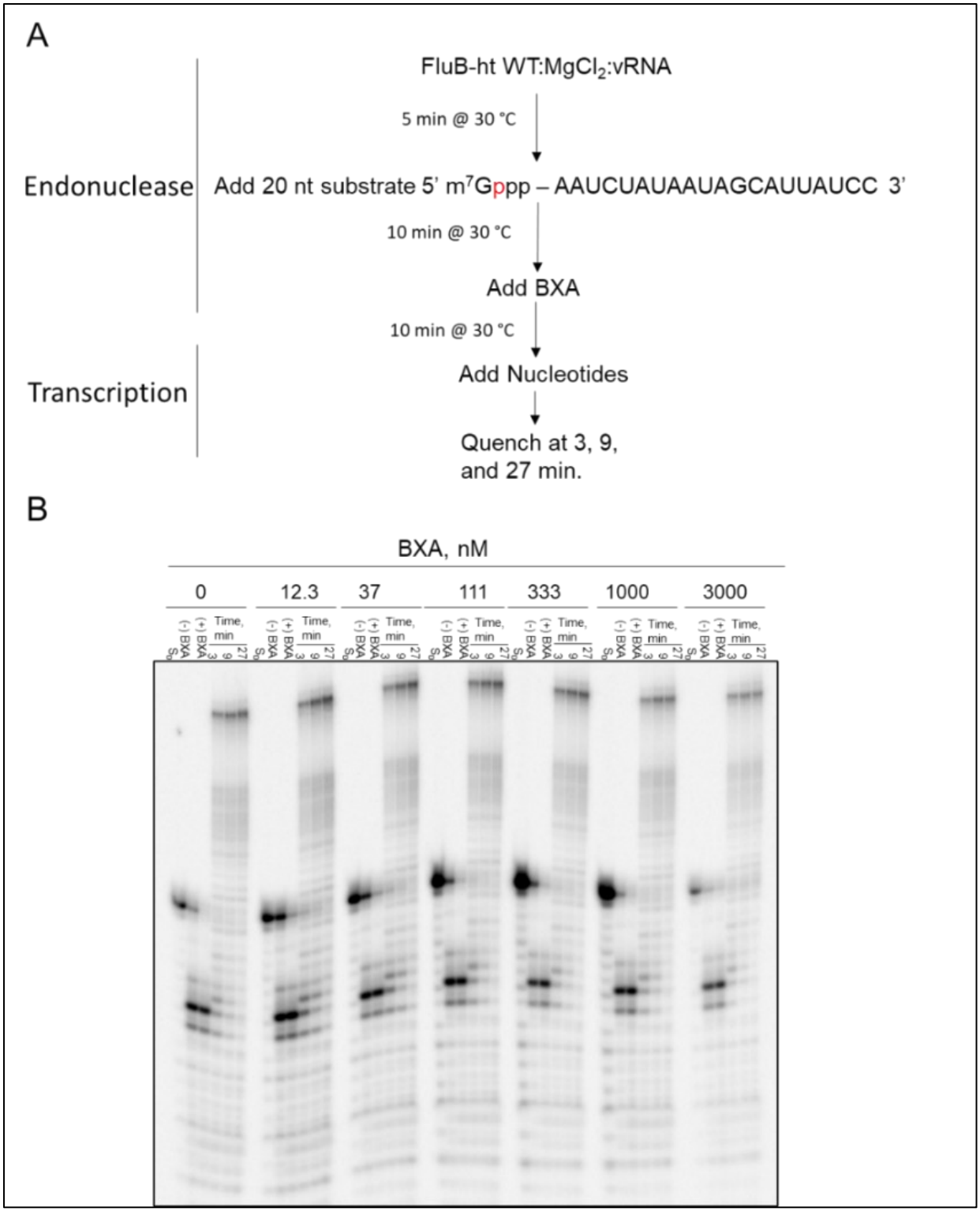


Figure 4.10 Effect of BXA on the RNA elongation activity of FluB-ht WT. (A) Schematic diagram detailing the reaction setup. In brief, a 20-nt capped radiolabeled substrate was

incubated with FluB-ht in the absence of nucleotides to generate a 12-nt capped primer (Shown as (-) BXA Lane) Following this BXA was added to the reaction mixture and incubated for 10 minutes (Shon as (+) BXA lane). Following this, nucleotides were added, and the reaction quenched at 3-, 9-, and 27-minutes post nucleotide addition. (B) PAGE showing the migration pattern of the products of the endonuclease and RNA elongation activities of FluB-ht.

4.2.7 Effect of NaCl Titration on the endonuclease activity of FluB-ht.

To ensure that the salt conditions used in the studies was appropriate we titrated both NaCl and KCl and looked for any effects on endonuclease activity (Figure 4.11). We observed that the activity was optimal between ~25- 50 mM NaCl or KCl (Figure 4.11). Therefore, we concluded that the NaCl concentration used in our assays of 25 mM was appropriate.

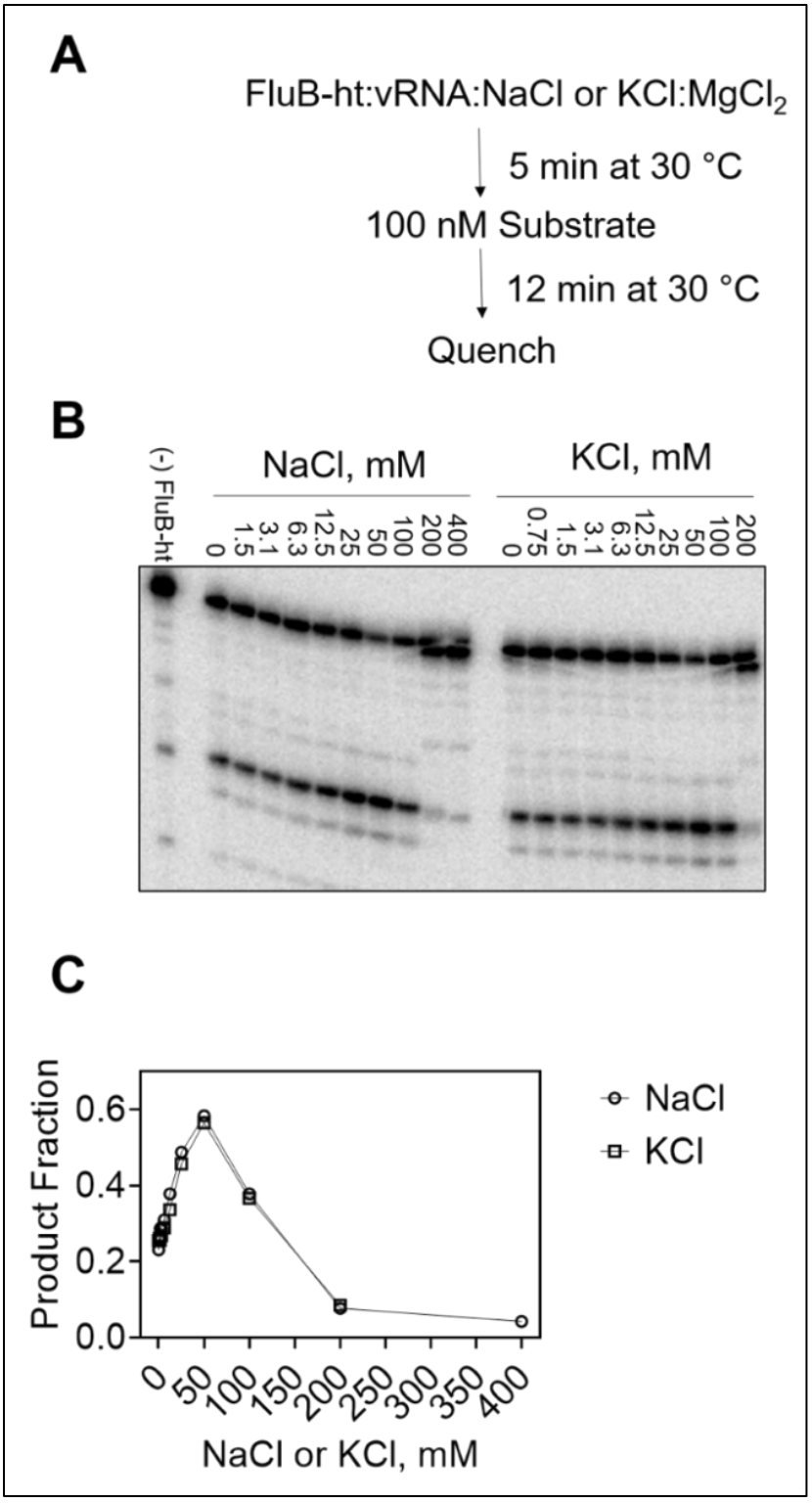


Figure 4.11 Effects of NaCl and KCl titration on FluB-ht endonuclease activity. 55 nM FluB-ht, 1.7 μM vRNA, 5 mM $MgCl_2$, and either NaCl or KCl at the indicated concentrations were

incubated for 5 minutes at 30 °C before the initiation of the reactions with 100 nM of substrate. (A) Schematic detailing the reaction setup. (B) PAGE migration pattern of the reaction products. (C) graphical representation of the data in B.

4.3 Discussion

BXM is a first-in-class inhibitor of the influenza cap-snatching endonuclease subunit of the RdRp complex. Clinical trials revealed reductions in viral load, a faster time of recovery as compared to the placebo group, and utility as post-exposure prophylaxis (163-166,172). BXM is now approved in several countries for the treatment of uncomplicated influenza A and B infection. The drug is highly potent with EC_{50} values in the low nanomolar range with subtle increases for influenza B as compared to influenza A (91,172,184). Mutations at position I38, predominantly I38T, were shown to confer resistance to BXM *in vitro* and *in vivo* (165,172). In this context, EC_{50} values can increase up to 50-fold for influenza A and up to 10-fold for influenza B (172). Structural studies with the isolated PA endonuclease domain bound to the active form of the drug, i.e., BXA, provided important insight on inhibitor binding (172).

Here we employed a biochemical approach to study mechanisms associated with the high potency of the drug and its reduction in the presence of I38T. The endonuclease reaction was monitored with purified FluB-ht WT and FluB-ht PA I38T in the presence of Mg^{2+} and a 20mer, capped model RNA. In contrast to the reported structural studies, this approach may help to better understand the role of the RNA substrate in drug binding and inhibition. We identified a 12-nt major cleavage product, which is consistent with previous reports (27). The FluB-ht PA I38T mutant shows the same pattern, albeit with reduced intensity when compared with WT. The reduction in enzyme activity can translate into diminished replication capacity, which is not

unusual for resistant mutant viruses (172). Indeed, mutant strains with amino acid substitutions at position 38 display fitness deficits in the absence of drug (91,172).

Kinetic parameters for BXA-mediated inhibition of the nuclease activity have not been determined previously. Instead, Noshi and colleagues reported IC_{50} values of ~2nM (91). The low IC_{50} value points to tight binding. For tight binding inhibitors, IC_{50} values increase with increasing enzyme concentration. Changes in IC_{50} values as a function of enzyme concentration correspond to the equation: $IC_{50} = ½ [E] + K_i^{app}$. where K_i^{app} represents an approximation of the true inhibitor constant Ki (183). The determination of K_i^{app} values allows one to compare the efficacy of inhibitors against WT and mutant enzymes. We observed an 18-fold difference in K_i^{app} between WT and I38T FluB-ht, while plots of IC_{50} as a function of enzyme concentration yielded in both cases linear plots with slopes of ~0.5.

Measurements of IC_{50} and K_i^{app} values do not reveal the mechanism of inhibition and Michaelis-Menten kinetics may also not distinguish between noncompetitive, competitive, and uncompetitive inhibition, because the steady-state assumption is not valid for tight binders as this analysis will show patterns consistent with non-competitive inhibition regardless of the true mode of action (Fig.4.3.1) (183).

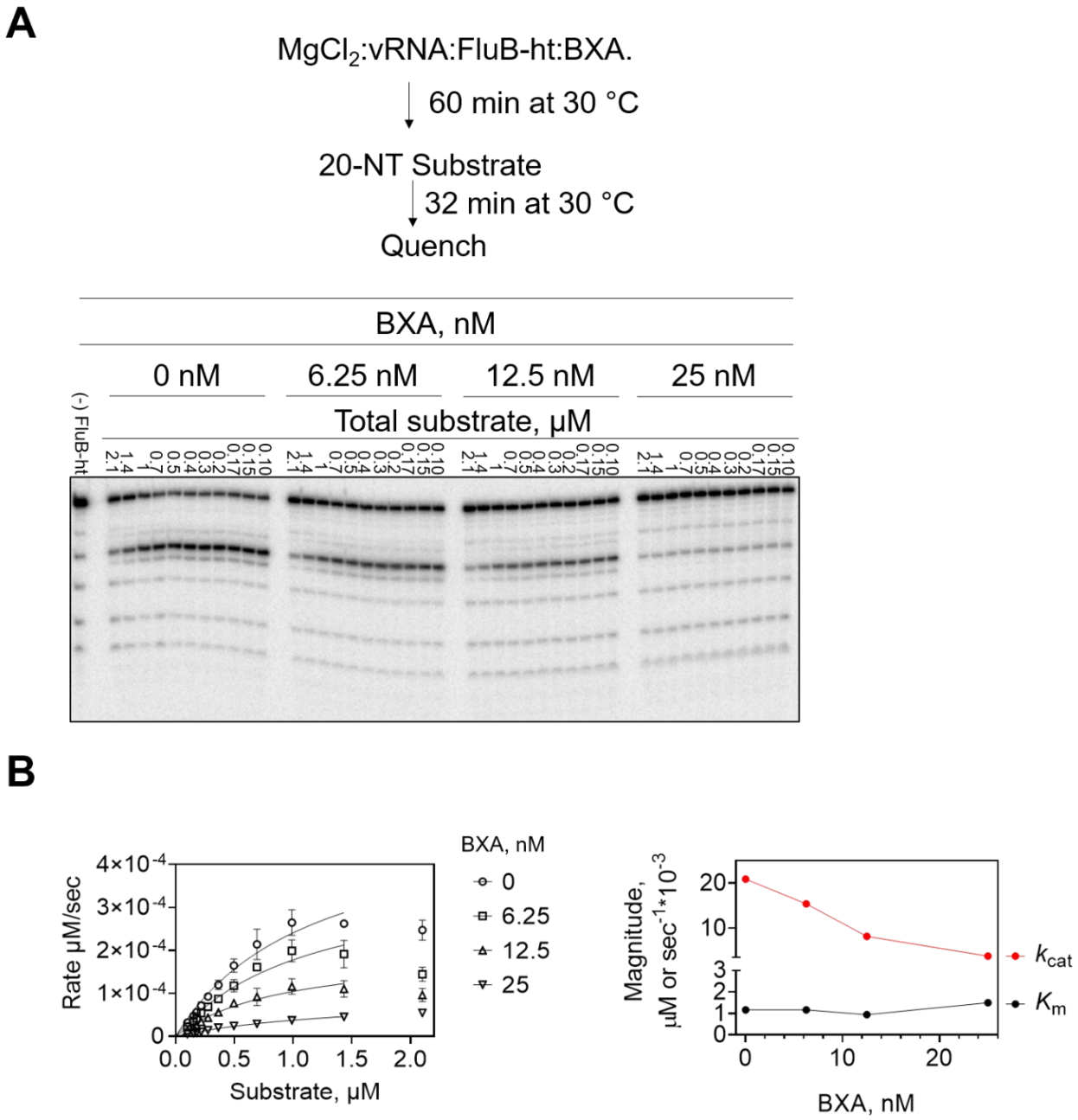


Figure 4.3.1 Michaelis-Menten kinetics show pattern of non-competitive inhibition of BXA. (A) Scheme detailing the experimental setup and representative data showing the effect of BXA on substrate utilization by FluB-ht. (B) Graphic representation and analysis of the data presented in panel A are shown on the left-hand side graph. The effects of increasing concentrations of

baloxavir acid on k_{cat} and K_m are illustrated on the right-hand side graph. Data represents the average of two independent experiments. The final reaction conditions were as follows: 30 mM Tris-HCL pH 7.5, 25 mM NaCl, 5 mM $MgCl_2$, 25 nM FluB-ht, 1.7 μM vRNA, 5% DMSO, BXA as indicated and 20-nt capped substrate as indicated. The indicated substrate concentrations were achieved by mixing 0.1 μM 20-nt radiolabeled capped substrate with unlabeled substrate.

To address this problem, we designed order-of-addition experiments and show that a pre-formed complex of enzyme and RNA substrate prevents inhibition of nuclease activity. Inhibition requires a pre-formed complex of enzyme and inhibitor in the presence of divalent, catalytic metal ions. Based on this data we conclude that the bound inhibitor prevents simultaneous binding of the RNA substrate. Our data further indicate that inhibitor dissociation and association are both slow processes, which are often seen with tight binding inhibitors. The described circumstances are reminiscent of the mechanism of action of HIV-1 RNase H inhibitors (26,44,185). The HIV-1 RNase H active site of reverse transcriptase (RT) also contains two divalent metal ions (37,186). Several RNase H inhibitors that bind the two metal ions have been described (37,187,188). Overlaying structures of HIV-1 RT in complex with a representative RNase H inhibitor and its nucleic acid substrate, respectively, revealed that the RNA would clash with the bound inhibitor (189). Structures of the N-terminal domain of the FluA PA subunit with a bound RNA oligonucleotide and BXA, respectively, also revealed a significant overlap of the two binding sites (190). Thus, BXA or potent RNase H inhibitors prevent RNA binding, at least locally at the nuclease active sites. However, the nucleic acid binding groove is in both cases large enough to allow partial binding of the RNA even in the presence of an inhibitor. Structures of FluB-ht with bound RNA substrate in the groove between PB2 and PA are necessary to address this problem

directly. Overall, the proposed competitive mechanism is in stark contrast to HIV-1 integrase inhibitors that use the DNA substrate for binding (43).

In conclusion, our data provide strong evidence to show that BXA is a tight binding inhibitor that competes with the RNA substrate at the active site of the influenza cap-snatching endonuclease. There are also several study limitations. Tight binding of small molecule inhibitors can involve two steps that comprise the initial complex formation and a subsequent conformational change that stabilizes the complex. The current data do not distinguish between one-step and two-step binding modes. Moreover, inhibitor dissociation and association were here measured only indirectly, which precluded the determination of rate constants. It will be interesting to see how the I38T mutation affects these parameters. Despite these limitations, the measurements of K_i^{app} values enable a quantitative assessment of both cap-snatching endonuclease inhibitors and mutant enzymes with resistance-conferring amino acid substitutions. The proposed mechanism of action of BXA guides future studies aimed at the development of inhibitors with improved properties.

Furthermore, our data show that BXA is selective for the PA endonuclease as 3 μM BXA failed to appreciably inhibit the transcriptional activity of FluB-ht. This concentration is ~ 120-fold greater than the IC_{50} (~25 nM) for the PA endonuclease at this FluB-ht concentration. Therefore, our data show that inhibiting the PA endonuclease (cap-snatching), and not PB1 (The RdRp), is likely the most significant mechanism through which BXA inhibits influenza transcription.

Chapter 5: Characterization of the SARS-CoV-2 proofreading exonuclease nsp14.

5.1 Introduction

In December of 2019, a novel respiratory virus appeared in Wuhan, China (137). Soon after, the virus, which is now named severe acute respiratory virus coronavirus two (SARS-CoV-2), spread worldwide leading the WHO to declare a global pandemic (140). At the time of writing, SARS-CoV-2 has infected 160 million people leading to approximately 3.3 million deaths worldwide. Therefore, there is an urgent need to develop new therapies against SARS-CoV-2. The viral RNA-dependent RNA polymerase or RdRp is a validated drug target as evidenced by the recent approval of the polymerase inhibitor remdesivir by both the FDA and Health Canada for the treatment of severe SARS-CoV-2 disease (138,145,191). Here, we aimed to develop assays to understand the role of the SARS-CoV-2 proofreading exonuclease in resistance to nucleoside analogue antivirals. We envision that these data will further our understanding of the mechanisms underlying the action of and development of resistance to nucleoside analogue antivirals such as remdesivir. This increased understanding would then serve to inform the development of more effective antivirals.

The SARS-CoV-2 viral polymerase complex performs RNA synthesis and is required for viral genome replication and transcription (141). Remdesivir inhibits these processes through a multifaceted mechanism. Firstly, remdesivir functions as an ATP analogue and outcompetes natural ATP for incorporation by the polymerase (141). After incorporation into the growing RNA strand, owing to a steric clash, RNA synthesis is terminated at a site 3 base pairs downstream of remdesivir incorporation (141). However, under high (uM) nucleoside triphosphate concentrations

synthesis can continue until full length product is generated (145). However, once incorporated in the template strand remdesivir provides a barrier to RNA extension by the polymerase. Specifically, RNA synthesis is terminated upon encountering template incorporated remdesivir (145). Thus, it is through a combination of these mechanisms that remdesivir inhibits RNA synthesis by the SARS-CoV-2 RdRp.

However, as nucleotide analogues are, by definition, modified nucleosides it is possible that they could be removed by RNA proofreading mechanisms that are responsible for the removal of misincorporated nucleotides. As coronaviruses encode a proofreading exonuclease complex (nsp10/nsp14), this can potentially represent an avenue by which resistance to antivirals may emerge (28). This is especially likely as the excision of nucleoside analogues has been demonstrated biochemically for the highly related SARS coronavirus (23). The exonuclease allows for the excision and repair of RNA/RNA mismatches increasing the fidelity of RNA genome replication allowing coronaviruses to have larger genomes than most other RNA viruses (24,95). However, a consequence of this is the possibility that nucleoside analogue RdRp inhibitors could be excised limiting their effectiveness as a therapy. To examine this possibility, we expressed and purified the SARS-COV-2 nsp10/nsp14 proofreading exonuclease complex using an insect baculovirus-based protein expression system. Our data suggest that the excision of nucleotide analogues is possible, and that nucleoside analogues with altered base-pairing are preferentially excised. Thus, our results represent the first demonstration of nucleoside analogue excision by NSP-14 and the first demonstration of “proofreading” in SARS-CoV-2. Thus, our results have implications for both our understanding of coronavirus RNA mismatch repair and the utility of nucleoside analogue therapies for the management of infection by SARS-CoV-2.

5.2 Results

5.2.1 Expression and purification of wt and active site mutant SARS-CoV-2 nsp 14/nsp10 exonuclease Complex (ExoN).

We expressed the SARS-CoV-2 RdRp complex (nsp7/nsp8/nsp12) using an insect baculovirus-based expression system as described previously (192-194). In our hands, this system has been highly suitable for the expression of protein complexes. Therefore, we used this system to express the SARS-CoV-2 Exonuclease complex nsp14/nsp10 (ExoN). To purify the ExoN complex, we made use of NiNTA affinity chromatography followed by gel filtration (Figure 5.1). This approach yielded binary complexes of nsp10/nsp14. We also expressed mutant enzymes with amino acid substitutions in the ExoN active site of nsp14 (D90N, E92Q, E191Q, D273N). These residues are responsible for coordinating magnesium ions in the exonuclease active site of nsp14 from SARS CoV (28). As nsp14 is highly conserved between SARS CoV and SARS-CoV-2 we hypothesized that mutations of these residues would also result in the loss of exonuclease activity in SARS-CoV-2 nsp14. In line with this hypothesis, we observed a complete loss of activity for the mutant (Fig. 5.2 and 5.3). This is strong evidence that the observed exonuclease activity is due to NSP-14 and not a co-purified contaminant.

Furthermore, by size exclusion chromatography we obtained complexes of approximately ~80 KDA. These complexes correspond to a 1:1 complex of nsp 10 and nsp 14 and are analogous to those which have been described by others for the highly related SARS coronavirus proofreading exonuclease complex (Figure 5.1 and 5.2) (23,28). As the amino-acid sequences are highly conserved between SARS-CoV-2 nsp 10/nsp14 and their homologues from SARS-Cov-1 it is likely that the complex we observed represents a 1:1 complex and this is supported by our gel

filtration chromatography data (Figure 5.2). However, structural studies involving nsp14 from SARS-CoV-2 are currently lacking and will be required to confirm this finding.

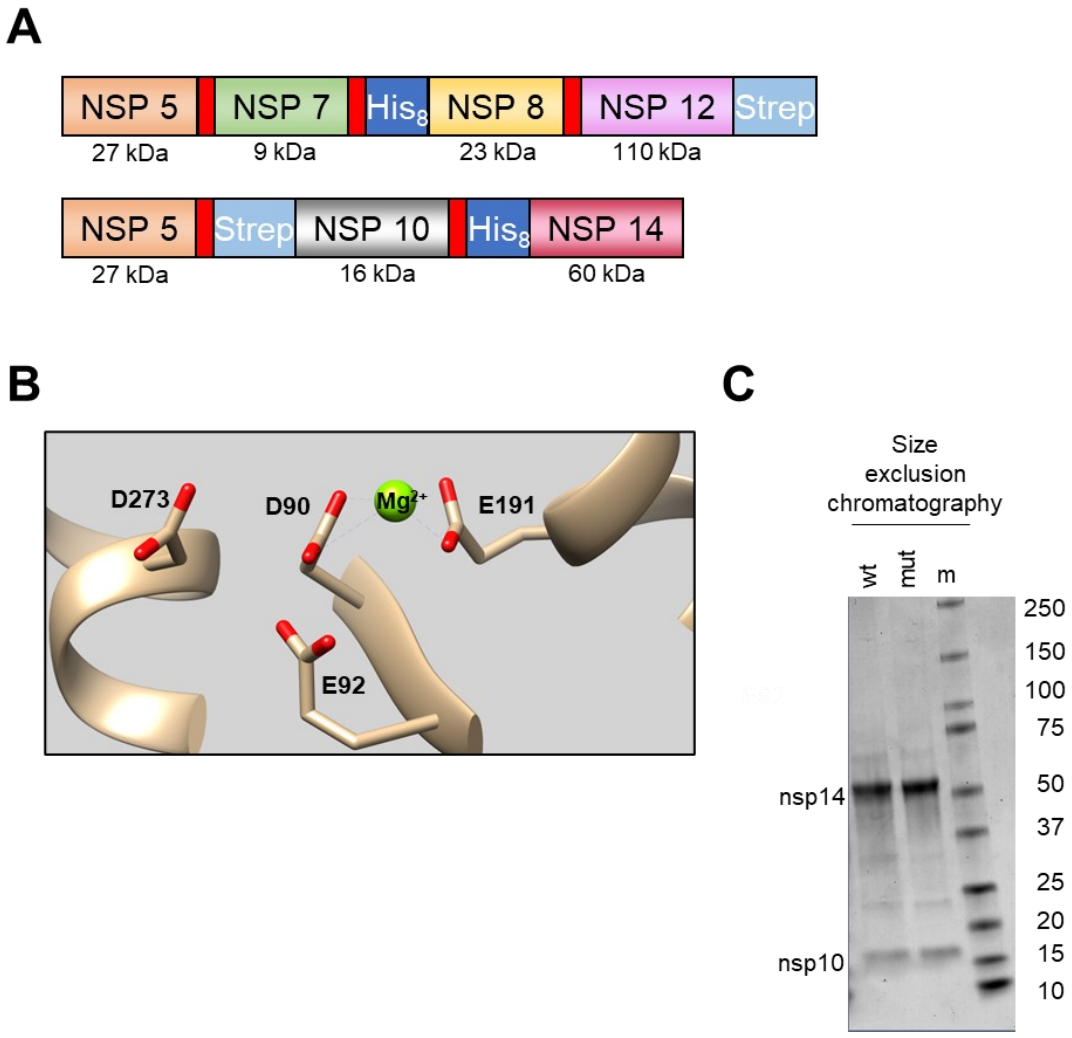
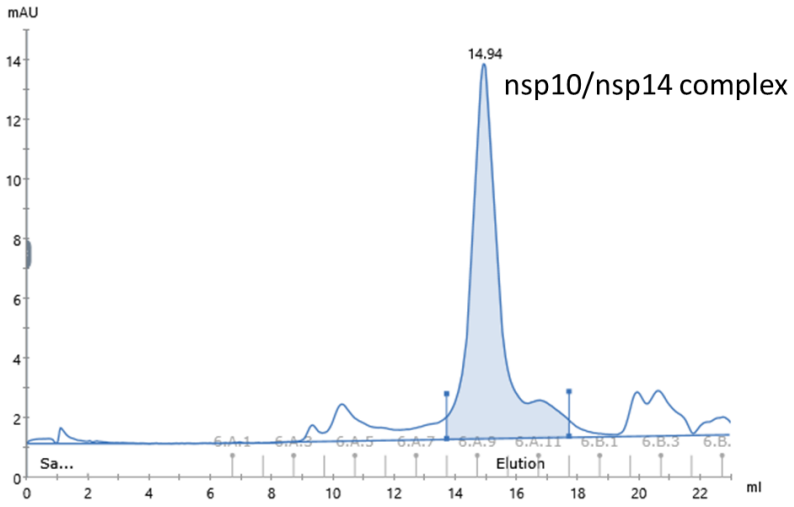


Figure 5.1. Expression, purification, and characterization of the SARS-CoV-2 nsp10/nsp14 exonuclease complex (ExoN). (**A**) Schematic representation of the expression constructs used to produce SARS-CoV-2 RdRp polymerase and ExoN complexes. Genes coding for non-structural

proteins, molecular weight in kDa, and affinity tags are indicated. Red rectangles indicate nsp5 cleavage sites. His_8 and Strep indicate the location of the histidine or strep tags, respectively. (**B**) A snapshot of the crystal structure of the ExoN nsp14 protein (PDB: 5c8u) showing active site residues that were mutated to generate an ExoN active site mutant. (**C**) SDS-PAGE migration pattern of the Ni-NTA/SEC-purified ExoN wild type (wt) and D90N/E92Q/D273N/E191Q nsp14 active site mutant (mut) enzyme preparations stained with Coomassie Brilliant Blue G-250 dye. Based on mass spectroscopy data the proteins migrating at ~ 60 and 16 kDa correspond to nsp14 and nsp10, respectively (Appendices 1). The identity of these proteins was confirmed by LC-MS/MS (appendices 1). Emma Woolner assisted in the purification of the nsp14/nsp10 protein complexes by size exclusion chromatography. Dr. Egor Tchesnokov assisted in the preparation of the figure.

A.



B.

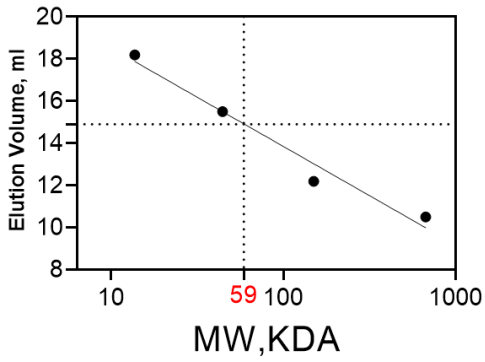


Figure 5.2. Size exclusion chromatography (SEC) analysis of the SARS-CoV-2 nsp10/nsp14 exonuclease complex. A. SEC chromatogram for wt nsp14/nsp10. B. SEC chromatogram for molecular weight standards of known size. The position at which nsp10/nsp14 eluted is indicated by the intersection of two broken lines and its relative size, 60 KDa, indicated on the X-axis. SEC chromatography was performed in collaboration with Emma Woolner.

5.2.2 Excision of adenosine triphosphate (ATP), remdesivir triphosphate (RDV), or favipiravir triphosphate incorporated at position 6 by the SARS CoV-2 nsp 10/nsp14 exonuclease complex.

To assay for the excision of NTPs and NTP analogues we first made use of a primer-template system that utilizes radiolabeled GTP as a signal generator. In short, this system allows for the insertion of ATP, RDV, or favipiravir immediately after the incorporation of radiolabeled GTP (Fig 5.4A). This leads to a 6 nucleotide (nt) product terminating with either ATP, RDV, or Favipiravir. The reasons why we chose to utilize this system are two-fold. Firstly, we chose to utilize this system as we had recently established this system for the study of RDV incorporation by the SARS-CoV-2 RdRp (141). Secondly, we chose this system as it would allow us to study the excision of incorporated nucleotides by the exonuclease while the polymerase is in the initiation phase of RNA synthesis. Structures of RdRp(s) in the initiation phase have been solved with 6-nt transcripts (195). Therefore, using conditions that halt the RNA-polymerase at 6-nt allows us to study the excision of nucleotides under conditions resembling the initiation phase of RNA synthesis (Figure 5.4). This is in contrast to later experiments where we produced complexes containing 12-nt of newly synthesized RNA, resembling the elongation phase.

The analogues used in this thesis were chosen to examine the effect of mismatched base pairs on the excision of nucleotides by nsp14. This was necessary because we were unable to optimize a system that allows for the misincorporation of a natural nucleotide. Specifically, RDV was chosen because it is structurally similar to the natural nucleotide ATP while favipiravir was selected because it base pairs in such a way that it represents a mismatch (Figure 5.3) (196). Specifically, favipiravir basepairs ambiguously allowing it to function as either an A or G analogue and because of this has been hypothesized to exert its antiviral action through a mechanism known as lethal

metagenesis (196). That is to say that incorporation of favipiravir into template RNA strand leads to catastrophic mutagenesis over successive rounds of replication. Therefore, incorporated favipiravir can be said to be functionally similar to a misincorporated base pair. Due to this, we hypothesized that the excision of favipiravir by the nsp14 proofreading exonuclease would be more efficient than ATP or RDV and this is what was observed (Figure 5.4 and 5.5). Specifically, after generating this product we added varying concentrations of the nsp10/nsp14 complex to the reaction mixture. Under increasing concentrations of nsp10/nsp14, we observed increased excision of ATP, RDV, and favipiravir. However, we observed that favipiravir was more readily excised than either ATP or RDV (Figure 5.4B, D). When we utilized an exonuclease deficient mutant where acidic residues known to coordinate divalent metal ions were mutated to asparagine or glutamine no excision was observed (Figure 5.4C). This suggests that the excision we observed is due to the addition of the nsp10/nsp14 complex and not a contaminating ribonuclease. Furthermore, as favipiravir and remdesivir are excised these data demonstrate that the excision of nucleoside analogues by nsp14 is chemically possible. A question that is highly relevant given the fact that some authors have suggested that RDV should be resistance to excision based on *in silico* analysis (197).

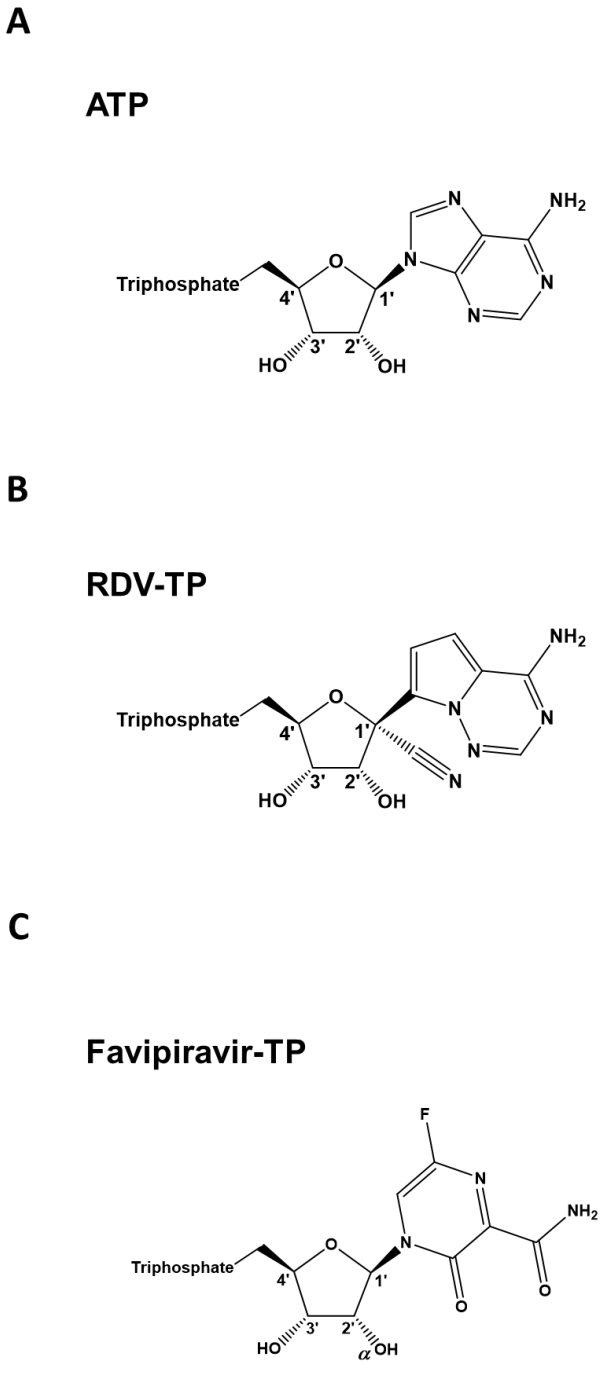


Figure 5.3 Structures of the nucleoside analogues used in the study. (A) Structure of Adenosine triphosphate (ATP). (B) Structure of Remdesivir Triphosphate (RDV-TP). (C) Structure of Favipiravir Triphosphate (Favipiravir-TP).

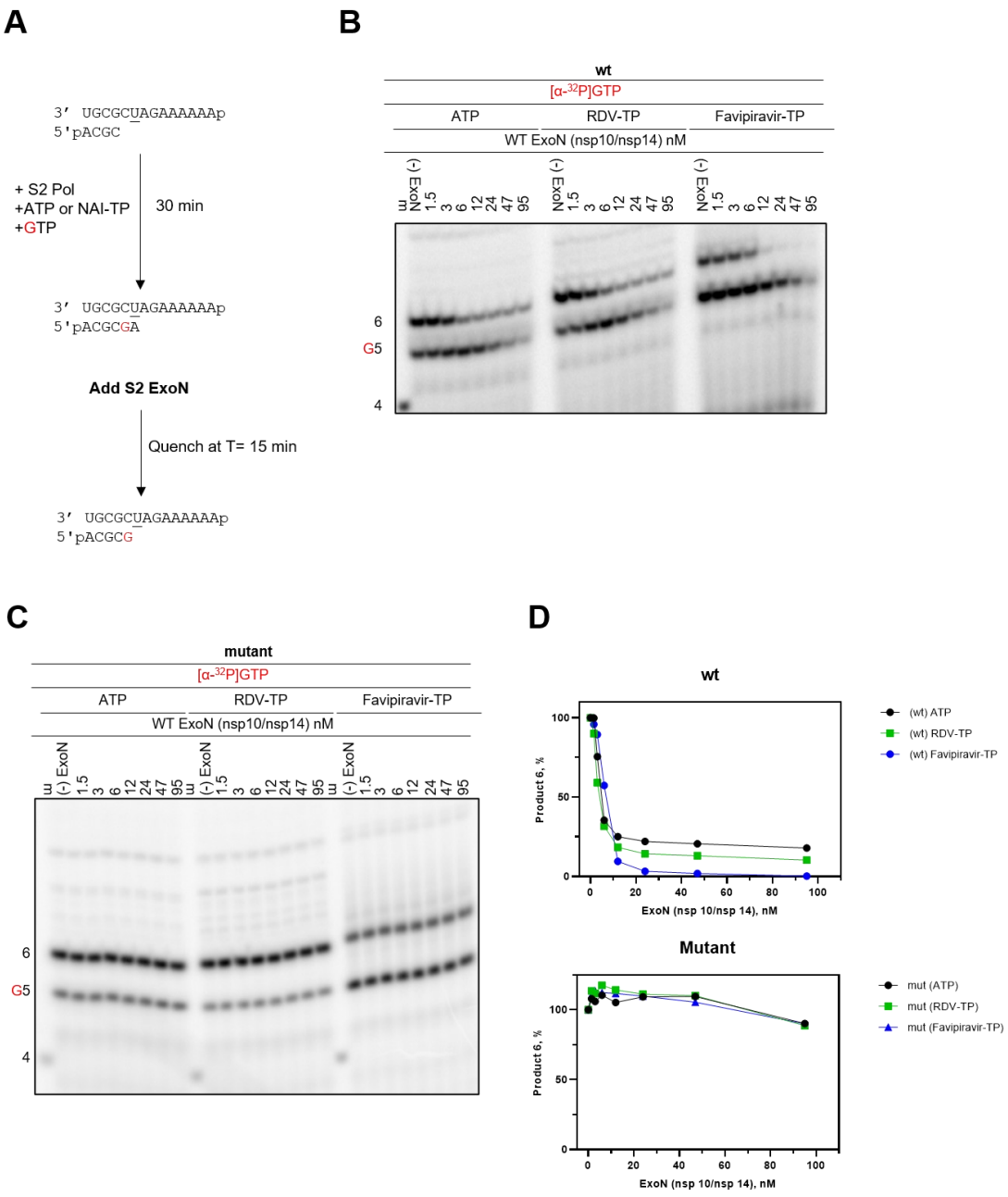


Figure 5.4 Excision of ATP or Nucleoside Analogue Inhibitors by wt and active site mutant ExoN complexes using a model primer/template system. **(A)** Schematic representation of the exonuclease assay. Both RNA template and primer are 5'-phosphorylated (p). RNA primer/template is incubated at 30 °C with SARS-CoV-2 RdRp complex in the presence of $MgCl_2$, [α-^{32}P]-GTP, UTP and one of either ATP, Remdesivir-TP (RDV-TP), or Favipiravir-TP for 30

minutes to allow the RdRp complex to the extent the RNA primer up to position 6. Various concentrations of ExoN complex are then added to the mixture and the reaction is allowed to proceed for 15 minutes. (B) Migration pattern of the products of the Exonuclease assay with wildtype (wt) ExoN complex. Lane m illustrates the migration pattern of the radiolabeled 4 nucleotide primer. Lanes (-) ExoN contain reaction products of RNA synthesis by RdRp before the addition of the ExoN complex (C) Migration pattern of the products of the exonuclease assay utilizing the ExoN complex active site mutant. (D) Graphic representation of product 6 (Panels B and C) as a function of ExoN complex concentration 15 minutes after the addition of ExoN complex. . Dr. Egor Tchesnokov assisted in the preparation of the figure.

5.2.3 Excision of adenosine monophosphate (AMP) or nucleoside analogues incorporated at position 12 by the SARS CoV-2 nsp10/nsp14 exonuclease complex.

We had previously demonstrated that nsp14 can act on small primers to excise incorporated nucleoside analogues (Figure 5.4). However, while favipiravir, which represents a mismatch, was the most readily excised the difference was small and we also saw the excision of correctly paired natural nucleotides (Figure 5.4). When we generated longer primers terminating with either AMP, RDV, or favipiravir this is no longer the case. Under these conditions, we observed efficient excision of favipiravir but not AMP or RDV (Figure 5.5 and 5.6). Thus, it is likely that our results mimic a biologically relevant scenario where mismatched base pairs, i.e.,favipiravir, are readily excised while natural nucleotides or nucleoside analogues that base pair similarly to a natural nucleotide (i.e., RDV) are not excised. Thus our results here support a proofreading function for nsp14 in SARS-CoV-2 as has been suggested elsewhere (95).

Interestingly, however, when RDV is incorporated followed by three additional base pairs, a situation mimicking the results of delayed chain termination, excision occurs (Figure 5.6) Thus, I hypothesize that the products of delayed chain termination may produce a structure similar to that which occurs in the presence of a misincorporated base pair and that this structure may drive excision by nsp14.

It has also been suggested by others that the nitrile group present in remdesivir may reduce the efficiency at which it is excised relative to natural nucleotide (197). Our data suggest that this hypothesis is incorrect as RDV and ATP are both resistant to excision at position 12 and excised with relatively equivalent efficiency at position six (Figures 5.4, 5.5, and 5.6). Thus, it does not appear that the nitrile group of RDV prevents any significant chemical barrier to excision. This is also supported by the fact that the generation of a delayed chain termination product containing RDV at position 9 and terminally at position 12 was also readily excised by nsp14.

When RDV is incorporated and the subsequent nucleotides are provided at low enough concentrations the polymerase will stall at a position three nucleotides downstream of the site of RDV incorporation (141). This is referred to as delayed chain termination and in the case of remdeisvir is thought to result from a steric clash between Serine 861 of the polymerase and the 1' nitrile group of remdesivir leading to stalling (141,198). We observed that a product containing RDV at position 9 followed by 3 additional nucleotides was readily cleaved by the nsp10/nsp14 exonuclease complex.

Thus, given that the excision of RDV is chemically possible, we hypothesized that the differences in excision that we have seen with 12-nt products representing chain terminated products (RDV), terminally incorporated RDV (inefficient excision), favipiravir (excised), or a 12-nt product

containing only natural nucleotides (inefficient excision) is due to either differences in the translocation state of the polymerase complex or to dissociation of the polymerase complex.

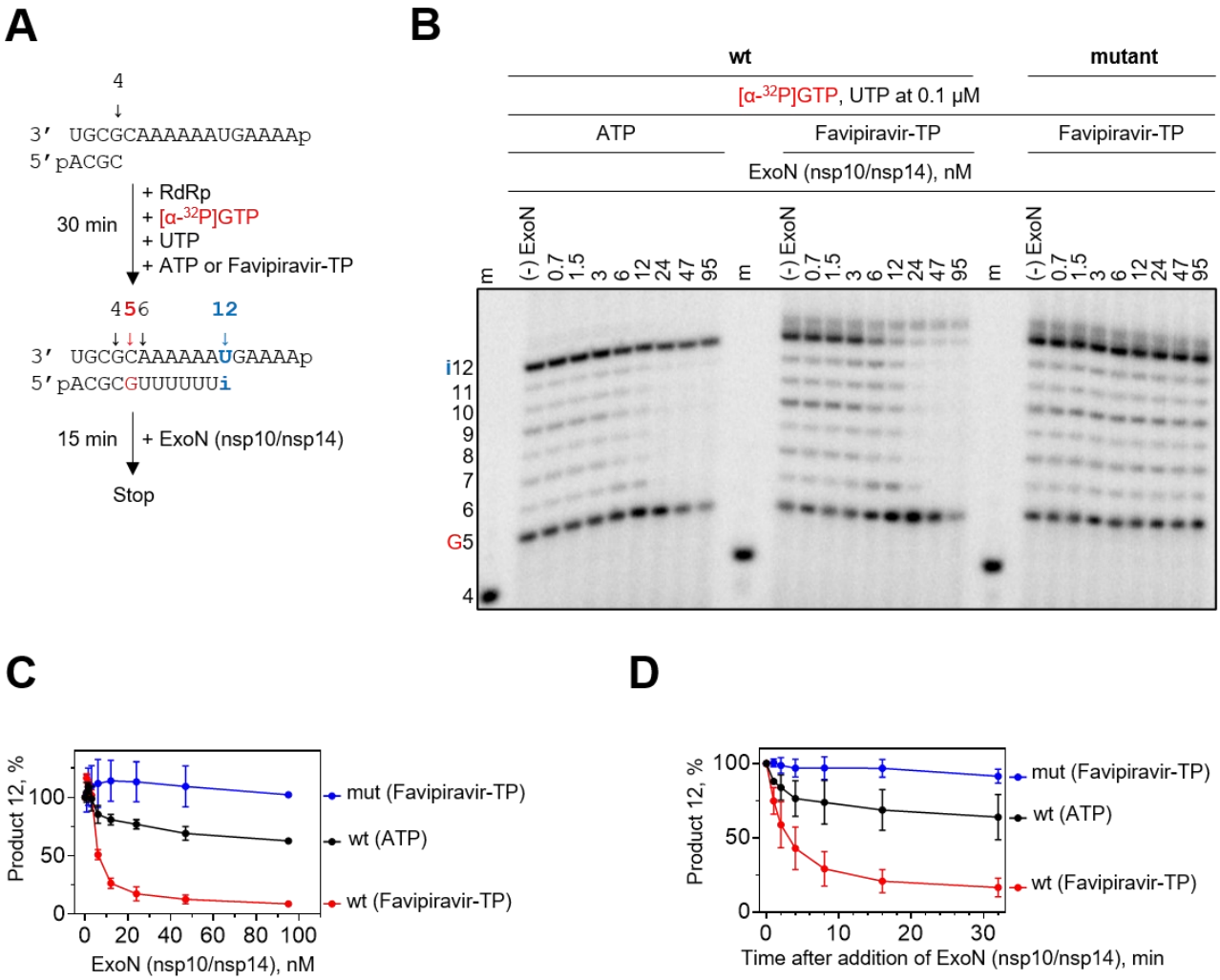


Figure 5.5 Exonuclease activity of wt and active site mutant ExoN complexes. (A) Schematic representation of the exonuclease assay. Both RNA template and primer are 5'-phosphorylated (p). RNA primer/template is incubated at 30˚C with SARS-CoV-2 RdRp complex in the presence of $MgCl_2$, [α-^{32}P]-GTP, UTP, and ATP or RDV-TP for 30 minutes to allow the RdRp complex to

extend RNA primer up to position 12 (i). ExoN complex is then added to the reaction mixture and the reaction is allowed to proceed for up to 15 minutes. (B) The migration pattern of products of exonuclease assay schematically depicted in panel A. Lanes m illustrates the migration pattern of the radiolabeled 4-nucleotide-long primer. Lanes (-) ExoN contains the reaction products of RNA synthesis by the RdRp complex before the addition of the ExoN complex. RNA synthesis was conducted for 30 minutes in the presence of indicated concentrations of GTP and UTP supplemented with ATP or Favipiravir-TP followed by the addition of various concentrations of ExoN wt and mutant complexes for another 15 minutes. (C), (D) Graphic representation of product 12 (panel B) as a function of ExoN complex concentration and as a function of time after the addition of ExoN complex, respectively. .Dr. Egor Tchesnokov assisted in the preparation of the figure.

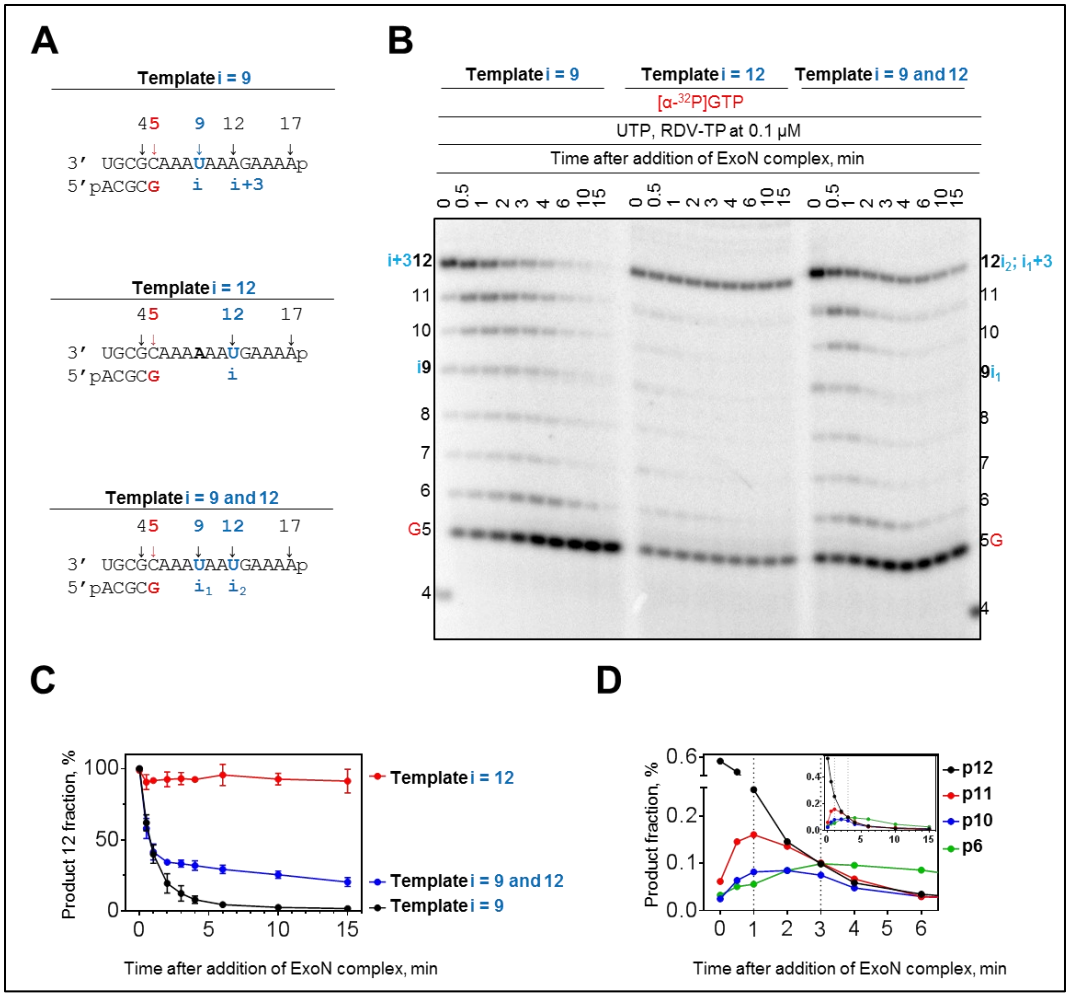


Figure 5.6 Excision of incorporated remdesivir as a function of its position in the template. (A) RNA primer/templates supporting incorporation of remdesivir (i) at position 9 (i = 9), position 12 (i = 12), and at positions 9 (i_1) and 12 (i_2) (i = 9 and 12). i+3 illustrates a delayed chain termination at position 12 when remdesivir is incorporated at position 9. (B) Incorporation of remdesivir based on the indicated primer/templates was catalyzed by CoV-SARS-2 RdRp complex in the presence of $MgCl_2$ and indicated concentrations of NTP and RDV-TP for 30 minutes (lane 0). ExoN complex was then added to the reaction mixture and aliquots withdrawn at indicated time points. 4 indicates the migration pattern of 5'-^{32}P-labeled 4-nt primer used here as a marker. G5 indicates the incorporation of [α-^{32}P]-GTP opposite template position 5. (C) Graphic representation of the product 12 fraction as a function of incubation time with ExoN complex.

Error bars represent the standard deviation of the data within three independent experiments involving two independent enzymes (RdRp and ExoN complexes) preparations. (D) Graphic representation of the various product fractions as a function of incubation time with ExoN complex. These experiments were done in collaboration with Dr. Egor Tchesnokov and Dr. Tchesnokov also assisted in the preparation of the figure.

5.2.4 Effect of Templated RDV on the excision by nsp14.

We previously observed correctly paired nucleotides or nucleoside analogues that are structurally similar to natural nucleotides (RDV) appeared to be resistant to excision compared to those that represent a mismatch (favipiravir). That is to say that, under the conditions we tested, it appeared that a mismatched nucleotide (such as favipiravir), can function as the switch that initiates excision by nsp14. Thus, we were interested in what would occur if instead of assaying for the excision of a nucleoside analogue from the primer strand, we instead looked at the excision of a natural nucleotide (U) incorporated across from templated RDV. We hypothesized that UTP would not be readily excised as RDV in the primer strand at this position is not excised. This is what we observed, suggesting that RDV: Uracil base pairs represent a natural base pair and do not lead to significant activation of the exonuclease (Figure 5.7). The notable exception being in the case of delayed chain termination where rapid excision is observed.

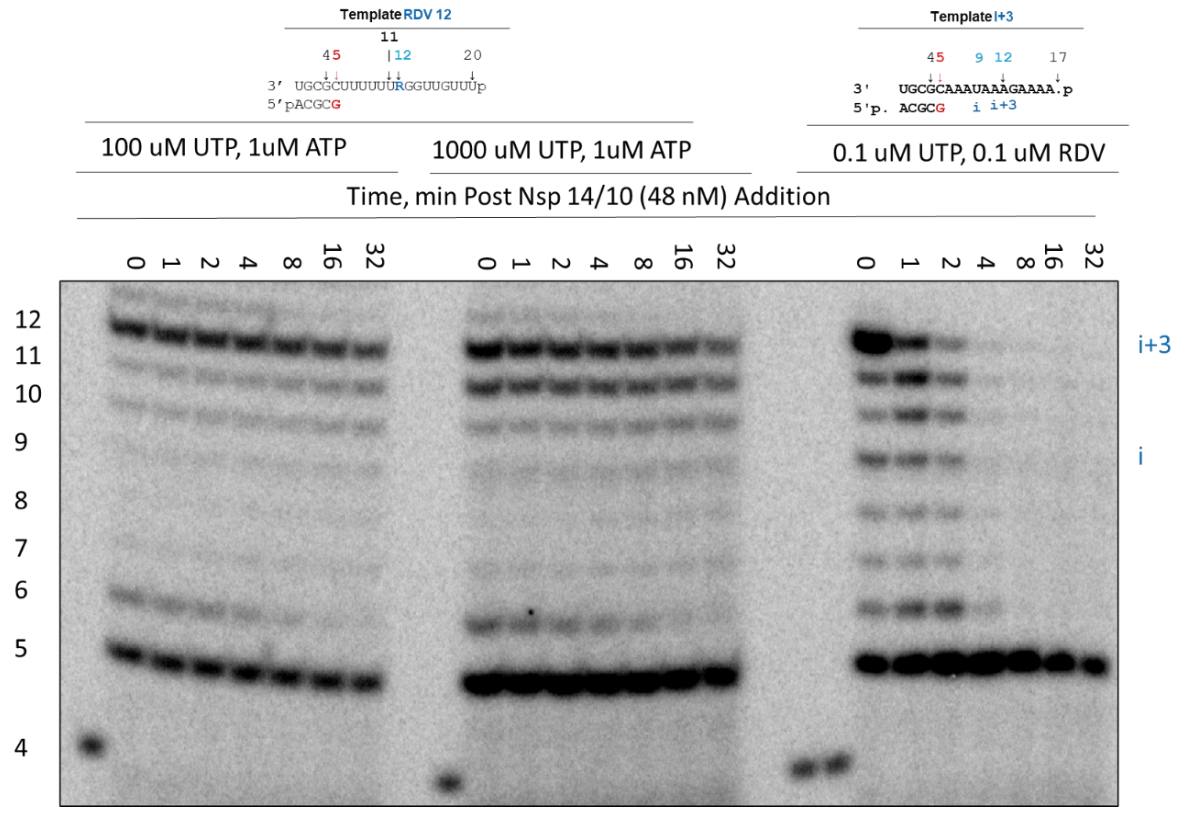


Figure 5.7 Incorporated UTP across from Templated RDV is not efficiently excised. Primers terminating with UTP across from RDV in the template, or the traditional products of delayed chain termination which served as the positive control, were generated, and exposed to 48 nM nsp 14/nsp10. Primer/template combinations used along with the nucleotide concentrations used for RNA synthesis are provided above the gel image. Inefficient excision of U paired across from RDV was observed. This is in stark contrast to the products of delayed chain termination which are readily excised.

5.3 Discussion

The excision of misincorporated nucleotides allows for increased viral replication fidelity and may have allowed coronaviruses to have larger genomes than many other RNA viruses (24,95). Most

FDA-approved antiviral medications are nucleoside analogues. Therefore, the presence of a proofreading exonuclease enzyme provides a potential pathway for resistance to these medications by SARS-CoV-2. Specifically, if nucleoside analogues can be readily excised by nsp14 at a rate that is higher than the rate at which they are incorporated into the genome these therapies will be ineffective. Therefore, there is an interest in examining the ability of nsp14 to excise existing nucleoside analogues such as RDV as a characterization of the mechanism of excision of these molecules will potentially allow for the development of more effective antivirals by elucidating the characteristics of nucleoside analogues which lead to inefficient excision.

Here, we have shown that the excision of natural nucleotides and nucleoside analogues such as RDV and favipiravir by nsp14 is chemically possible. This contradicts previous modeling studies which suggested that RDV, due to the presence of its nitrile group, would be highly resistant to excision by nsp14 (197). Our data have also shown that nsp 14 favors the excision of products that represent mismatched bases. Specifically, we have shown under all conditions tested that favipiravir, a nucleoside analog that represents a mismatched base pair is excised more readily than either RDV (an analogue that base pairs very similarly to ATP) or natural ATP.

Thus, our data support a model by which NSP-14 is activated by RNA structures corresponding to misincorporated bases (i.e., favipiravir) (196) , or by a kinked RNA strand as occurs in the case of delayed chain termination caused by remdesivir (141,145) (Figure 5.8). Specifically, in the case of delayed chain termination by RDV the RNA strand becomes kinked such that the 1’ nitrile group of remdesivir clashes with Serine 861 in the interior of the viral polymerase resulting in stalling (141,145,198). This stalling is hypothesized to lock the polymerase into the pre-translocation state (145). Stalling in the pre-translocation state is also known to occur after base misincorporation (199). As, the pre-translocation state is known to facilitate strand transfer to the

exonuclease in DNA polymerases, we hypothesized that RDV I+3 and favipiravir are rapidly excised because they may shift the viral polymerase into the pre-translocation state facilitating strand transfer to the exonuclease (99). However, in order to verify that this is the case further experiments will be required to ensure that the nsp10/nsp14 complex and the viral polymerase interact. This could be accomplished by making use of experiments utilizing high concentrations of heparin.

Specifically, by incubating the reaction components in the presence of heparin prior to the initiation of the reaction any enzymes that dissociate from the RNA template will be bound by the highly charged heparin molecules and inactivated (42). This approach has been previously used by our group in the study of HIV-1 reverse transcriptase and could be used here to ensure that patterns of exonuclease activity that we observe here are due to strand transfer between the SARS-CoV-2 RdRp and nsp10/nsp14 exonuclease complex and not simply due to the dissociation of the polymerase and interaction of the newly liberated substrate with free nsp0/nsp14.

Nevertheless, these data, the resulting models, and the assays developed here provide insight into the mechanisms of RNA proofreading and nucleoside analogue drug resistance in SARS-CoV-2. Specifically, these data suggest that the excision of nucleosides is chemically possible and that this is a viable mechanism by which resistance to nucleoside analogues may emerge within SARS-CoV-2.

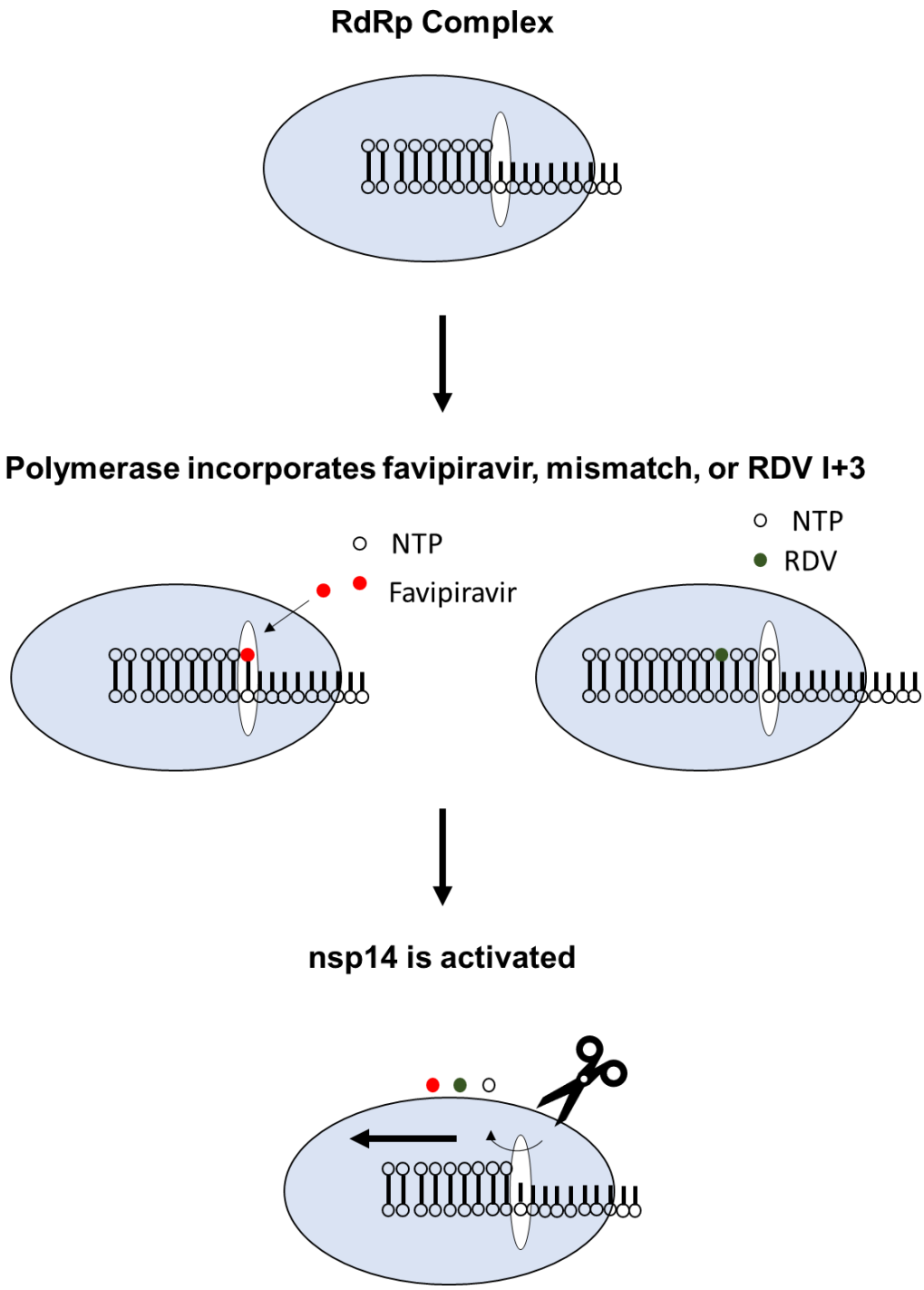


Figure 5.8 Model detailing nucleotide excision by nsp14. Favipiravir, mismatched nucleotides, or the RNA structure produced by RDV I+3 activate the nsp14 proofreading exonuclease and lead to the excision of incorporated nucleotides and nucleotide analogues.

Chapter 6: Summary and Future Directions

6.1 Summary

Viral nucleases are essential to the lifecycles of many medically relevant viruses including CMV, HIV, Influenza, and SARS-CoV-2 (25,26,36,90,197). Furthermore, nucleases are validated targets for the development of antivirals as evidenced by the fact that the study of these enzymes has led to the FDA approval of several antiviral agents including Letermovir (CMV), Baloxavir (Influenza), and several Integrase inhibitors (i.e., raltegravir, HIV) (30,40,91). However, biochemical assays to study the action of and resistance to viral nuclease inhibitors were lacking at the time that I began work on this thesis. To address this, we worked to both express viral nucleases using recombinant protein expression systems and develop biochemical assays to study the enzymology of these proteins and the mechanism of action of inhibitors.

To this end, in **chapter 3** I successfully purified the CMV terminase complex and developed assays to measure its nuclease, ATPase, and DNA binding activities. These assays are amenable to the characterization and refinement of CMV, and herpesviral, terminase inhibitors in general and may allow for the identification of novel inhibitors *in vitro*.

In **chapter 4** I successfully purified the Influenza B RdRp heterotrimer and developed assays to study cap snatching and transcription. These assays revealed that the active form of the influenza endonuclease inhibitor baloxavir marboxil, baloxavir acid, is a slow-tight binding competitive inhibitor of the influenza PA endonuclease. Furthermore, these assays revealed that baloxavir selectively inhibits the PA endonuclease rather than the PB1 RdRp active site. Additionally, the methodologies developed here provide a framework by which novel tight binding inhibitors can

be identified and characterized. Such novel inhibitors are urgently required given the threat posed by the emergence of influenza strains resistant to currently available antivirals.

In **chapter 5** I successfully purified the SARS-CoV-2 proofreading exoribonuclease complex nsp14/nsp10. I discovered that this complex could excise incorporated nucleotide analogue inhibitors, including remdesivir which, at the time of writing, is the only FDA approved small molecule inhibitor for the treatment of SARS-CoV-2 infection. Furthermore, I was able to show that inhibitors that resemble base pair mismatches are more readily excised by the complex providing biochemical evidence for the proofreading function that has been hypothesized to be encoded by the nsp14 exonuclease. This provides a biochemical basis for the increase in mutation frequency observed by others when this protein is mutated, in related viruses, to render it nonfunctional (93). Furthermore, the assays developed here provide a platform by which both inhibitors of nsp14 and nucleoside analogue polymerase inhibitors resistant to excision can be characterized and refined.

6.2 Future Directions

There are several potentially interesting future directions to pursue with regards to the research presented in this thesis. These directions are especially exciting given the importance of developing better antivirals in the context of being able to combat future viral pandemics.

Firstly, much remains to be done with regards to characterizing the CMV terminase complex and understanding the mechanism of action of letermovir. Specifically, we were unable to generate constructs where the NTPase or nuclease activities of the terminase were knocked out. Thus, it is possible that the activities observed in chapter 3 are due to copurified contaminants. Thus, it will be necessary to refine the expression and purification processes. This could perhaps be accomplished by adding additional purification steps, such as ionic exchange chromatography or

ammonium sulfate precipitation. If the addition of these steps allows the purification of terminase variants that are inactive with respect to the NTPase or exonuclease activities one would then be able to use the assays presented in this thesis to finalize a biochemical characterization of full length CMV terminase. Once this is completed, mechanisms of letermovir action and resistance could then be explored. Furthermore, the assays developed in this chapter are, at the time of writing, the only biochemical assays available for the in-vitro study of the CMV terminase. For this reason, the assays described in this chapter represent a significant advancement in that they provide a platform by which novel anti-CMV inhibitors can be identified *in vitro* and mechanisms of drug action and resistance characterized.

Additionally, the ATPase and nuclease activities of the CMV terminase complex have yet to be unambiguously assigned to terminase subunits. One major reason surrounding this is the fact that to date, no structure of the full length CMV terminase complex has been solved. However, the structure of the homologous HSV-1 terminase complex was recently solved using cryo-electron microscopy (cryo-EM) (76). Given that our CMV terminase preparations are similar in purity to the HSV-1 terminase preparations showcased in this publication it is likely that the expression protocols that we developed for the CMV terminase complex are suitable for structure determination via cryo-EM. Given this, it would be useful to work towards the generation of structures of the CMV terminase complex in complex with metal ions, nucleic acid, and ATP. These structures would indicate the residues involved in the ATPase, nuclease, and DNA binding activities of the complex and would provide us with the necessary information to make mutants to eliminate these activities. Additionally, depending on these results, this structural approach would allow us to detect the presence of non-canonical ATPase and nuclease domains that cannot yet be

predicted via bioinformatic approaches. Therefore, this work is of paramount importance if we aim to unambiguously assign the enzymatic activities of the CMV terminase to individual subunits. This approach could also be supplemented with reverse genetic approaches making use of CMV bacmid systems which have previously been developed (159). Specifically, we could generate constructs containing the various mutations in the ATPase and nuclease domains that have been proposed to be present in both Ul89 and UL56. As cleavage and packaging are essential to the vial lifecycle the recovery of infectious progeny would suggest that these domains are not required for viral concatemer resolution and packaging.

Secondly, with regards to our characterization of the biochemical mechanism of action of baloxavir acid we measured binding only indirectly precluding the determination of rate constants. Future work could address this using either surface plasmon resonance or isothermal calorimetry titration to determine constants for BXA binding. These data would serve to strengthen our conclusion that BXA binding prevents substrate binding by binding to the active site, and therefore, BXA, acts as a competitive inhibitor of the influenza PA endonuclease. Furthermore, this work represents the first system developed for the study of tight binding inhibitors of the influenza PA "cap-snatching" endonuclease. It is highly probable that this system could be expanded to identify other tight binding inhibitors that function by a similar mechanism. This is especially important given that resistance to baloxavir marboxil has been described in the clinic and novel antivirals targeting PA could, potentially, address this problem.

Finally, with regards to our study of the SARS-CoV-2 nsp10/nsp14, we were able to demonstrate that the SARS-CoV-2 proofreading exonuclease nsp14 can excise incorporated nucleotide analogue inhibitors and that those resembling basepair mismatches appear to be more readily

excised. However, a model addressing the biochemical mechanism of action underlying these observations is currently lacking. Specifically, it is not clear whether our observations are a result of changes in the translocation state of the polymerase which pass the 3' end of the substrate to the exonuclease, or simply due to dissociation of the polymerase from newly synthesized RNA. To address these questions, it will be necessary to perform heparin trap experiments. High-concentrations of heparin will inactivate DNA binding proteins that are not actively engaged with substrate (200). This would allow us to determine whether the observed results are due to dissociation of the polymerase.

Additionally, it would be interesting to optimize conditions such that the switch from excision back to extension can be captured. If this can be done this would, in essence, represent the first biochemical demonstration of "proof-reading" and error correction in SARS-CoV-2.

Finally, the systems described in chapter 5 provide a platform by which both nucleoside analogue inhibitors resistant to excision by nsp14 and the inhibition of the nsp14 proofreading exonuclease can be studied. The existence of these assays opens two related avenues for the development of antivirals against SARS-CoV-2 which may be vital in combating the ongoing pandemic. Firstly, modifying nucleoside analogues so that they are resistant to excision by nsp14 may increase their efficacy provided that these analogues are still incorporated efficiently by the viral RNA dependent RNA polymerase. Secondly, the exonuclease activity of nsp14 is essential to viral replication (24). Therefore, the identification of inhibitors of this process may lead to new small molecule antivirals for the treatment of SARS-CoV-2 infection.

Works Cited

1. Taubenberger, J. K., and Morens, D. M. (2006) 1918 Influenza: the mother of all pandemics. *Emerg Infect Dis* **12**, 15-22
2. Reid, A. H., Taubenberger, J. K., and Fanning, T. G. (2004) Evidence of an absence: the genetic origins of the 1918 pandemic influenza virus. *Nature reviews. Microbiology* **2**, 909-914
3. Taubenberger, J. K., Reid, A. H., Lourens, R. M., Wang, R., Jin, G., and Fanning, T. G. (2005) Characterization of the 1918 influenza virus polymerase genes. *Nature* **437**, 889-893
4. Ye, Z. W., Yuan, S., Yuen, K. S., Fung, S. Y., Chan, C. P., and Jin, D. Y. (2020) Zoonotic origins of human coronaviruses. *Int J Biol Sci* **16**, 1686-1697
5. Sharp, P. M., and Hahn, B. H. (2010) The evolution of HIV-1 and the origin of AIDS. *Philos Trans R Soc Lond B Biol Sci* **365**, 2487-2494
6. Latinne, A., Hu, B., Olival, K. J., Zhu, G., Zhang, L., Li, H., Chmura, A. A., Field, H. E., Zambrana-Torrelio, C., Epstein, J. H., Li, B., Zhang, W., Wang, L. F., Shi, Z. L., and Daszak, P. (2020) Origin and cross-species transmission of bat coronaviruses in China. *Nat Commun* **11**, 4235
7. Mari Saez, A., Weiss, S., Nowak, K., Lapeyre, V., Zimmermann, F., Dux, A., Kuhl, H. S., Kaba, M., Regnaut, S., Merkel, K., Sachse, A., Thiesen, U., Villanyi, L., Boesch, C., Dabrowski, P. W., Radonic, A., Nitsche, A., Leendertz, S. A., Petterson, S., Becker, S., Krahling, V., Couacy-Hymann, E., Akoua-Koffi, C., Weber, N., Schaade, L., Fahr, J., Borchert, M., Gogarten, J. F., Calvignac-Spencer, S., and Leendertz, F. H. (2015)

Investigating the zoonotic origin of the West African Ebola epidemic. *EMBO molecular medicine* **7**, 17-23

8. Gallo, R. C., Salahuddin, S. Z., Popovic, M., Shearer, G. M., Kaplan, M., Haynes, B. F., Palker, T. J., Redfield, R., Oleske, J., Safai, B., and et al. (1984) Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224**, 500-503
9. Barré-Sinoussi, F., Chermann, J.-C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., and Rouzioux, C. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**, 868-871
10. Levy, J. A., Hoffman, A. D., Kramer, S. M., Landis, J. A., Shimabukuro, J. M., and Oshiro, L. S. (1984) Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* **225**, 840-842
11. Poiesz, B. J., Ruscetti, F. W., Gazdar, A. F., Bunn, P. A., Minna, J. D., and Gallo, R. C. (1980) Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci U S A* **77**, 7415-7419
12. Becerra, J. C., Bildstein, L. S., and Gach, J. S. (2016) Recent Insights into the HIV/AIDS Pandemic. *Microb Cell* **3**, 451-475
13. Arts, E. J., and Hazuda, D. J. (2012) HIV-1 antiretroviral drug therapy. *Cold Spring Harb Perspect Med* **2**, a007161
14. Dabis, F., and Bekker, L.-G. (2017) We still need to beat HIV. American Association for the Advancement of Science

15. Maina, E. K., Adan, A. A., Mureithi, H., Muriuki, J., and Lwembe, R. M. (2021) A Review of Current Strategies Towards the Elimination of Latent HIV-1 and Subsequent HIV-1 Cure. *Current HIV research* **19**, 14-26

16. UNAIDS. (2020) Global HIV/AIDS 2020 factsheet. https://www.unaids.org/en/resources/fact-sheet

17. Dror, A. A., Eisenbach, N., Taiber, S., Morozov, N. G., Mizrachi, M., Zigron, A., Srouji, S., and Sela, E. (2020) Vaccine hesitancy: the next challenge in the fight against COVID-19. *Eur J Epidemiol* **35**, 775-779

18. Sallam, M. (2021) COVID-19 Vaccine Hesitancy Worldwide: A Concise Systematic Review of Vaccine Acceptance Rates. *Vaccines (Basel)* **9**, 160

19. Szemiel, A. M., Merits, A., Orton, R. J., MacLean, O., Pinto, R. M., Wickenhagen, A., Lieber, G., Turnbull, M. L., Wang, S., Mair, D., da Silva Filipe, A., Willett, B. J., Wilson, S. J., Patel, A. H., Thomson, E. C., Palmarini, M., Kohl, A., and Stewart, M. E. (2021) In vitro evolution of Remdesivir resistance reveals genome plasticity of SARS-CoV-2. *bioRxiv*, 2021.2002.2001.429199

20. Yang, W. (2011) Nucleases: diversity of structure, function and mechanism. *Q Rev Biophys* **44**, 1-93

21. Doudna, J. A., and Cech, T. R. (2002) The chemical repertoire of natural ribozymes. *Nature* **418**, 222-228

22. Wyatt, H. D., and West, S. C. (2014) Holliday junction resolvases. *Cold Spring Harb Perspect Biol* **6**, a023192

23. Ferron, F., Subissi, L., Silveira De Morais, A. T., Le, N. T. T., Sevajol, M., Gluais, L., Decroly, E., Vonrhein, C., Bricogne, G., Canard, B., and Imbert, I. (2018) Structural and

molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *Proc Natl Acad Sci U S A* **115**, E162-E171

24. Ogando, N. S., Ferron, F., Decroly, E., Canard, B., Posthuma, C. C., and Snijder, E. J. (2019) The Curious Case of the Nidovirus Exoribonuclease: Its Role in RNA Synthesis and Replication Fidelity. *Front Microbiol* **10**, 1813

25. Nadal, M., Mas, P. J., Blanco, A. G., Arnan, C., Sola, M., Hart, D. J., and Coll, M. (2010) Structure and inhibition of herpesvirus DNA packaging terminase nuclease domain. *Proc Natl Acad Sci U S A* **107**, 16078-16083

26. Beilhartz, G. L., and Gotte, M. (2010) HIV-1 Ribonuclease H: Structure, Catalytic Mechanism and Inhibitors. *Viruses* **2**, 900-926

27. Reich, S., Guilligay, D., Pflug, A., Malet, H., Berger, I., Crepin, T., Hart, D., Lunardi, T., Nanao, M., Ruigrok, R. W., and Cusack, S. (2014) Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature* **516**, 361-366

28. Ma, Y., Wu, L., Shaw, N., Gao, Y., Wang, J., Sun, Y., Lou, Z., Yan, L., Zhang, R., and Rao, Z. (2015) Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. *Proc Natl Acad Sci U S A* **112**, 9436-9441

29. Ferro, S., Gitto, R., Buemi, M. R., Karamanou, S., Stevaert, A., Naesens, L., and De Luca, L. (2018) Identification of influenza PA-Nter endonuclease inhibitors using pharmacophore- and docking-based virtual screening. *Bioorganic & medicinal chemistry* **26**, 4544-4550

30. Goldner, T., Hewlett, G., Ettischer, N., Ruebsamen-Schaeff, H., Zimmermann, H., and Lischka, P. (2011) The novel anticytomegalovirus compound AIC246 (Letermovir)

inhibits human cytomegalovirus replication through a specific antiviral mechanism that involves the viral terminase. *J Virol* **85**, 10884-10893

31. Noshi, T., Kitano, M., Taniguchi, K., Yamamoto, A., Omoto, S., Baba, K., Hashimoto, T., Ishida, K., Kushima, Y., and Hattori, K. (2018) In vitro characterization of baloxavir acid, a first-in-class cap-dependent endonuclease inhibitor of the influenza virus polymerase PA subunit. *Antiviral research* **160**, 109-117
32. Fauci, A. S., and Lane, H. C. (2020) Four Decades of HIV/AIDS - Much Accomplished, Much to Do. *N Engl J Med* **383**, 1-4
33. Walker, B. D., and Hirsch, M. S. (2013) Antiretroviral therapy in early HIV infection. *N Engl J Med* **368**, 279-281
34. Nakagawa, F., Lodwick, R. K., Smith, C. J., Smith, R., Cambiano, V., Lundgren, J. D., Delpech, V., and Phillips, A. N. (2012) Projected life expectancy of people with HIV according to timing of diagnosis. *AIDS (London, England)* **26**, 335-343
35. Pau, A. K., and George, J. M. (2014) Antiretroviral therapy: current drugs. *Infect Dis Clin North Am* **28**, 371-402
36. Andreola, M. L. (2004) Closely related antiretroviral agents as inhibitors of two HIV-1 enzymes, ribonuclease H and integrase: "killing two birds with one stone". *Current pharmaceutical design* **10**, 3713-3723
37. Klumpp, K., Hang, J. Q., Rajendran, S., Yang, Y., Derosier, A., Wong Kai In, P., Overton, H., Parkes, K. E., Cammack, N., and Martin, J. A. (2003) Two-metal ion mechanism of RNA cleavage by HIV RNase H and mechanism-based design of selective HIV RNase H inhibitors. *Nucleic Acids Res* **31**, 6852-6859

38. Diamond, T. L., and Bushman, F. D. (2006) Role of metal ions in catalysis by HIV integrase analyzed using a quantitative PCR disintegration assay. *Nucleic Acids Res* **34**, 6116-6125

39. Hu, W. S., and Hughes, S. H. (2012) HIV-1 reverse transcription. *Cold Spring Harb Perspect Med* **2**, a006882

40. Scarsi, K. K., Havens, J. P., Podany, A. T., Avedissian, S. N., and Fletcher, C. V. (2020) HIV-1 Integrase Inhibitors: A Comparative Review of Efficacy and Safety. *Drugs* **80**, 1649-1676

41. Boyer, P. L., Smith, S. J., Zhao, X. Z., Das, K., Gruber, K., Arnold, E., Burke, T. R., Jr., and Hughes, S. H. (2018) Developing and Evaluating Inhibitors against the RNase H Active Site of HIV-1 Reverse Transcriptase. *J Virol* **92**, e02203-02217

42. Beilhartz, G. L., Wendeler, M., Baichoo, N., Rausch, J., Le Grice, S., and Gotte, M. (2009) HIV-1 reverse transcriptase can simultaneously engage its DNA/RNA substrate at both DNA polymerase and RNase H active sites: implications for RNase H inhibition. *Journal of molecular biology* **388**, 462-474

43. Engelman, A. N., and Cherepanov, P. (2021) Close-up: HIV/SIV intasome structures shed new light on integrase inhibitor binding and viral escape mechanisms. *FEBS J* **288**, 427-433

44. Beilhartz, G. L., Ngure, M., Johns, B. A., DeAnda, F., Gerondelis, P., and Gotte, M. (2014) Inhibition of the ribonuclease H activity of HIV-1 reverse transcriptase by GSK5750 correlates with slow enzyme-inhibitor dissociation. *J Biol Chem* **289**, 16270-16277

45. Alvarez, D. M., Castillo, E., Duarte, L. F., Arriagada, J., Corrales, N., Farias, M. A., Henriquez, A., Agurto-Munoz, C., and Gonzalez, P. A. (2020) Current Antivirals and Novel Botanical Molecules Interfering With Herpes Simplex Virus Infection. *Front Microbiol* **11**, 139

46. Umene, K., Oohashi, S., Yoshida, M., and Fukumaki, Y. (2008) Diversity of the a sequence of herpes simplex virus type 1 developed during evolution. *Journal of General Virology* **89**, 841-852

47. Deiss, L. P., Chou, J., and Frenkel, N. (1986) Functional domains within the a sequence involved in the cleavage-packaging of herpes simplex virus DNA. *Journal of virology* **59**, 605-618

48. Tamashiro, J., Filpula, D., Friedmann, T., and Spector, D. (1984) Structure of the heterogeneous LS junction region of human cytomegalovirus strain AD169 DNA. *Journal of virology* **52**, 541-548

49. Chowdhury, S., Buhk, H., Ludwig, H., and Hammerschmidt, W. (1990) Genomic termini of equine herpesvirus 1. *Journal of virology* **64**, 873-880

50. Davison, A. (1984) Structure of the genome termini of varicella-zoster virus. *Journal of general virology* **65**, 1969-1977

51. Deiss, L. P., and Frenkel, N. (1986) Herpes simplex virus amplicon: cleavage of concatemeric DNA is linked to packaging and involves amplification of the terminally reiterated a sequence. *Journal of virology* **57**, 933-941

52. Hammerschmidt, W., Ludwig, H., and Buhk, H. (1988) Specificity of cleavage in replicative-form DNA of bovine herpesvirus 1. *Journal of virology* **62**, 1355-1363

53. Hodge, P. D., and Stow, N. D. (2001) Effects of mutations within the herpes simplex virus type 1 DNA encapsidation signal on packaging efficiency. *Journal of virology* **75**, 8977-8986

54. McVoy, M. A., Nixon, D. E., and Adler, S. P. (1997) Circularization and cleavage of guinea pig cytomegalovirus genomes. *Journal of virology* **71**, 4209-4217

55. Nasseri, M., and Mocarski, E. (1988) The cleavage recognition signal is contained within sequences surrounding an aa junction in herpes simplex virus DNA. *Virology* **167**, 25-30

56. Smiley, J. R., Duncan, J., and Howes, M. (1990) Sequence requirements for DNA rearrangements induced by the terminal repeat of herpes simplex virus type 1 KOS DNA. *Journal of virology* **64**, 5036-5050

57. Tong, L., and Stow, N. D. (2010) Analysis of herpes simplex virus type 1 DNA packaging signal mutations in the context of the viral genome. *Journal of virology* **84**, 321-329

58. Varmuza, S. L., and Smiley, J. R. (1985) Signals for site-specific cleavage of HSV DNA: maturation involves two separate cleavage events at sites distal to the recognition sequences. *Cell* **41**, 793-802

59. Zimmermann, J., and Hammerschmidt, W. (1995) Structure and role of the terminal repeats of Epstein-Barr virus in processing and packaging of virion DNA. *Journal of Virology* **69**, 3147-3155

60. Wang, J. B., and McVoy, M. A. (2011) A 128-bp sequence containing pac1 and a presumed cryptic pac2 includes the cis elements sufficient to mediate efficient genome maturation of human cytomegalovirus. *Journal of Virology*

61. Ligat, G., Cazal, R., Hantz, S., and Alain, S. (2018) The human cytomegalovirus terminase complex as an antiviral target: a close-up view. *FEMS Microbiol Rev* **42**, 137-145

62. McVoy, M. A., Nixon, D. E., Adler, S. P., and Mocarski, E. S. (1998) Sequences within the herpesvirus-conserved pac1 and pac2 motifs are required for cleavage and packaging of the murine cytomegalovirus genome. *J Virol* **72**, 48-56

63. Muylaert, I., Tang, K. W., and Elias, P. (2011) Replication and recombination of herpes simplex virus DNA. *J Biol Chem* **286**, 15619-15624

64. McVoy, M. A., and Adler, S. P. (1994) Human cytomegalovirus DNA replicates after early circularization by concatemer formation, and inversion occurs within the concatemer. *J Virol* **68**, 1040-1051

65. Full, F., and Ensser, A. (2019) Early Nuclear Events after Herpesviral Infection. *J Clin Med* **8**, 1408

66. Borst, E. M., Kleine-Albers, J., Gabaev, I., Babic, M., Wagner, K., Binz, A., Degenhardt, I., Kalesse, M., Jonjic, S., Bauerfeind, R., and Messerle, M. (2013) The human cytomegalovirus UL51 protein is essential for viral genome cleavage-packaging and interacts with the terminase subunits pUL56 and pUL89. *J Virol* **87**, 1720-1732

67. Neuber, S., Wagner, K., Goldner, T., Lischka, P., Steinbrueck, L., Messerle, M., and Borst, E. M. (2017) Mutual Interplay between the Human Cytomegalovirus Terminase Subunits pUL51, pUL56, and pUL89 Promotes Terminase Complex Formation. *J Virol* **91**

68. Chelikani, V., Ranjan, T., and Kondabagil, K. (2014) Revisiting the genome packaging in viruses with lessons from the "Giants". *Virology* **466-467**, 15-26

69. Selvarajan Sigamani, S., Zhao, H., Kamau, Y. N., Baines, J. D., and Tang, L. (2013) The structure of the herpes simplex virus DNA-packaging terminase pUL15 nuclease domain suggests an evolutionary lineage among eukaryotic and prokaryotic viruses. *J Virol* **87**, 7140-7148

70. Catalano, C. E. (2005) Viral genome packaging machines. *Viral Genome Packaging Machines: Genetics, Structure, and Mechanism*, 1-4

71. Tsuchida, S. I., Kokubo, H., Tasaka, M., and Fujisawa, H. (1996) DNA sequences responsible for specificity of DNA packaging and phage growth interference of bacteriophages T3 and T7. *Virology* **217**, 332-337

72. Alam, T. I., Draper, B., Kondabagil, K., Rentas, F. J., Ghosh-Kumar, M., Sun, S., Rossmann, M. G., and Rao, V. B. (2008) The headful packaging nuclease of bacteriophage T4. *Mol Microbiol* **69**, 1180-1190

73. Heming, J. D., Conway, J. F., and Homa, F. L. (2017) Herpesvirus Capsid Assembly and DNA Packaging. in *Cell Biology of Herpes Viruses* (Osterrieder, K. ed.), Springer International Publishing, Cham. pp 119-142

74. Wang, J. B., and McVoy, M. A. (2011) A 128-base-pair sequence containing the pac1 and a presumed cryptic pac2 sequence includes cis elements sufficient to mediate efficient genome maturation of human cytomegalovirus. *J Virol* **85**, 4432-4439

75. Sun, L., Zhang, X., Gao, S., Rao, P. A., Padilla-Sanchez, V., Chen, Z., Sun, S., Xiang, Y., Subramaniam, S., and Rao, V. B. (2015) Cryo-EM structure of the bacteriophage T4 portal protein assembly at near-atomic resolution. *Nature communications* **6**, 1-11

76. Yang, Y., Yang, P., Wang, N., Chen, Z., Su, D., Zhou, Z. H., Rao, Z., and Wang, X. (2020) Architecture of the herpesvirus genome-packaging complex and implications for DNA translocation. *Protein & cell* **11**, 339-351

77. Heming, J. D., Conway, J. F., and Homa, F. L. (2017) Herpesvirus Capsid Assembly and DNA Packaging. *Adv Anat Embryol Cell Biol* **223**, 119-142

78. Koslowski, K. M., Shaver, P. R., Casey, J. T., 2nd, Wilson, T., Yamanaka, G., Sheaffer, A. K., Tenney, D. J., and Pederson, N. E. (1999) Physical and functional interactions between the herpes simplex virus UL15 and UL28 DNA cleavage and packaging proteins. *J Virol* **73**, 1704-1707

79. Abbotts, A. P., Preston, V. G., Hughes, M., Patel, A. H., and Stow, N. D. (2000) Interaction of the herpes simplex virus type 1 packaging protein UL15 with full-length and deleted forms of the UL28 protein. *J Gen Virol* **81**, 2999-3009

80. Wang, J. B., Zhu, Y., McVoy, M. A., and Parris, D. S. (2012) Changes in subcellular localization reveal interactions between human cytomegalovirus terminase subunits. *Virol J* **9**, 315

81. Bogner, E., Radsak, K., and Stinski, M. F. (1998) The gene product of human cytomegalovirus open reading frame UL56 binds the pac motif and has specific nuclease activity. *J Virol* **72**, 2259-2264

82. Hwang, J. S., and Bogner, E. (2002) ATPase activity of the terminase subunit pUL56 of human cytomegalovirus. *J Biol Chem* **277**, 6943-6948

83. Scholz, B., Rechter, S., Drach, J. C., Townsend, L. B., and Bogner, E. (2003) Identification of the ATP-binding site in the terminase subunit pUL56 of human cytomegalovirus. *Nucleic Acids Res* **31**, 1426-1433

84. Theiß, J., Sung, M. W., Holzenburg, A., and Bogner, E. (2019) Full-length human cytomegalovirus terminase pUL89 adopts a two-domain structure specific for DNA packaging. *PLoS pathogens* **15**, e1008175

85. Scheffczik, H., Savva, C. G., Holzenburg, A., Kolesnikova, L., and Bogner, E. (2002) The terminase subunits pUL56 and pUL89 of human cytomegalovirus are DNA-metabolizing proteins with toroidal structure. *Nucleic Acids Res* **30**, 1695-1703

86. Champier, G., Hantz, S., Couvreux, A., Stuppfler, S., Mazeron, M. C., Bouaziz, S., Denis, F., and Alain, S. (2007) New functional domains of human cytomegalovirus pUL89 predicted by sequence analysis and three-dimensional modelling of the catalytic site DEXDc. *Antiviral therapy* **12**, 217-232

87. Fodor, E., and te Velthuis, A. J. (2019) Structure and function of the influenza virus transcription and replication machinery. *Cold Spring Harbor Perspectives in Medicine*, a038398

88. te Velthuis, A. J., and Fodor, E. (2016) Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis. *Nature Reviews Microbiology* **14**, 479

89. Yuan, P., Bartlam, M., Lou, Z., Chen, S., Zhou, J., He, X., Lv, Z., Ge, R., Li, X., Deng, T., Fodor, E., Rao, Z., and Liu, Y. (2009) Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. *Nature* **458**, 909-913

90. Fodor, E., and Te Velthuis, A. J. W. (2020) Structure and Function of the Influenza Virus Transcription and Replication Machinery. *Cold Spring Harb Perspect Med* **10**

91. Noshi, T., Kitano, M., Taniguchi, K., Yamamoto, A., Omoto, S., Baba, K., Hashimoto, T., Ishida, K., Kushima, Y., Hattori, K., Kawai, M., Yoshida, R., Kobayashi, M., Yoshinaga, T., Sato, A., Okamatsu, M., Sakoda, Y., Kida, H., Shishido, T., and Naito, A.

(2018) In vitro characterization of baloxavir acid, a first-in-class cap-dependent endonuclease inhibitor of the influenza virus polymerase PA subunit. *Antiviral Res* **160**, 109-117

92. Smith, E. C. (2017) The not-so-infinite malleability of RNA viruses: Viral and cellular determinants of RNA virus mutation rates. *PLoS Pathog* **13**, e1006254

93. Eckerle, L. D., Lu, X., Sperry, S. M., Choi, L., and Denison, M. R. (2007) High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J Virol* **81**, 12135-12144

94. Becares, M., Pascual-Iglesias, A., Nogales, A., Sola, I., Enjuanes, L., and Zuniga, S. (2016) Mutagenesis of Coronavirus nsp14 Reveals Its Potential Role in Modulation of the Innate Immune Response. *J Virol* **90**, 5399-5414

95. Eskier, D., Suner, A., Oktay, Y., and Karakulah, G. (2020) Mutations of SARS-CoV-2 nsp14 exhibit strong association with increased genome-wide mutation load. *PeerJ* **8**, e10181

96. Ulferts, R., and Ziebuhr, J. (2011) Nidovirus ribonucleases: Structures and functions in viral replication. *RNA Biol* **8**, 295-304

97. Johansson, E., and Dixon, N. (2013) Replicative DNA polymerases. *Cold Spring Harb Perspect Biol* **5**, a012799

98. Kunkel, T. A. (2004) DNA replication fidelity. *J Biol Chem* **279**, 16895-16898

99. Singh, A., Pandey, M., Nandakumar, D., Raney, K. D., Yin, Y. W., and Patel, S. S. (2020) Excessive excision of correct nucleotides during DNA synthesis explained by replication hurdles. *EMBO J* **39**, e103367

100. Krafcikova, P., Silhan, J., Nencka, R., and Boura, E. (2020) Structural analysis of the SARS-CoV-2 methyltransferase complex involved in RNA cap creation bound to sinefungin. *Nat Commun* **11**, 3717

101. Bouvet, M., Imbert, I., Subissi, L., Gluais, L., Canard, B., and Decroly, E. (2012) 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

102. DuBois, R. M., Slavish, P. J., Baughman, B. M., Yun, M. K., Bao, J., Webby, R. J., Webb, T. R., and White, S. W. (2012) Structural and biochemical basis for development of influenza virus inhibitors targeting the PA endonuclease. *PLoS Pathog* **8**, e1002830

103. Tischer, B. K., and Osterrieder, N. (2010) Herpesviruses--a zoonotic threat? *Vet Microbiol* **140**, 266-270

104. Gilden, D. H., Mahalingam, R., Cohrs, R. J., and Tyler, K. L. (2007) Herpesvirus infections of the nervous system. *Nature clinical practice. Neurology* **3**, 82-94

105. Ramanan, P., and Razonable, R. R. (2013) Cytomegalovirus infections in solid organ transplantation: a review. *Infect Chemother* **45**, 260-271

106. Dollard, S. C., Grosse, S. D., and Ross, D. S. (2007) New estimates of the prevalence of neurological and sensory sequelae and mortality associated with congenital cytomegalovirus infection. *Rev Med Virol* **17**, 355-363

107. Bartlett, A. W., McMullan, B., Rawlinson, W. D., and Palasanthiran, P. (2017) Hearing and neurodevelopmental outcomes for children with asymptomatic congenital cytomegalovirus infection: A systematic review. *Rev Med Virol*

108. Cannon, M. J., Schmid, D. S., and Hyde, T. B. (2010) Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. *Rev Med Virol* **20**, 202-213

109. Bate, S. L., Dollard, S. C., and Cannon, M. J. (2010) Cytomegalovirus seroprevalence in the United States: the national health and nutrition examination surveys, 1988–2004. *Clinical infectious diseases* **50**, 1439-1447

110. Douglas, C. M., Barnard, R., Holder, D., Leavitt, R., Levitan, D., Maguire, M., Nickle, D., Teal, V., Wan, H., van Alewijk, D., van Doorn, L. J., Chou, S., and Strizki, J. (2020) Letermovir Resistance Analysis in a Clinical Trial of Cytomegalovirus Prophylaxis for Hematopoietic Stem Cell Transplant Recipients. *J Infect Dis* **221**, 1117-1126

111. Collaborators, G. B. D. I. (2019) Mortality, morbidity, and hospitalisations due to influenza lower respiratory tract infections, 2017: an analysis for the Global Burden of Disease Study 2017. *The Lancet. Respiratory medicine* **7**, 69-89

112. Molinari, N. A., Ortega-Sanchez, I. R., Messonnier, M. L., Thompson, W. W., Wortley, P. M., Weintraub, E., and Bridges, C. B. (2007) The annual impact of seasonal influenza in the US: measuring disease burden and costs. *Vaccine* **25**, 5086-5096

113. WHO. (2020) Influenza (Seasonal). https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal)

114. Plotkin, S. A. (2005) Vaccines: past, present and future. *Nature medicine* **11**, S5-11

115. Gubareva, L. V., Kaiser, L., and Hayden, F. G. (2000) Influenza virus neuraminidase inhibitors. *The Lancet* **355**, 827-835

116. CDC. (2020) Influenza Antiviral Medications: Summary for Clinicians.

117. Pflug, A., Guilligay, D., Reich, S., and Cusack, S. (2014) Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* **516**, 355-360

118. Pflug, A., Guilligay, D., Reich, S., and Cusack, S. (2014) Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* **516**, 355

119. Dou, D., Revol, R., Östbye, H., Wang, H., and Daniels, R. (2018) Influenza A virus cell entry, replication, virion assembly and movement. *Frontiers in immunology* **9**, 1581

120. Dias, A., Bouvier, D., Crépin, T., McCarthy, A. A., Hart, D. J., Baudin, F., Cusack, S., and Ruigrok, R. W. H. (2009) The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **458**, 914-918

121. Guilligay, D., Tarendeau, F., Resa-Infante, P., Coloma, R., Crepin, T., Sehr, P., Lewis, J., Ruigrok, R. W., Ortin, J., and Hart, D. J. (2008) The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nature structural & molecular biology* **15**, 500

122. Hayden, F. G., Sugaya, N., Hirotsu, N., Lee, N., de Jong, M. D., Hurt, A. C., Ishida, T., Sekino, H., Yamada, K., and Portsmouth, S. (2018) Baloxavir marboxil for uncomplicated influenza in adults and adolescents. *New England Journal of Medicine* **379**, 913-923

123. Omoto, S., Speranzini, V., Hashimoto, T., Noshi, T., Yamaguchi, H., Kawai, M., Kawaguchi, K., Uehara, T., Shishido, T., and Naito, A. (2018) Characterization of influenza virus variants induced by treatment with the endonuclease inhibitor baloxavir marboxil. *Scientific reports* **8**, 1-15

124. Mishin, V. P., Patel, M. C., Chesnokov, A., De La Cruz, J., Nguyen, H. T., Lollis, L., Hodges, E., Jang, Y., Barnes, J., and Uyeki, T. (2019) Susceptibility of Influenza A, B, C, and D Viruses to Baloxavir. *Emerging infectious diseases* **25**, 1969

125. Takashita, E., Ichikawa, M., Morita, H., Ogawa, R., Fujisaki, S., Shirakura, M., Miura, H., Nakamura, K., Kishida, N., and Kuwahara, T. (2019) Human-to-human transmission of influenza A (H3N2) virus with reduced susceptibility to baloxavir, Japan, February 2019. *Emerging infectious diseases* **25**, 2108

126. Delang, L., Abdelnabi, R., and Neyts, J. (2018) Favipiravir as a potential countermeasure against neglected and emerging RNA viruses. *Antiviral research* **153**, 85-94

127. Baranovich, T., Wong, S.-S., Armstrong, J., Marjuki, H., Webby, R. J., Webster, R. G., and Govorkova, E. A. (2013) T-705 (favipiravir) induces lethal mutagenesis in influenza A H1N1 viruses in vitro. *Journal of virology* **87**, 3741-3751

128. Goldhill, D. H., te Velthuis, A. J. W., Fletcher, R. A., Langat, P., Zambon, M., Lackenby, A., and Barclay, W. S. (2018) The mechanism of resistance to favipiravir in influenza. *Proceedings of the National Academy of Sciences* **115**, 11613-11618

129. Hayden, F. G., and Shindo, N. (2019) Influenza virus polymerase inhibitors in clinical development. *Curr Opin Infect Dis* **32**, 176-186

130. Trevejo, J. M., Asmal, M., Vingerhoets, J., Polo, R., Robertson, S., Jiang, Y., Kieffer, T. L., and Leopold, L. (2018) Pimodivir treatment in adult volunteers experimentally inoculated with live influenza virus: a Phase IIa, randomized, double-blind, placebo-controlled study. *Antiviral therapy* **23**, 335-344

131. Clark, M. P., Ledeboer, M. W., Davies, I., Byrn, R. A., Jones, S. M., Perola, E., Tsai, A., Jacobs, M., Nti-Addae, K., Bandarage, U. K., Boyd, M. J., Bethiel, R. S., Court, J. J., Deng, H., Duffy, J. P., Dorsch, W. A., Farmer, L. J., Gao, H., Gu, W., Jackson, K., Jacobs, D. H., Kennedy, J. M., Ledford, B., Liang, J., Maltais, F., Murcko, M., Wang, T., Wannamaker, M. W., Bennett, H. B., Leeman, J. R., McNeil, C., Taylor, W. P.,

Memmott, C., Jiang, M., Rijnbrand, R., Bral, C., Germann, U., Nezami, A., Zhang, Y., Salituro, F. G., Bennani, Y. L., and Charifson, P. S. (2014) Discovery of a Novel, First-in-Class, Orally Bioavailable Azaindole Inhibitor (VX-787) of Influenza PB2. *Journal of Medicinal Chemistry* **57**, 6668-6678

132. Byrn, R. A., Jones, S. M., Bennett, H. B., Bral, C., Clark, M. P., Jacobs, M. D., Kwong, A. D., Ledeboer, M. W., Leeman, J. R., McNeil, C. F., Murcko, M. A., Nezami, A., Perola, E., Rijnbrand, R., Saxena, K., Tsai, A. W., Zhou, Y., and Charifson, P. S. (2015) Preclinical activity of VX-787, a first-in-class, orally bioavailable inhibitor of the influenza virus polymerase PB2 subunit. *Antimicrobial agents and chemotherapy* **59**, 1569-1582

133. Finberg, R. W., Lanno, R., Anderson, D., Fleischhackl, R., van Duijnhoven, W., Kauffman, R. S., Kosoglou, T., Vingerhoets, J., and Leopold, L. (2019) Phase 2b Study of Pimodivir (JNJ-63623872) as Monotherapy or in Combination With Oseltamivir for Treatment of Acute Uncomplicated Seasonal Influenza A: TOPAZ Trial. *The Journal of infectious diseases* **219**, 1026-1034

134. Eastman, R. T., Roth, J. S., Brimacombe, K. R., Simeonov, A., Shen, M., Patnaik, S., and Hall, M. D. (2020) Remdesivir: A Review of Its Discovery and Development Leading to Emergency Use Authorization for Treatment of COVID-19. *Acs Central Sci* **6**, 672-683

135. Helmy, Y. A., Fawzy, M., Elaswad, A., Sobieh, A., Kenney, S. P., and Shehata, A. A. (2020) The COVID-19 Pandemic: A Comprehensive Review of Taxonomy, Genetics, Epidemiology, Diagnosis, Treatment, and Control. *J Clin Med* **9**, 1225

136. Singhal, T. (2020) A Review of Coronavirus Disease-2019 (COVID-19). *Indian J Pediatr* **87**, 281-286

137. Hu, B., Guo, H., Zhou, P., and Shi, Z. L. (2021) Characteristics of SARS-CoV-2 and COVID-19. *Nature reviews. Microbiology* **19**, 141-154

138. Wiersinga, W. J., Rhodes, A., Cheng, A. C., Peacock, S. J., and Prescott, H. C. (2020) Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review. *Jama* **324**, 782-793

139. Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., Si, H. R., Zhu, Y., Li, B., Huang, C. L., Chen, H. D., Chen, J., Luo, Y., Guo, H., Jiang, R. D., Liu, M. Q., Chen, Y., Shen, X. R., Wang, X., Zheng, X. S., Zhao, K., Chen, Q. J., Deng, F., Liu, L. L., Yan, B., Zhan, F. X., Wang, Y. Y., Xiao, G. F., and Shi, Z. L. (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273

140. Cucinotta, D., and Vanelli, M. (2020) WHO Declares COVID-19 a Pandemic. *Acta bio-medica : Atenei Parmensis* **91**, 157-160

141. Gordon, C. J., Tchesnokov, E. P., Woolner, E., Perry, J. K., Feng, J. Y., Porter, D. P., and Gotte, M. (2020) Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. *J Biol Chem* **295**, 6785-6797

142. Tchesnokov, E. P., Raeisimakiani, P., Ngure, M., Marchant, D., and Gotte, M. (2018) Recombinant RNA-Dependent RNA Polymerase Complex of Ebola Virus. *Sci Rep* **8**, 3970

143. Warren, T., Jordan, R., Lo, M., Soloveva, V., Ray, A., Bannister, R., Mackman, R., Perron, M., Stray, K., and Feng, J. (2015) Nucleotide prodrug GS-5734 is a broad-spectrum filovirus inhibitor that provides complete therapeutic protection against the

development of Ebola virus disease (EVD) in infected non-human primates. in *Open Forum Infectious Diseases*, Infectious Diseases Society of America

144. Agostini, M. L., Andres, E. L., Sims, A. C., Graham, R. L., Sheahan, T. P., Lu, X., Smith, E. C., Case, J. B., Feng, J. Y., Jordan, R., Ray, A. S., Cihlar, T., Siegel, D., Mackman, R. L., Clarke, M. O., Baric, R. S., and Denison, M. R. (2018) Coronavirus Susceptibility to the Antiviral Remdesivir (GS-5734) Is Mediated by the Viral Polymerase and the Proofreading Exoribonuclease. *mBio* **9**, e00221-00218

145. Tchesnokov, E. P., Gordon, C. J., Woolner, E., Kocinkova, D., Perry, J. K., Feng, J. Y., Porter, D. P., and Gotte, M. (2020) Template-dependent inhibition of coronavirus RNA-dependent RNA polymerase by remdesivir reveals a second mechanism of action. *J Biol Chem* **295**, 16156-16165

146. Garzoni, F., Bieniossek, C., and Berger, I. (2012) The MultiBac BEVS for producing proteins and their complexes (Prot54).

147. Gugliesi, F., Coscia, A., Griffante, G., Galitska, G., Pasquero, S., Albano, C., and Biolatti, M. (2020) Where do we Stand after Decades of Studying Human Cytomegalovirus? *Microorganisms* **8**

148. Rawlinson, W. D., Boppana, S. B., Fowler, K. B., Kimberlin, D. W., Lazzarotto, T., Alain, S., Daly, K., Doutre, S., Gibson, L., Giles, M. L., Greenlee, J., Hamilton, S. T., Harrison, G. J., Hui, L., Jones, C. A., Palasanthiran, P., Schleiss, M. R., Shand, A. W., and van Zuylen, W. J. (2017) Congenital cytomegalovirus infection in pregnancy and the neonate: consensus recommendations for prevention, diagnosis, and therapy. *Lancet Infect Dis* **17**, e177-e188

149. Fakhreddine, A. Y., Frenette, C. T., and Konijeti, G. G. (2019) A Practical Review of Cytomegalovirus in Gastroenterology and Hepatology. *Gastroenterology research and practice* **2019**, 6156581

150. Komatsu, T. E., Pikis, A., Naeger, L. K., and Harrington, P. R. (2014) Resistance of human cytomegalovirus to ganciclovir/valganciclovir: a comprehensive review of putative resistance pathways. *Antiviral Res* **101**, 12-25

151. Poole, C. L., and James, S. H. (2018) Antiviral Therapies for Herpesviruses: Current Agents and New Directions. *Clin Ther* **40**, 1282-1298

152. Gordon, C. J., Tchesnokov, E. P., Feng, J. Y., Porter, D. P., and Gotte, M. (2020) The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J Biol Chem* **295**, 4773-4779

153. Todd, B., Tchesnokov, E. P., and Gotte, M. (2021) The active form of the influenza cap-snatching endonuclease inhibitor baloxavir marboxil is a tight binding inhibitor. *J Biol Chem*, 100486

154. Gerlach, P., Malet, H., Cusack, S., and Reguera, J. (2015) Structural Insights into Bunyavirus Replication and Its Regulation by the vRNA Promoter. *Cell* **161**, 1267-1279

155. Murphy, C. I., Piwnica-Worms, H., Grunwald, S., Romanow, W. G., Francis, N., and Fan, H. Y. (2004) Overview of the baculovirus expression system. *Current protocols in molecular biology* **Chapter 16**, Unit 16 19

156. Hilbert, B. J., Hayes, J. A., Stone, N. P., Xu, R. G., and Kelch, B. A. (2017) The large terminase DNA packaging motor grips DNA with its ATPase domain for cleavage by the flexible nuclease domain. *Nucleic Acids Res* **45**, 3591-3605

157. Hilbert, B. J., Hayes, J. A., Stone, N. P., Duffy, C. M., Sankaran, B., and Kelch, B. A. (2015) Structure and mechanism of the ATPase that powers viral genome packaging. *Proceedings of the National Academy of Sciences* **112**, E3792-E3799

158. Rao, V. B., and Feiss, M. (2008) The bacteriophage DNA packaging motor. *Annual review of genetics* **42**, 647-681

159. Paredes, A. M., and Yu, D. (2012) Human cytomegalovirus: bacterial artificial chromosome (BAC) cloning and genetic manipulation. *Curr Protoc Microbiol* **Chapter 14**, Unit14E 14

160. Iuliano, A. D., Roguski, K. M., Chang, H. H., Muscatello, D. J., Palekar, R., Tempia, S., Cohen, C., Gran, J. M., Schanzer, D., Cowling, B. J., Wu, P., Kyncl, J., Ang, L. W., Park, M., Redlberger-Fritz, M., Yu, H., Espenhain, L., Krishnan, A., Emukule, G., van Asten, L., Pereira da Silva, S., Aungkulanon, S., Buchholz, U., Widdowson, M. A., Bresee, J. S., and Global Seasonal Influenza-associated Mortality Collaborator, N. (2018) Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet* **391**, 1285-1300

161. Belongia, E. A., and Osterholm, M. T. (2020) COVID-19 and flu, a perfect storm. *Science* **368**, 1163

162. Paules, C., and Subbarao, K. (2017) Influenza. *The Lancet* **390**, 697-708

163. Hayden, F. G., Sugaya, N., Hirotsu, N., Lee, N., de Jong, M. D., Hurt, A. C., Ishida, T., Sekino, H., Yamada, K., Portsmouth, S., Kawaguchi, K., Shishido, T., Arai, M., Tsuchiya, K., Uehara, T., Watanabe, A., and Baloxavir Marboxil Investigators, G. (2018) Baloxavir Marboxil for Uncomplicated Influenza in Adults and Adolescents. *N Engl J Med* **379**, 913-923

164. Ison, M. G., Portsmouth, S., Yoshida, Y., Shishido, T., Mitchener, M., Tsuchiya, K., Uehara, T., and Hayden, F. G. (2020) Early treatment with baloxavir marboxil in high-risk adolescent and adult outpatients with uncomplicated influenza (CAPSTONE-2): a randomised, placebo-controlled, phase 3 trial. *Lancet Infect Dis* **20**, 1204-1214

165. Uehara, T., Hayden, F. G., Kawaguchi, K., Omoto, S., Hurt, A. C., De Jong, M. D., Hirotsu, N., Sugaya, N., Lee, N., Baba, K., Shishido, T., Tsuchiya, K., Portsmouth, S., and Kida, H. (2020) Treatment-Emergent Influenza Variant Viruses With Reduced Baloxavir Susceptibility: Impact on Clinical and Virologic Outcomes in Uncomplicated Influenza. *J Infect Dis* **221**, 346-355

166. Ikematsu, H., Hayden, F. G., Kawaguchi, K., Kinoshita, M., de Jong, M. D., Lee, N., Takashima, S., Noshi, T., Tsuchiya, K., and Uehara, T. (2020) Baloxavir Marboxil for Prophylaxis against Influenza in Household Contacts. *N Engl J Med* **383**, 309-320

167. Guilligay, D., Tarendeau, F., Resa-Infante, P., Coloma, R., Crepin, T., Sehr, P., Lewis, J., Ruigrok, R. W., Ortin, J., Hart, D. J., and Cusack, S. (2008) The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat Struct Mol Biol* **15**, 500-506

168. te Velthuis, A. J., and Fodor, E. (2016) Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis. *Nature Reviews Microbiology* **14**, 479-493

169. Dias, A., Bouvier, D., Crepin, T., McCarthy, A. A., Hart, D. J., Baudin, F., Cusack, S., and Ruigrok, R. W. (2009) The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **458**, 914-918

170. Fodor, E., Crow, M., Mingay, L. J., Deng, T., Sharps, J., Fechter, P., and Brownlee, G. G. (2002) A single amino acid mutation in the PA subunit of the influenza virus RNA polymerase inhibits endonucleolytic cleavage of capped RNAs. *J Virol* **76**, 8989-9001

171. Kobayashi, M., Toyoda, T., and Ishihama, A. (1996) Influenza virus PB1 protein is the minimal and essential subunit of RNA polymerase. *Archives of virology* **141**, 525-539

172. Omoto, S., Speranzini, V., Hashimoto, T., Noshi, T., Yamaguchi, H., Kawai, M., Kawaguchi, K., Uehara, T., Shishido, T., Naito, A., and Cusack, S. (2018) Characterization of influenza virus variants induced by treatment with the endonuclease inhibitor baloxavir marboxil. *Sci Rep* **8**, 9633

173. Pommier, Y., Johnson, A. A., and Marchand, C. (2005) Integrase inhibitors to treat HIV/AIDS. *Nat Rev Drug Discov* **4**, 236-248

174. Engelman, A. N. (2019) Multifaceted HIV integrase functionalities and therapeutic strategies for their inhibition. *J Biol Chem* **294**, 15137-15157

175. Ilina, T., Labarge, K., Sarafianos, S. G., Ishima, R., and Parniak, M. A. (2012) Inhibitors of HIV-1 Reverse Transcriptase-Associated Ribonuclease H Activity. *Biology (Basel)* **1**, 521-541

176. Powdrill, M. H., Bernatchez, J. A., and Gotte, M. (2010) Inhibitors of the Hepatitis C Virus RNA-Dependent RNA Polymerase NS5B. *Viruses* **2**, 2169-2195

177. Yoshino, R., Yasuo, N., and Sekijima, M. (2019) Molecular Dynamics Simulation reveals the mechanism by which the Influenza Cap-dependent Endonuclease acquires resistance against Baloxavir marboxil. *Sci Rep* **9**, 17464

178. Thierry, E., Guilligay, D., Kosinski, J., Bock, T., Gaudon, S., Round, A., Pflug, A., Hengrung, N., El Omari, K., Baudin, F., Hart, D. J., Beck, M., and Cusack, S. (2016) Influenza Polymerase Can Adopt an Alternative Configuration Involving a Radical Repacking of PB2 Domains. *Molecular Cell* **61**, 125-137

179. Rao, P., Yuan, W., and Krug, R. M. (2003) Crucial role of CA cleavage sites in the cap-snatching mechanism for initiating viral mRNA synthesis. *EMBO J* **22**, 1188-1198

180. Cianci, C., Tiley, L., and Krystal, M. (1995) Differential activation of the influenza virus polymerase via template RNA binding. *Journal of virology* **69**, 3995-3999

181. Hagen, M., Chung, T. D. Y., Butcher, J. A., and Krystal, M. (1994) Recombinant Influenza-Virus Polymerase - Requirement of Both 5' and 3' Viral Ends for Endonuclease Activity. *Journal of Virology* **68**, 1509-1515

182. Li, M. L., Ramirez, B. C., and Krug, R. M. (1998) RNA-dependent activation of primer RNA production by influenza virus polymerase: different regions of the same protein subunit constitute the two required RNA-binding sites. *EMBO J* **17**, 5844-5852

183. Copeland, R. A. (2000) A practical introduction to structure, mechanism, and data analysis. *Enzymes, 2nd ed. John Wiley & Sons, New York, NY*, 104

184. Takashita, E., Morita, H., Ogawa, R., Nakamura, K., Fujisaki, S., Shirakura, M., Kuwahara, T., Kishida, N., Watanabe, S., and Odagiri, T. (2018) Susceptibility of Influenza Viruses to the Novel Cap-Dependent Endonuclease Inhibitor Baloxavir Marboxil. *Front Microbiol* **9**, 3026

185. Beilhartz, G. L., Wendeler, M., Baichoo, N., Rausch, J., Le Grice, S., and Gotte, M. (2009) HIV-1 Reverse Transcriptase Can Simultaneously Engage Its DNA/RNA Substrate at Both DNA Polymerase and RNase H Active Sites: Implications for RNase H Inhibition. *Journal of Molecular Biology* **388**, 462-474

186. Davies, J. F., 2nd, Hostomska, Z., Hostomsky, Z., Jordan, S. R., and Matthews, D. A. (1991) Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. *Science* **252**, 88-95

187. Shaw-Reid, C. A., Munshi, V., Graham, P., Wolfe, A., Witmer, M., Danzeisen, R., Olsen, D. B., Carroll, S. S., Embrey, M., Wai, J. S., Miller, M. D., Cole, J. L., and Hazuda, D. J. (2003) Inhibition of HIV-1 ribonuclease H by a novel diketo acid, 4-[5-(benzoylamino)thien-2-yl]-2,4-dioxobutanoic acid. *J Biol Chem* **278**, 2777-2780

188. Budihas, S. R., Gorshkova, I., Gaidamakov, S., Wamiru, A., Bona, M. K., Parniak, M. A., Crouch, R. J., McMahon, J. B., Beutler, J. A., and Le Grice, S. F. (2005) Selective inhibition of HIV-1 reverse transcriptase-associated ribonuclease H activity by hydroxylated tropolones. *Nucleic Acids Res* **33**, 1249-1256

189. Himmel, D. M., Maegley, K. A., Pauly, T. A., Bauman, J. D., Das, K., Dharia, C., Clark, A. D., Jr., Ryan, K., Hickey, M. J., Love, R. A., Hughes, S. H., Bergqvist, S., and Arnold, E. (2009) Structure of HIV-1 reverse transcriptase with the inhibitor beta-Thujaplicinol bound at the RNase H active site. *Structure* **17**, 1625-1635

190. Kumar, G., Cuypers, M., Webby, R. R., Webb, T. R., and White, S. W. (2021) Structural insights into the substrate specificity of the endonuclease activity of the influenza virus cap-snatching mechanism. *Nucleic Acids Res* **49**, 1609-1618

191. Canada, G. o. (2021) Regulatory Decision Summary - Veklury - Health Canada.

192. Gordon, C. J., Tchesnokov, E. P., Woolner, E., Perry, J. K., Feng, J. Y., Porter, D. P., and Gotte, M. (2020) Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. *Journal of Biological Chemistry*

193. Tchesnokov, E. P., Raeisimakiani, P., Ngure, M., Marchant, D., and Götte, M. (2018) Recombinant RNA-dependent RNA polymerase complex of Ebola virus. *Scientific reports* **8**, 1-9

194. Gordon, C. J., Tchesnokov, E. P., Feng, J. Y., Porter, D. P., and Götte, M. (2020) The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J Biol Chem* **295**, 4773-4779

195. Sainsbury, S., Bernecky, C., and Cramer, P. (2015) Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**, 129-143

196. Jin, Z., Smith, L. K., Rajwanshi, V. K., Kim, B., and Deval, J. (2013) The ambiguous base-pairing and high substrate efficiency of T-705 (Favipiravir) Ribofuranosyl 5'-triphosphate towards influenza A virus polymerase. *PLoS One* **8**, e68347

197. Shannon, A., Le, N. T., Selisko, B., Eydoux, C., Alvarez, K., Guillemot, J. C., Decroly, E., Peersen, O., Ferron, F., and Canard, B. (2020) Remdesivir and SARS-CoV-2: Structural requirements at both nsp12 RdRp and nsp14 Exonuclease active-sites. *Antiviral Res* **178**, 104793

198. Kokic, G., Hillen, H. S., Tegunov, D., Dienemann, C., Seitz, F., Schmitzova, J., Farnung, L., Siewert, A., Hobartner, C., and Cramer, P. (2021) Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. *Nat Commun* **12**, 279

199. Golosov, A. A., Warren, J. J., Beese, L. S., and Karplus, M. (2010) The mechanism of the translocation step in DNA replication by DNA polymerase I: a computer simulation analysis. *Structure* **18**, 83-93

200. Fidalgo da Silva, E., and Reha-Krantz, L. J. (2007) DNA polymerase proofreading: active site switching catalyzed by the bacteriophage T4 DNA polymerase. *Nucleic Acids Res* **35**, 5452-5463

Appendices 1.

A.

UL51

Accession	Description	Score	Coverage	# Proteins	# Unique Peptides
P16792	Tripartite terminase subunit 2 OS=Human cytomegalovirus (strain AD169) GN=TRM2 PE=3 SV=1 - [TRM2_HCMVA]	1376.94	46.50	1	10

UL56

Accession	Description	Score	Coverage	# Proteins	# Unique Peptides
Q2F7S6	Tripartite terminase subunit 1 OS=Human cytomegalovirus GN=UL56 PE=3 SV=1 - [Q2F7S6_HCMV]	4767.24	91.65	1	6

UL89

Accession	Description	Score	Coverage	# Proteins	# Unique Peptides
A0A0G2T9T6	Tripartite terminase subunit 3 OS=Human cytomegalovirus GN	52.15	20.47	34	11

B.

FluB-ht

Accession	Description	Score	Coverage	# Proteins	# Unique Peptides
PB2	PB2	1272.94	73.86	1	75
PA	PA	1050.80	56.32	1	56
PB1	PB1	978.85	52.85	1	45

C.

nsp10

Accession	Description	Score	Coverage	# Proteins	# Unique Peptides
SARS2NSP10	SARS2NSP10	204.94	99.31	1	11

nsp14

Accession	Description	Score	Coverage	# Proteins	# Unique Peptides
SARS2NSP14	SARS2NSP14	453.98	60.89	1	39

Figure A1. LC-MS/MS results for the protein complexes used in this thesis.