

Georgia State University

ScholarWorks @ Georgia State University

Chemistry Dissertations

Department of Chemistry

12-16-2020

Unveiling the Polymerase Complex of Negative Stranded RNA Viruses

Chelsea Severin

Follow this and additional works at: https://scholarworks.gsu.edu/chemistry_diss

Recommended Citation

Severin, Chelsea, "Unveiling the Polymerase Complex of Negative Stranded RNA Viruses." Dissertation, Georgia State University, 2020.

doi: <https://doi.org/10.57709/20469146>

This Dissertation is brought to you for free and open access by the Department of Chemistry at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Chemistry Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

UNVEILING THE POLYMERASE COMPLEX OF NEGATIVE STRANDED RNA VIRUSES

by

CHELSEA SEVERIN

Under the Direction of Ming Luo, Ph.D.

ABSTRACT

Negative stranded RNA viruses (NSVs) are among the most common human pathogens which cause pandemics and epidemics. This group includes many notable members such as influenza, mumps and Ebola viruses. These viruses are identifiable by their negative polarity genome which is associated with the nucleocapsid (NP) protein and assembled into higher order structures. The RNA-nucleocapsid complex or ribonucleoprotein (RNP) serves as the template for transcription and replication by the viral RNA-dependent RNA polymerase (vRdRp). Though progress has been made in the study of these viruses, knowledge is lacking with regards to the polymerase complex. Here, we utilize structural biology and mutational analysis to identify components of the polymerase complex that will be targets for drug design.

NSVs typically cause high mortality outbreaks by transmission from animal reservoirs. In fact, in 2013 H7N9 avian influenza A virus emerged as human infections and in 2017 the number of infections raised to 688. This reaffirms that influenza virus is a global health threat and requires antiviral drugs in the effort to control influenza virus. Frequently used anti-influenza drugs target neuraminidase; however, there have been strains that show resistance to these neuraminidase inhibitors. The PB2cap binding domain of the influenza RNA polymerase is an innovative target for development of anti-influenza drugs. In this study, we have solved the crystal structure of the PB2cap binding domain of influenza A H1N1 virus alone and in complex with its binding partner. Utilizing this structure, we have identified critical interactions that will aid in the design of antivirals.

The emergence of mumps virus outbreaks throughout the United States in the past five years indicates that the MMR vaccine is not the most efficient source of protection and reaffirms the need for inhibitors that target the virus. Here, we have utilized cryogenic electron microscopy (cryoEM) to analyze the RNA encapsidation of mumps virus nucleocapsid and mutational analysis of the phosphoprotein to probe the interactions involved in uncoiling the nucleocapsid. This data adds to the available knowledge about mumps virus infection and could potentially aid in the design of inhibitors.

INDEX WORDS: Negative-stranded RNA virus, Influenza virus, Mumps virus, RNA-dependent RNA polymerase, Nucleocapsid, Phosphoprotein, Antiviral, Drug design

UNVEILING THE POLYMERASE COMPLEX OF NEGATIVE STRANDED RNA VIRUSES

by

CHELSEA SEVERIN

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2020

Copyright by
Chelsea Severin
2020

UNVEILING THE POLYMERASE COMPLEX OF NEGATIVE STRANDED RNA VIRUSES

by

CHELSEA SEVERIN

Committee Chair: Ming Luo

Committee: Jenny Yang

Kathryn Grant

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

December 2020

DEDICATION

I owe everything that I am and everything that I am becoming to the strongest women I know - my mother Lynda and aunt Shirley. I would not have been able to complete this journey without your support.

“As iron sharpens iron, so one man sharpens another” Proverbs 27:17

ACKNOWLEDGEMENTS

After graduating with my bachelor's degree, I knew I wanted to do work that impacted the world and that a PhD degree would be a stepping-stone on that journey. I would not have been able to see my dream become a reality it was not for the supportive group of individuals that have provided a listening ear, a comforting shoulder and stern words. None of this work would have been possible without the financial support of the Chemistry Department and the Center for Diagnostics and Therapeutics at Georgia State University.

I am deeply indebted to my advisor Dr. Ming Luo, who has been patient with me throughout my PhD while I found my footing. When I started in August 2014, I was doe-eyed and blissfully unaware about the effort that this degree would require. Dr. Luo never allowed me to settle for less and taught me what I needed to do to succeed as a doctoral researcher. Through his actions, he has shown me what it takes to be not only a good researcher but a good mentor. I would also like to thank my committee members Dr. Katherine Grant and Dr. Jenny Yang. Dr. Grant and Dr. Yang exemplify what it means to be a successful woman in science.

To the Luo lab members, past and present, thank you. I do not think I would have been able to accomplish anything if it was not for the amazing group of individuals who I have shared a workspace with. Over these past years I have gotten to know some amazing men and women some of which have become close friends. Thank you all for being a sounding board for my ideas and for providing valued advice.

Lastly, I want to salute my family. To my cancer fighting mother, I am so grateful to have you around. Mom, you have shown me the meaning of resilience. I am also thankful for my aunt Shirley who has supported me in all ways throughout this time. Finally, to my dear husband

Rotando you are the epitome of a compassionate spouse. You are more supportive than I could ever give you credit for. Thank you all for sticking through this journey with me.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS		V
LIST OF TABLES		XI
LIST OF FIGURES		XII
LIST OF ABBREVIATIONS		XIII
1. INTRODUCTION		1
1.1 Group V Viruses: Negative Strand RNA Viruses		2
1.2 <i>Orthomyxoviridae</i> – Influenza Viruses		3
<i>1.2.1 Overview of Virus Structure and Genome Organization</i>		4
<i>1.2.2 Influenza A Virus Lifecycle</i>		5
<i>1.2.3 Inhibitors of Influenza Virus</i>		8
1.3 <i>Paramyxoviridae</i> – Mumps Virus		11
<i>1.3.1 Mumps Viral Replication Cycle</i>		12
<i>1.3.2 Polymerase Complex</i>		13
<i>1.3.3 Mumps Virus Prevention - MMR Vaccine</i>		15
1.4 Scope of Research		16
2. THE CAP-BINDING SITE OF INFLUENZA VIRUS PROTEIN PB2 AS A DRUG TARGET		26
2.1 Abstract		26
2.2 Introduction		27

2.3	Materials and methods	30
2.3.1	<i>Protein expression and purification</i>	30
2.3.2	<i>Crystallization and structure analyses</i>	30
2.4	Results and discussion	31
2.4.1	<i>The cap-binding site in PB2 of A/California/07/2009</i>	31
2.4.2	<i>Interactions between m⁷GTP and the cap-binding site</i>	33
2.4.3	<i>Comparisons with PB2cap inhibitors</i>	36
2.4.4	<i>Differences from PB2cap of influenza B virus</i>	37
2.4.5	<i>Comparisons with human cap-binding proteins</i>	38
3.	RELEASING THE GENOMIC RNA SEQUESTERED IN THE MUMPS VIRUS NUCLEOCAPSID	47
3.1	Abstract	47
3.1.1	<i>Importance</i>	48
3.2	Introduction	48
3.3	Materials and methods	51
3.3.1	<i>Expression and purification of recombinant MuV N protein</i>	51
3.3.2	<i>Cryo-EM structure of the empty MuV N assembly.</i>	51
3.3.3	<i>Minigenome assays</i>	52
3.4	Results	53
3.4.1	<i>Structure of a truncated empty capsid.</i>	53

3.4.2	<i>Comparison with the authentic nucleocapsid.</i>	54
3.4.3	<i>How is the encapsidated genomic RNA unveiled by the viral polymerase?</i>	55
3.4.4	<i>Residues in the loop-helix $\alpha 7$ are critical.</i>	57
3.5	Discussion	58
4.	PROBING THE ROLE OF THE PHOSPHOPROTEIN AMINO TERMINAL END IN THE UNCOILING OF THE NUCLEOCAPSID	64
4.1	Abstract	64
4.2	Introduction	64
4.3	Materials and Methods	67
4.3.1	<i>Molecular Biology</i>	67
4.3.2	<i>Computer aided studies</i>	68
4.3.3	<i>Protein expression and purification</i>	68
4.3.4	<i>Protein binding assay</i>	70
4.3.5	<i>Thermal shift assay</i>	70
4.4	Results	71
4.4.1	<i>Identification of N-terminal residues of P involved in the potential NLP binding site</i>	71
4.4.2	<i>Residues in loop are necessary for NLP-P110 interaction in pulldown assay</i>	72
4.4.3	<i>Effects of targeted mutations on thermal stability of NLP-P110 interaction</i>	72
4.4.4	<i>Alanine variants affect the thermal release of RNA from the NLP</i>	73

4.5	Discussion.....	74
5.	RESEARCH CHALLENGES AND IMPACT.....	81
5.1	Crystallization of PB2 cap binding protein	81
5.2	Implications for Design of inhibitors that target influenza viruses A & B.....	81
5.3	Viral RNA sequestered within the Paramyxovirus Nucleocapsid	82
5.4	Uncoiling the MuV Nucleocapsid	83
5.5	Implications for design of antivirals and vaccine development of MuV.....	83
6.	CONCLUSIONS	84
	REFERENCES.....	86

LIST OF TABLES

Table 1-1 FDA Approved Influenza Drugs	21
Table 2-1 Data-collection and refinement statistics.....	41
Table 4-1 Summary of Affinities and Thermal Release of RNA	80

LIST OF FIGURES

Figure 1-1 Baltimore Viral Classification	18
Figure 1-2 Genome Organization of Influenza Viruses	19
Figure 1-3 Cap Snatching Polymerase	20
Figure 1-4 Drug Resistance in Neuraminidase	22
Figure 1-5 Drug Resistance in M2 channel	23
Figure 1-6 The Family Paramyxoviridae	24
Figure 1-7 Mumps virus Outbreaks	25
Figure 2-1 The cap-binding site in PB2cap	40
Figure 2-2 Structure of Influenza A cap binding site	42
Figure 2-3 Comparing bound and unbound PB2cap	43
Figure 2-4 Comparison of the PB2cap cap-binding site	44
Figure 2-5 Comparing Influenza A and B cap binding	45
Figure 2-6 Human cap binding proteins	46
Figure 3-1 CryoEM Structure of N₃₇₉	60
Figure 3-2 Comparing empty capsid with that of the authentic	61
Figure 3-3 RNA binding per N subunit	62
Figure 3-4 Minigenome assay of mutants	63
Figure 4-1 Homology Model of N_{CORE}-P110 complex	77
Figure 4-2 Pulldown assay	78
Figure 4-3 Thermal Stability Assay	79

LIST OF ABBREVIATIONS

Centers for Disease Control (CDC)
Cryogenic electron microscopy (cryoEM)
Deoxyribonucleic Acid (DNA)
Double Stranded DNA (dsDNA)
Double Stranded RNA (dsRNA)
Glycoprotein (G)
Large protein (L)
Matrix-protein (M)
Messenger Ribonucleic Acid (mRNA)
Mumps Virus (MuV)^a
Negative stranded RNA viruses (NSVs)
Nucleocapsid protein (NP)
Nucleocapsid-like Particle (NLP)
Phosphoprotein (P)
Polymerase Basic 1 protein (PB1)
Polymerase Basic 2 cap binding protein (PB2cap)
Protein Databank (PDB)
Ribonucleic Acid (RNA)
Ribonucleoprotein (RNP)
Single Stranded DNA (ssDNA)
Single Stranded RNA (ssRNA)
Thermal Shift Assay (TSA)
Viral Ribonucleic Acid (vRNA)
Viral RNA Dependent RNA polymerase (vRdRp)
World Health Organization (WHO)

1. INTRODUCTION

The current coronavirus pandemic highlights how unprepared we are to tackle emerging viruses. These pathogens are notorious for having a negative impact on public health and agricultural commerce. The International Committee on Taxonomy of Viruses (ICTV) is responsible for the formal taxonomic classification of viruses. Although this method of classification is necessary to organize newly discovered viruses to track their evolutionary relationships, the Baltimore system is a simpler method of grouping viruses (1).

In his landmark article, David Baltimore proposed a viral classification system based on the viral genome and the method of replication (**Figure 1-1**). For efficient viral infection, all viruses must synthesize mRNA to produce viral proteins. Therefore, Baltimore designed his classification system to give mRNA a central role in viral replication.

Group I viruses are pathogens which have a double-stranded DNA (dsDNA) genome. These viruses utilize the host cell replication machinery (RNA polymerase II) to transcribe the viral DNA into mRNA. Examples of Group I viruses include herpes simplex I (HSV-1) and varicella zoster virus (2, 3). Viruses belonging to the group II include the family *Parvoviridae* which includes canine parvovirus which is a highly contagious infection in dogs. This group possesses a single-stranded DNA (ssDNA) which replicates mostly within the nucleus forming a dsDNA intermediate in the process.

Viruses of group III to group V contain an RNA genome and replicate primarily in the cytoplasm of the host cell. Group III viruses possess a double-stranded RNA (dsRNA) genome which is used to synthesize mRNA and protein synthesis. Group IV viruses have a positive sense single-stranded RNA (ssRNA) as their genome. For replication, these viruses must generate a negative sense anti-genome which acts as a template for generation of the positive sense genome.

Additionally, the positive sense genome can be directly transcribed by host ribosomes to produce viral proteins. Examples of this class of viruses include the *Flaviviridae* and *Picornaviridae* families. The *Flaviviridae* family includes several important human pathogens such as Zika virus, dengue virus and yellow fever virus. Viruses which possess a negative sense single-stranded RNA (nsRNA) as their genome are placed into group V. Prototypical members of this group include vesicular stomatitis virus (VSV), mumps virus (MuV) and influenza viruses. The viral genome of group V viruses must first be transcribed into the positive sense for translation of viral proteins.

Reverse transcription is utilized by groups VI and VII viruses to replicate their genomes. Viruses belonging to group VI have a positive sense ssRNA as their viral genome. These viruses use reverse transcriptase to transcribe their RNA genome into a DNA/RNA hybrid intermediate which serves as a template to form dsDNA which is then converted to mRNA. One of the most notable members of this group is human immunodeficiency virus (HIV). Unlike group VI viruses, viruses classified into group VII possess a dsDNA genome and utilize reverse transcription to generate mRNA.

1.1 Group V Viruses: Negative Strand RNA Viruses

Viruses belonging to group V are also referred to as Negative Stranded RNA Viruses (NSVs). These viruses consist of many notable human pathogens such as Ebola, influenza and mumps viruses and agricultural pathogens such as VSV. These viruses utilize a negative polarity ssRNA as their genome which is associated with nucleocapsid (N) protein to form a ribonucleoprotein (RNP) complex. NSVs can further be divided into two groups based on the segmented nature of their genome. The non-segmented NSVs (nsNSVs) include the *Bornaviridae* (Borne disease virus), *Filoviridae* (Ebola virus), *Paramyxoviridae* (mumps and

measles viruses) and *Rhabdoviridae* (rabies virus) families. Whereas the segmented NSVs (sNSVs) include the *Arenaviridae* (Machupo virus), *Bunyaviridae* (Rift valley fever virus) and *Orthomyxoviridae* (influenza viruses).

Despite their differences in genome segments, the viral RNA of NSVs is encapsidated by the nucleocapsid protein during the span of the viral life cycle. This ribonucleoprotein (RNP) complex is the only functional form of the genome for viral replication. In nsNSVs, the RNP complexes form helical or ring shaped structures which are typically linear and relatively rigid (4). The RNPs of sNSVs have a more flexible circular conformation which is induced by the binding of the polymerase to the 5' and 3' end of the viral RNA (vRNA) segments (5, 6).

Although both nsNSVs and sNSVs package a polymerase within the virus particle, they differ in the capping mechanism of mRNA transcripts. For sNSVs transcription is initiated by acquiring a capped primer derived from the host cell nascent mRNA (7–11). The polymerase of nsNSVs possess the ability to cap mRNA transcripts in a mechanism that is distinct from viral and eukaryotic systems (10, 12, 13). The unique nature of the polymerase complex of NSVs make it a viable target for the development of inhibitors that disrupt viral replication and transcription.

1.2 *Orthomyxoviridae* – Influenza Viruses

Influenza viruses are amongst the most contagious respiratory viruses that infect humans and are notorious for yearly outbreaks. In addition to annual spread, occasionally these viruses are responsible for severe pandemics. The 1918 H1N1 often referred to as “Spanish flu” was responsible for about 50 million deaths worldwide with about 675,000 occurring in the United States (14, 15). More recently, the 2009 H1N1 pandemic resulted in an estimated 61 million infections and over 12000 deaths in the United States (16). These viruses pose a significant

public health burden; therefore, an efficient arsenal of therapeutics is necessary to treat this disease.

Influenza viruses belong to the *Orthomyxoviridae* family of viruses. These viruses are characterized by their segmented, negative-sense, single-stranded RNA genome. Within this family there are five different genera: the influenza viruses A, B, C and D; Thogotovirus; and Isavirus. Influenza A viruses are responsible for epidemics and pandemics and are known pathogens of humans, horses and fowls. Influenza B viruses occasionally cause epidemics and only infect humans. Influenza C viruses are known human and pig pathogens, but they rarely cause disease. Influenza D virus was recently isolated in cattle and swine; no human infections have been observed. Although, Thogotovirus and Isavirus pose a significant threat to human health and agricultural economies, the primary focus of this review will be on influenza A viruses and to a lesser extent influenza B virus. Furthermore, influenza C viruses are morphologically and genetically distinct from influenza A and B viruses and are generally asymptomatic and will not be discussed extensively.

1.2.1 Overview of Virus Structure and Genome Organization

Types A and B influenza viruses are indistinguishable when analyzed by electron microscopy, but the virion organization differs slightly (6, 17). Influenza A virus has the hemagglutinin (HA), neuraminidase (NA) and M2 proteins inserted into the host cell derived lipid membrane. The matrix protein (M1) lies beneath the envelope and the core of the particle consists of 8 viral RNA segments, the NEP/NS2 (nuclear export protein/nonstructural protein 2) and the polymerase complex (polymerase acid [PA]; polymerase basic 1 [PB1]; and polymerase basic 2 [PB2]). Influenza B virions have four proteins in their envelope: HA, NA, NB and BM2

(18). With regards to genome organization, influenza A and B viruses both possess only eight viral RNA segments that code for one or more proteins (**Figure 1-2**).

1.2.2 Influenza A Virus Lifecycle

1.2.2.1 Virus Entry and Nuclear Import

Viral infection is initiated by HA binding to sialic acid moieties on the cell surface. Human influenza viruses typically bind to N-acetylneuraminic acid attached to the penultimate galactose sugar by an $\alpha 2,6$ linkage. This binding induces endocytosis of the virion through clathrin-mediated process or micropinocytosis (18).

Once the virion is enclosed in an endosome, fusion of the viral membrane is induced by structural changes in HA due to the low pH. The HA0 precursor is cleaved into HA1 and HA2 subunits. The structural change of the HA exposes the HA2 fusion peptide which enables the HA to interact with the endosomal membrane. Additionally, the conformational change in HA opens the M2 ion channel which allows an influx of H^+ ions from the endosome into the viral core. This decrease in pH releases the viral RNP (vRNP) from the M1 protein and allows free vRNP complex to be released into the cytoplasm

Viral replication and transcription are dependent on the host cell's nuclear functions; therefore, upon release into the cytoplasm the vRNPs are transported to the nucleus. The vRNP complex consists of the vRNA encapsidated by NP and the polymerase complex which binds to the ends of the vRNA. Each protein within the vRNP contains a nuclear localization signal (NLS); however, it seems that the NLS on NP is necessary for viral RNA import (19). The transport of vRNPs across the nuclear membrane is mediated by the importin α family. Karyopherin α binds directly to the NLS on NP and then recruits karyopherin β which opens the nuclear pore and allows vRNP transport into the nucleus.

1.2.2.2 *Cap Snatching Viral RdRp*

Once within the nucleus, the viral RNA serves as a template for mRNA synthesis. As previously described, the vRNP exists as a complex in which the RNA is encapsidated by NP and forms a helical hairpin that is attached to the polymerase complex. It is the NP coated RNA not the naked RNA that serves as a template for transcription. Influenza viral RNA dependent RNA polymerase (vRdRp) is composed of three proteins (**Figure 1-3**): PB1, PB2 and PA (20). This polymerase complex hijacks the host cell machinery during the synthesis of mRNA, a process referred to as cap snatching. PB1 initiates transcription by binding to the terminal ends of the vRNA. This binding induces a conformational change in the complex which allows the PB2 protein to bind to the cap of host cell pre-mRNAs (21). Studies have shown that the vRdRp interacts with the large subunit of cellular DNA-dependent RNA polymerase II (Pol II) in its transcriptionally active, hyperphosphorylated form. The PB2 protein initiates transcription by binding to the 5',7-methylguanosine cap of nascent cell mRNA molecules. PA endonuclease domain then cleaves the cellular mRNA about 10 to 15 nucleotides downstream of the cap. This capped primer is then transferred to the active site of the PB1 subunit and is used to initiate viral transcription. The addition of a guanine or cytosine residue to the end of the primer initiates elongation by PB1 until a stretch of five to seven uridine residues are encountered which is a polyadenylation signal.

In addition to serving as a template for mRNA synthesis, the vRNA also acts as a template for the synthesis of positive sense replication intermediate RNA which is used for RNA genome synthesis. This process is a two-step primer independent mechanism. The switch from transcriptase to replicase mode is poorly understood, but some studies suggest that the concentration of soluble NP plays a role (18). Firstly, the vRNPs are used to generate a positive

sense complementary RNA (cRNA) intermediate. To begin, the 3' end of the vRNA is transported to the active site of the complex, where a pppApG dinucleotide is aligned to residues U1 and C2 of the vRNA. The PB1 protein elongates the cRNA strand and the cRNA is encapsidated by NP as it is synthesized. This complementary RNP is then used as a template for vRNA synthesis. For vRNA production, a pppApG dinucleotide is formed on residues U4 and C5. The pppApG product is then subsequently U1 and C2 for elongation. It is unknown at where and how the 5' end of the cRNA binds to the polymerase; however, several models have been proposed. The newly synthesized vRNA is encapsidated by NP.

1.2.2.3 Nuclear Export of Ribonucleoproteins

The plasma membrane is the site of assembly and budding for influenza viruses. Therefore, it is necessary to export the newly synthesized vRNPs to the cytoplasm. Transport of vRNPs is facilitated by two proteins: M1 and NEP/NS2. Current models of nuclear export suggest that M1 binds to vRNPs and causes dissociation of vRNPs from the nuclear matrix, thereby downregulating transcription(22, 23). Furthermore, it has been shown that the NEP/NS2 protein interacts with chromosome region maintenance 1 (CRM1) and several nucleoporins. The data suggests that a 'daisy-chain' complex of vRNP, M1, NEP/NS2 and CRM1 is formed for nuclear export (24).

Influenza virus particles assemble and bud from the apical membrane of polarized cells. The HA, NA, and M2 proteins are transported to the assembly site where later the vRNPs are translocated by Rab11. The eight segments of the viral genome are packaged, and budding is induced by the M1 protein. During budding, HA binds to the sialic acid containing receptors on the cell surface. For release of the virion from the infected cell, NA cleaves these sialic acid residues.

1.2.3 Inhibitors of Influenza Virus

The first line of defense against influenza epidemics and pandemics is the annual flu vaccine. This seasonal vaccine is designed each year using surveillance data of circulating viruses and predictions about the ones which are most likely to spread in the upcoming season; however, there are some limitations of the annual vaccine (25). Candidate viruses for vaccines are primarily produced in eggs which provides a problem for preparing vaccines for viruses that grow poorly using this method (e.g. H3N2 viruses). Furthermore, FDA regulations makes it difficult to prepare a candidate vaccine for viruses that spread later in the season. Fortunately, there are three classes of FDA approved influenza antivirals: neuraminidase inhibitors; M2 ion channel inhibitors; and cap-dependent endonuclease inhibitors (**Table 1-1**).

1.2.3.1 Neuraminidase inhibitors

Neuraminidase supports viral infection by cleaving sialic acid on cell surface receptors to allow release of newly formed viral progeny. The 3D structure of NA allowed structure guided design of therapeutics that inhibit the enzymatic activity of the protein. Historically, the first neuraminidase inhibitor (NAI) discovered was 2,3-dehydro-2-deoxy-N-acetylneuraminic acid (DANA); however, it has low inhibition of NA (31). The crystal structure of NA in complex with DANA served as a basis for the design of the three currently approved NAIs: zanamivir (Relenza); oseltamivir (Tamiflu); and peramivir (Rapivab).

The crystal structure of the NA-DANA complex revealed that the active site of the protein contained an empty pocket and the C4 hydroxyl group position could accommodate a larger substitution. Subsequently, zanamivir was designed with a 4-guanidino group at the C4 hydroxyl group position to improve the affinity for the enzyme. Zanamivir is indicated for treatment of type A and B viruses in adults and children seven years and older. The drug is

administered directly to the site of infection as an oral inhalation powder due to its highly polar nature and rapid excretion.

Oseltamivir was developed as an orally bioavailable inhibitor of NA. The drug was based upon the structure of zanamivir, in which the C4 hydroxyl group was replaced by an amino group and the glycerol side chain was replaced with a pentyl ether side chain (18, 26). Although oseltamivir was designed to increase oral bioavailability, studies showed that in its active form the drug had poor oral bioavailability. Therefore, the drug is delivered as the ethyl-ester prodrug which is converted to the active drug by hepatic esterases.

Peramivir is the third FDA approved NAI. This drug is a cyclopentane derivative with a C4-guanidino substitution and bulky hydrophobic side chain which resembles the structures of both zanamivir and oseltamivir (18). Peramivir is administered intravenously for treatment of influenza A viruses.

Resistance to NAIs is primarily due to mutations in NA. To date, viruses have been isolated from patients which have reduced susceptibility to all approved NAIs (27). A(H3N2) viruses with R292K and E119V mutants have reduced susceptibility to oseltamivir and peramivir. Zanamivir remains effective against E119V mutants but has reduced binding to R292K mutants. Additionally, an influenza B isolate with S250G mutation was found to have decreased effectiveness; however, oseltamivir remained a potent inhibitor of NA (28).

A H274Y mutation near the active site of the enzyme confers high resistance to oseltamivir in N1 viruses including 2009 pandemic H1N1. This H275Y mutant confers resistance to peramivir to a lesser extent but viruses with this mutation remain susceptible to treatment with zanamivir. Structural studies indicate that the tyrosine residue moves the carboxyl

group of Glu27 further into the binding site (**Figure 1-4**) which disrupts the hydrophobic pocket (29).

1.2.3.2 M2 Ion Channel Inhibitors Resistance

Amantadine (Symmetrel) and rimantadine (Flumadine), also referred to as the adamantanes, inhibit viral infection by binding to the M2 protein thereby preventing uncoating (30). Adamantanes were the first approved drugs for treatment of influenza virus infections. These antivirals are indicated for treatment for type A viruses and are not effective against influenza B viruses. Approximately 100% of circulating influenza A viruses have conferred resistance against the adamantanes. As a result, the Centers for Disease Control and Prevention has issued an advisory against the use of the adamantanes for influenza A treatment. The reduced susceptibility of influenza A virus to the adamantanes has been associated with a S31N mutation (31, 32). The presence of Asn seems to open the pore to allow protons to flow through the channel in the presence of the drug (**Figure 1-5**).

1.2.3.3 Cap-dependent endonuclease inhibitors

At the time of our studies, there were no FDA approved drugs that targeted the influenza polymerase complex. However, in 2018, baloxavir marboxil (BXM), a prodrug of baloxavir acid (BXA), received regulatory approval for the treatment of influenza A and B viruses in patients 12 and older. BXM inhibits the PA endonuclease of the polymerase complex and prevents viral transcription. The drug was developed by rational drug design using the pharmacophore structure of the strand transfer inhibitor of HIV, dolutegravir.

Reduced susceptibility of BXM has been observed in type A and B viruses carrying an I38T mutation. Structural analyses of the mutant and wild type PA bound to BXA illustrated the effects of this change on drug binding (33). The structure of the PA I38T mutant bound to BXA

indicates that the presence of the more polar threonine residue at position 38 reduces Van der Waals interactions with the compound.

1.3 *Paramyxoviridae* – Mumps Virus

The *Paramyxoviridae* family are enveloped viruses which possess a non-segmented, single stranded negative polarity RNA genome. This family includes many notable pathogens which have an impact on agricultural livestock and public health such as measles, mumps and Newcastle disease. The family *Paramyxoviridae* is further divided into the *Avulavirinae*, *Rubulavirinae*, *Metaparamyxovirinae* and *Orthoparamyxovirinae* subfamilies (**Figure 1-6**).

Mumps virus (MuV) belongs to the subfamily *Rubulavirinae* and *Orthorubulavirus* genus. The first documented infection of MuV was described by Hippocrates in the 5th century BCE (34), in which he described an illness associated with parotitis and orchitis. However, viral etiology was not described until 1934 when Johnson and Goodpasture showed that the filter sterilized virus particles could be transmitted from infected patients to rhesus monkeys (35). Viral infection is characterized by inflammation of the parotid glands but swelling in other organs including the brain, heart and testicles has been observed.

MuV virions are pleomorphic and range in size from 100 to 600 nm (36). Within the viral particles is the viral genome which is encapsidated by the N protein to form a long helical RNP complex. The MuV genome is a negative sense, nonsegmented genome consisting of 15,384 nucleotides. The genome consists of seven continuously linked genes that encode nine proteins in the order: nucleocapsid protein, V protein/phosphoprotein (P)/I protein, matrix (M) protein, fusion (F) protein, small hydrophobic (SH) protein, hemagglutinin-neuraminidase (HN), and large (L) protein.

1.3.1 Mumps Viral Replication Cycle

Mumps virus infections enter host cells via receptor mediated endocytosis. Entry is mediated by attachment of HN proteins to sialic acid receptors on the cell surface. In addition to receptor binding to sialic acid, HN protein is also responsible for cleavage of sialic acid and fusion. Once the virus has been attached to the host cell, the F proteins initiate viral entry by fusion of the virion envelope and the host cell plasma membrane. The fusion process is driven by conformational changes in the F proteins into lower energy hairpin states at a neutral pH (36). Consequently, the negative sense nucleocapsid is released into the cytoplasm. Primary intracellular replication then occurs with the vRdRp transcribing the viral genome within the ribonucleoprotein (RNP) into 5' capped and 3' polyadenylated mRNAs. These viral mRNAs are then translated by host ribosomes to produce viral proteins for genome replication. The accumulation of high levels of N protein mediates the switch of the vRdRp to replication mode and synthesizes full length positive-sense antigenome which is subsequently used for synthesis of full-length negative-sense progeny genomes (37). These progeny genomes can then be utilized as a template for secondary transcription for production of additional antigenomes or assembled into virions. The M protein orchestrates assembly of the viral proteins and RNPs at the plasma membrane of the infected cell and the progeny virions are released via a budding process.

1.3.1.1 Nucleocapsid Protein

The vRdRp initiates transcription by binding to the NP encapsidated viral RNA to transcribe the genome into 5' capped and 3' polyadenylated mRNAs. The NP protein of MuV is the first transcribed gene in the viral genome and is 549 amino acid residues long. NP binds to the full-length negative-sense genome and positive-sense anti-genome to form the transcriptionally active template for replication. Formation of the helical RNP is suggested to

serve several purposes including protection from nucleases. Structural imaging of the RNP indicates that the NP protein binds to approximately six nucleotides per subunit and 13 NP subunits constitutes one turn of the helical RNP. Studies have shown that the overall genome length must be a multiple of six for efficient replication (38).

The structure of the nucleocapsid protein has been solved for several members of the *Paramyxoviridae* family. Analyses of these structures suggest that they share common features: an N-terminal domain and C-terminal domain in its core which is composed of mostly α -helices. The 3D structure of the NP protein of MuV has not been solved; however, recently the structure of the N-terminal core of NP of parainfluenza virus-5 (PIV) was solved to 3.11 Å (39).

Analyses suggest that the P protein tethers the polymerase complex onto the nucleocapsid template through interaction with the C-terminal domain of P (P_{CTD}). Furthermore, electron microscopy has revealed that the helical RNP is unwound by binding of the N-terminal domain of P (P_{NTD}) to the nucleocapsid protein (40). In infected cells, the nucleocapsid protein also exists in a soluble, monomeric form (N^0) that is not associated with RNA. These N^0 molecules have been found associated with the P protein for many paramyxoviruses including measles virus (MeV), sendai virus (SeV) and PIV5. The current data suggests that N^0 binds to nascent viral RNA and P binds to N^0 to prevent binding to non-specific cellular RNA.

1.3.2 Polymerase Complex

The vRdRp is a complex formed between the large (L) and P proteins. The L protein of mumps virus is a 2261 amino acid residues long which ranges in size from 160 to 200 kDa. Catalytic activity for RNA synthesis, polyadenylation, capping and methylation is included in the L protein. Analyses of the L proteins of paramyxoviruses indicate that the viruses share six conserved domains (I-VI) which are responsible for the multiple enzymatic capabilities of the

protein. However, the exact functions of each domain in the protein has not been extensively studied. Domain II has been suggested to be a possible RNA binding site due to its large net positive charge. Domain III is involved in phosphodiester bond formation. Domain V is thought to be responsible for the unique capping mechanism of the viral polymerase. Domain VI has been associated with methyltransferase activity. The function of Domains I, IV but mutagenic studies have suggested that these domains may play a role in the switch from replicase to transcriptase mode.

Self-oligomerization in the L protein and in complex with other viral proteins is also observed. The L-P protein complex is required for viral replication. The complex is necessary to tether the polymerase onto the nucleocapsid template. Furthermore, studies show that P-L interaction is necessary to stabilize the L protein; however, the actual domain of L which interacts with the P protein has not been determined.

The P protein is an important cofactor of the L protein in MuV. The P protein is transcribed from the V/P/I gene by addition of two guanine residues at the site 155. This edited mRNA is translated into a polypeptide which is 41 to 47 kDa in size. In addition to its functions in transcription and replication, the P protein binds to the nucleocapsid and tethers the polymerase onto the nucleocapsid template. Extensive analyses indicate that the P protein is a modular protein consisting of C- and N- terminal domains which are essential for transcription and nascent chain assembly respectively (41). The interaction between the N-terminal domain of the P protein and the N⁰ nascent chain leads to the formation of the complex that encapsidates RNA during replication.

The MuV P protein has not been extensively studied like that of other paramyxoviruses such as Sendai virus (SeV) and human respiratory syncytial virus (RSV). Although it is certain

that the P protein is essential for viral RNA synthesis, the exact mechanism of the P protein has not been elucidated.

1.3.3 Mumps Virus Prevention - MMR Vaccine

There are no FDA approved drugs or drugs that are being tested in clinical trials that target MuV. Vaccination is the only method of protection against viral infection. Currently two vaccines are approved to prevent MuV infection in the United States: M-M-R II and ProQuad. Historically, M-M-R II was the first approved vaccine in 1967 for prevention of mumps viral infection. M-M-R II is a trivalent combination live attenuated virus indicated for vaccination against MeV, MuV and Rubella which is administered in a two-dose regimen. M-M-R II contains sterile lyophilized preparations of ATTENUVAX (a line MeV derived from Enders' attenuated Edmonston strain), MUMPSVAX (Jeryl Lynn strain of MuV) and MERUVAX II (Wistart RA 27/3 strain of rubella virus). A single vaccine dose is 78% effective against the virus, whereas two doses are 88% effective. ProQuad vaccine was FDA approved in 2005 for prevention of measles, mumps, rubella and varicella in children 12 months through 12 years.

At the time of introduction of the M-M-R II vaccine, there were 152,209 reported cases of mumps virus and 25 deaths associated with infection. Since the implementation of vaccination measures, the incidence of mumps virus infections has steadily declined (**Figure 1-7**). Typically, once immunized according to the ACIP recommended schedule, natural immunity to the virus is lifelong.

Despite the efficacy of the vaccine large outbreaks of mumps virus has been recorded in the United States. In 2006 a mumps virus epidemic was observed in Iowa where a massive outbreak was noted on university campuses which ultimately spread to other states and resulted in 6584 cases and one death due to infection. Most cases were observed in patients with an

average age of 21 years of age of which 65% had a documented immunization history of receiving two doses of the M-M-R II vaccine. Since infection spread to patients with the recommended doses of vaccine, the ACIP recommended that a third dose of a mumps virus containing vaccine be used in populations at an increased risk for infection.

1.4 Scope of Research

Waning immunity to the M-M-R II vaccine and the lack of efficient therapeutics that target NSVs could become a public health crisis during epidemics and pandemics caused by these viruses. The polymerase complex of these viruses has not been extensively studied. The polymerase complex of NSVs is a potential target for drug design since viral infection requires the enzymatic activity of the polymerase and there is a low risk of off-target effects since the structural organization of NSV polymerases are significantly different from those in mammalian cells.

This body of work aims to investigate the biological properties of the vRdRp of NSVs for design of more efficient antivirals. The interactions between the viral proteins of the polymerase complex are necessary for viral infection. Thorough knowledge of the relationship between these proteins is necessary to develop inhibitors that target the protein complex. This research includes the following objectives:

1. To analyze the structure of the PB2cap binding protein for design of novel influenza therapeutics. The structure of the PB2cap binding protein was studied alone and in complex with a cap-analog. Several key residues were identified that would be critical for rational drug design.
2. Structural analysis of RNA sequestered in the nucleocapsid protein of MuV. The structure of the nucleocapsid protein has been solved for several members of the

Paramyxoviridae family. Cryo-electron microscopy was used to identify a region within the protein that is necessary for RNA release.

3. To characterize important residues in the amino terminal domain of the phosphoprotein which are necessary for unveiling the encapsidated RNA. To identify interactions between the nucleocapsid and phosphoproteins models of MuV nucleocapsid core and phosphoprotein several alanine mutants were generated and probed by binding assays. Furthermore, the effects of the variants on the thermal release of RNA was assessed.

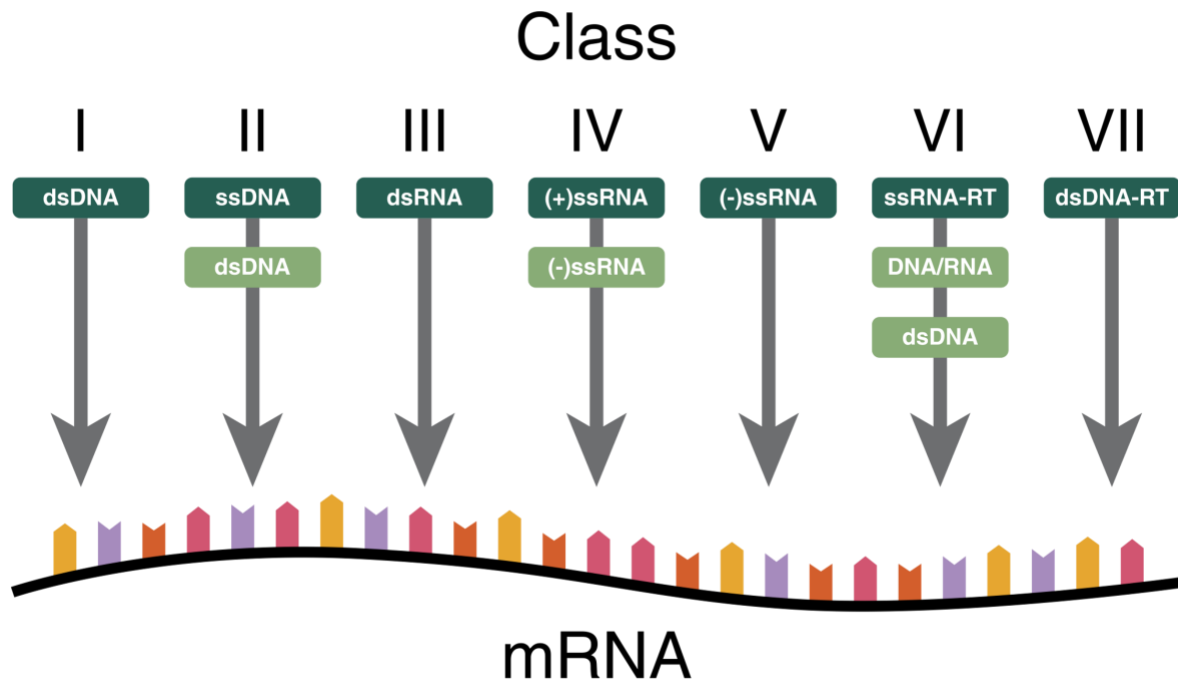


Figure 1-1 Baltimore Viral Classification

Schematic diagram of the Baltimore classification of viruses in which viruses are grouped by method of viral mRNA synthesis (Virus, Baltimore Classification by Thomas Spletstoeser licensed under CC BY-SA 3.0).

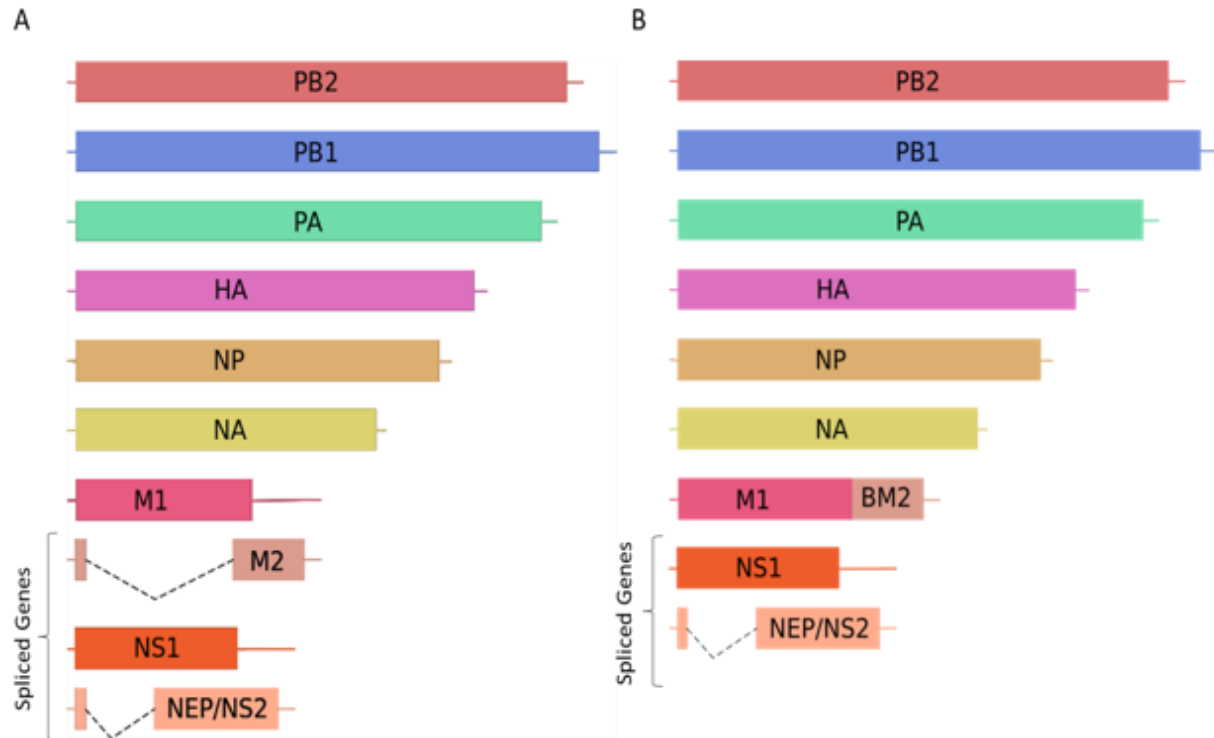


Figure 1-2 Genome Organization of Influenza Viruses

Schematic representation of the viral RNA segments of influenza A (A) and influenza B (B) viruses. Dashed lines represent introns.

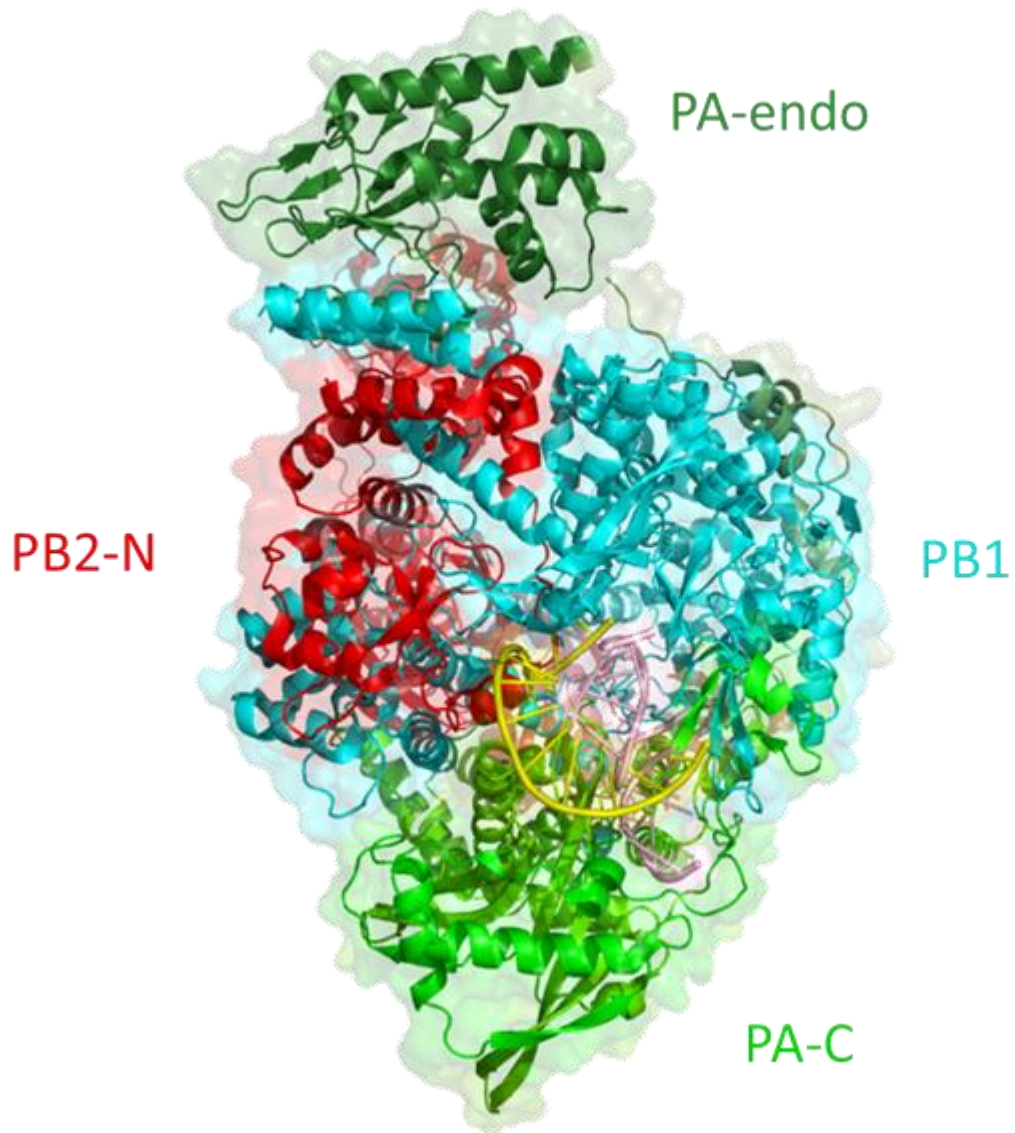
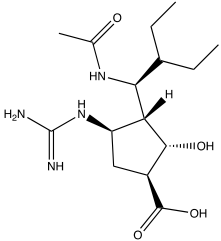
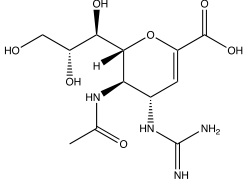
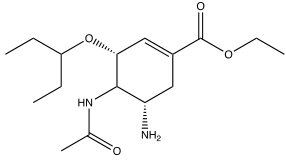
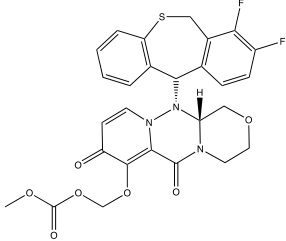
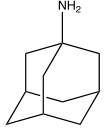
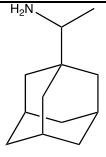


Figure 1-3 Cap Snatching Polymerase
Structure of bat influenza A polymerase in complex with vRNA promoter (PDB 4WSB)

Table 1-1 FDA Approved Influenza Drugs

	Structure	Activity	Drug Resistant Mutations
Neuraminidase inhibitors			
Peramivir (Rapivab®)		Types A and B	<ul style="list-style-type: none"> • H275Y mutation in NA confers resistance to influenza A/H1N1 • R292K, N294S, E119V in NA reduces viral susceptibility in influenza A/H3N2
Zanamivir (Relenza®)		Types A and B	<ul style="list-style-type: none"> • NA R292K mutant decreases efficacy against influenza A/H3N2 • D151G + H275Y substitutions in NA of pandemic 2009 H1N1 strains • S250G substitution in NA of influenza B viruses
Osetamivir (Tamiflu®)		Types A and B	<ul style="list-style-type: none"> • H275Y mutation in NA confers resistance to influenza A/H1N1 • R292K, N294S, E119V in NA reduces viral susceptibility in influenza A/H3N2
Polymerase inhibitors			
Baloxavir marboxil (Xofluza®)		Types A and B	<ul style="list-style-type: none"> • E23K, I38T substitutions in PA of A/H1N1, A/H3N2 and B viruses are associated with decreased efficacy of drugs
M2 ion channel inhibitors			
Amantadine		Type A	<ul style="list-style-type: none"> • S31N mutation in circulating viruses confers resistance; CDC does not recommend use for treatment of influenza viral infections
Rimantadine (Flumadine®)		Type A	<ul style="list-style-type: none"> • S31N mutation in circulating viruses confers resistance; CDC does not recommend use for treatment of influenza viral infections

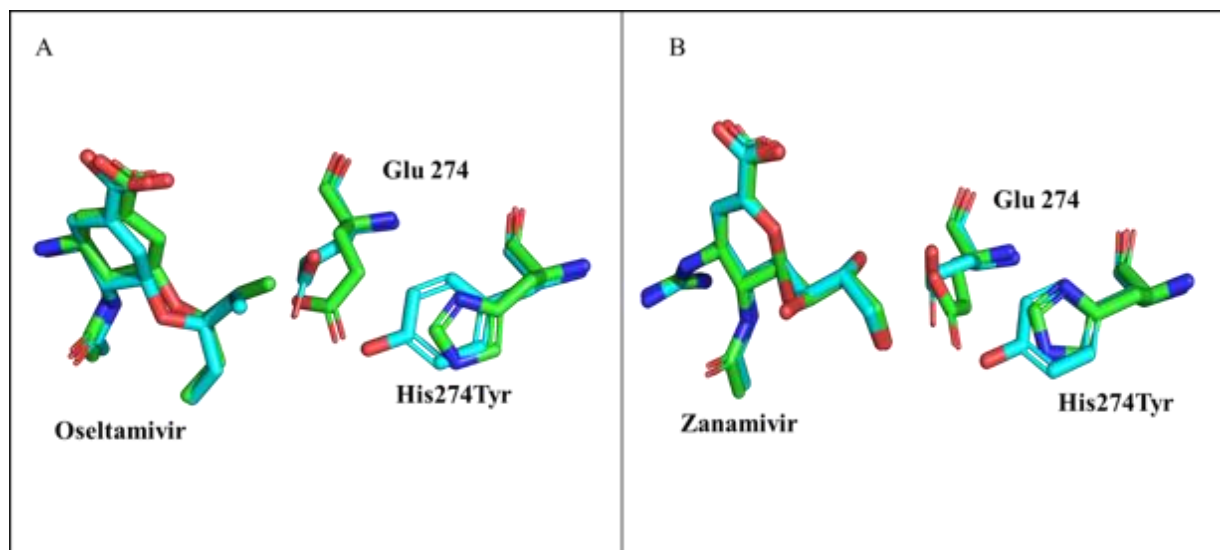
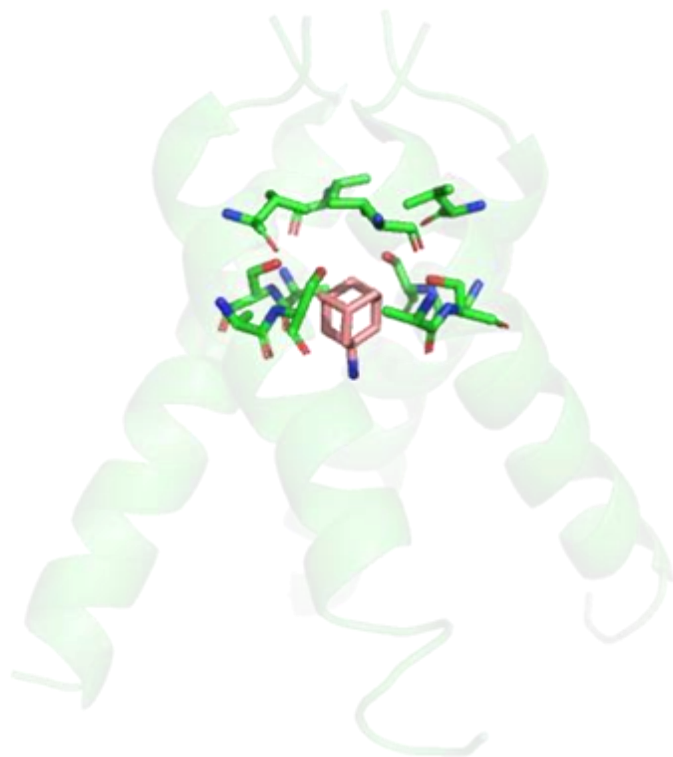


Figure 1-4 Drug Resistance in Neuraminidase

Superimposition of the active site of wild-type (green) and mutant (blue) bound to oseltamivir (A) and zanamivir (B). The H274Y structures indicate that the presence of the bulky tyrosine group pushes Glu 276 further into the binding site which disrupts the hydrophobic pocket.

A



B

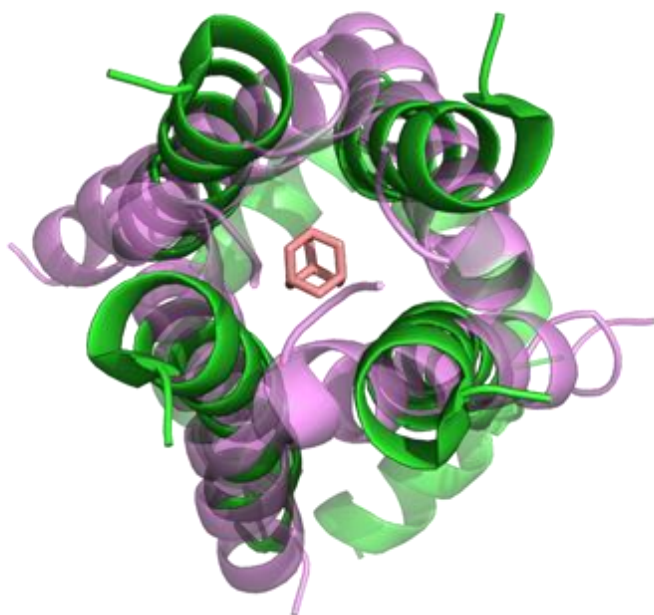


Figure 1-5 Drug Resistance in M2 channel

Amantadine binding site within the M2 ion channel of an influenza A virus (A). Superimposition of wild type (green) and mutant virus (purple) containing the S31N substitution indicates that the mutation opens the channel allowing protons to enter.

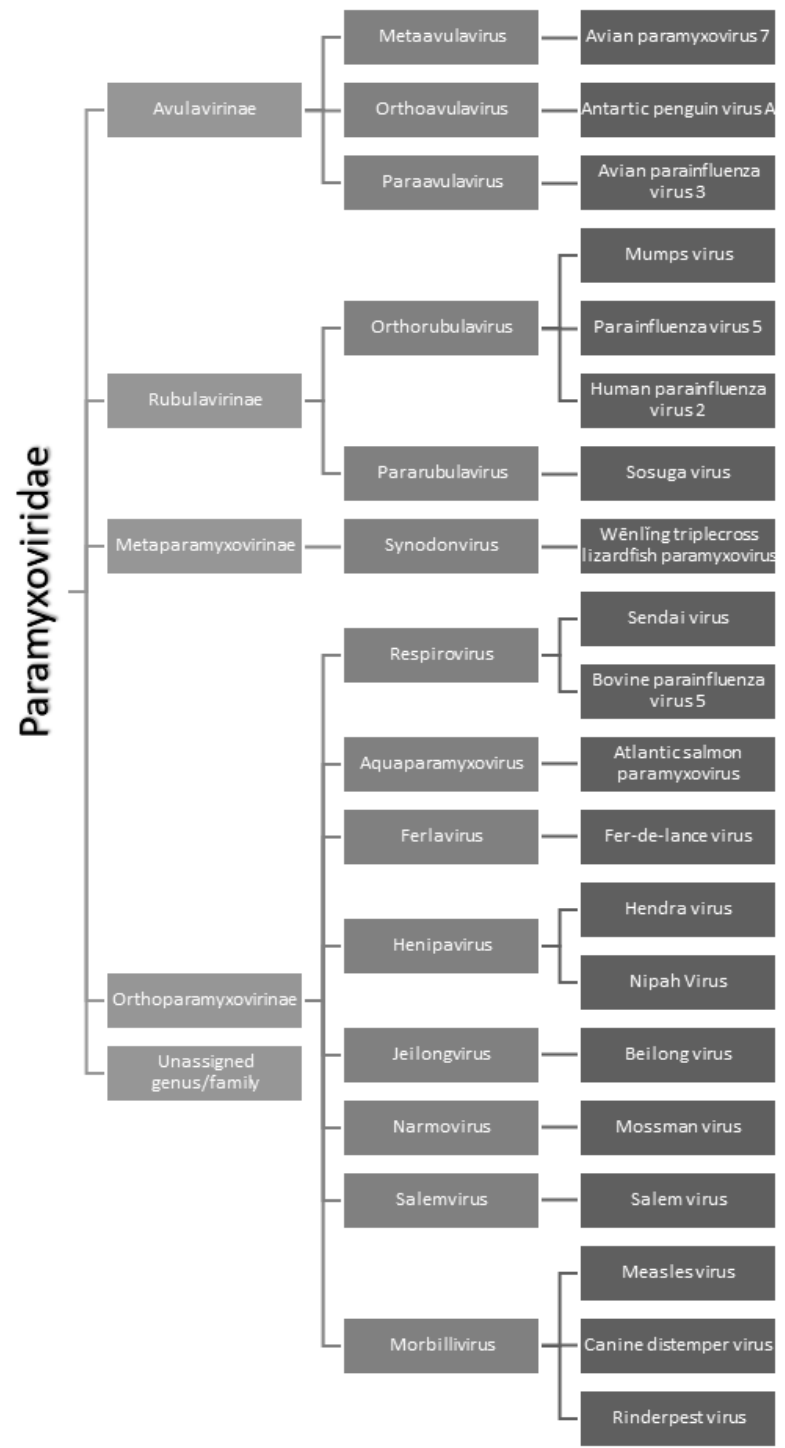


Figure 1-6 The Family Paramyxoviridae
The Paramyxoviridae family is subdivided into four subfamilies and 13 genera. Representative members of each genus are represented in the dark gray boxes.

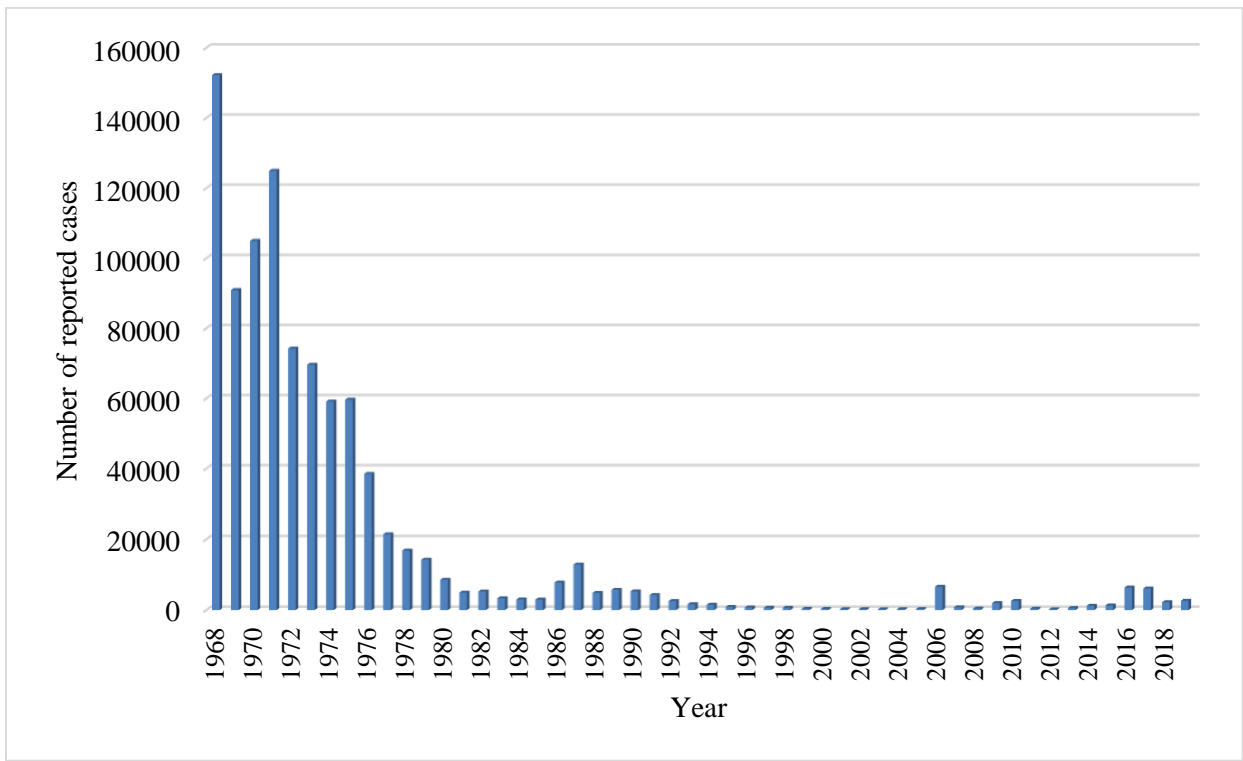


Figure 1-7 Mumps virus Outbreaks

Reported mumps virus outbreaks in the United States from 1968 to 2019. The mumps vaccine was introduced in 1968. Data as of October 11, 2019. Adapted from Morbidity and Mortality Weekly Report (MMWR), Notifiable Diseases and Mortality Tables

2. THE CAP-BINDING SITE OF INFLUENZA VIRUS PROTEIN PB2 AS A DRUG TARGET

Chelsea Severin, Tales Rocha de Moura, Yong Liu, Keqin Li, Xiaofeng Zheng and Ming Luo

Severin et al. *Acta Crystallographica D*. **2016**. DOI: 10.1107/S2059798316000085.

Copyright © 2016 International Union of Crystallography

2.1 Abstract

The RNA polymerase of influenza virus consists of three subunits: PA, PB1 and PB2. It uses a unique ‘cap-snatching’ mechanism for the transcription of viral mRNAs. The cap-binding domain of the PB2 subunit (PB2cap) in the viral polymerase binds the cap of a host pre-mRNA molecule, while the endonuclease of the PA subunit cleaves the RNA 10–13 nucleotides downstream from the cap. The capped RNA fragment is then used as the primer for viral mRNA transcription. The structure of PB2cap from influenza virus H1N1 A/California/ 07/2009 and of its complex with the cap analog m⁷GTP were solved at high resolution. Structural changes are observed in the cap-binding site of this new pandemic influenza virus strain, especially the hydrophobic interactions between the ligand and the target protein. m⁷GTP binds deeper in the

pocket than some other virus strains, much deeper than the host cap-binding proteins. Analysis of the new H1N1 structures and comparisons with other structures provide new insights into the design of small-molecule inhibitors that will be effective against multiple strains of both type A and type B influenza viruses.

2.2 Introduction

Influenza virus causes seasonal epidemics that affect the health of millions of people every year (41). Occasionally, a new pandemic strain emerges that can circulate the world in a short period of time, such as the pandemic strain of H1N1 influenza A virus in 2009 (pH1N1 2009; (42)). Since influenza virus mutates rapidly and new strains emerge frequently through the reassortment of viral strains originating from different hosts, new pandemic strains are expected. Control of influenza virus is therefore a major public health task. In addition to vaccines, antiviral drugs are a cost-effective means of controlling the spread of the virus. These drugs usually have a broad spectrum of activities against multiple influenza virus strains and can be stored on a relatively long term in strategic sites for rapid responses when new virus strains emerge. However, resistant strains of influenza virus have regularly been identified (43). Frequently prescribed anti-influenza drugs such as oseltamivir and zanamivir that are on the market today primarily target one viral protein: influenza virus neuraminidase. Furthermore, the Centers for Disease Control and Protection (CDC) report that high levels of resistance to the adamantanes (amantadine and rimantadine) persist among the influenza A viruses currently circulating. Adamantanes are ineffective against influenza B viruses. Consequently, it is desirable to have a panel of antiviral drugs that target other viral proteins.

Influenza virus has a unique mechanism for its viral transcription (44). The viral RNA-dependent RNA polymerase (vRdRp) of influenza virus has three virus encoded subunits: PA,

PB1 and PB2. The vRdRp does not synthesize the cap of viral mRNAs; instead, it uses a primer dependent mechanism for stealing a cap from the host pre-mRNA. The PB2 cap-binding domain (PB2cap) is responsible for binding the 5'-cap on host pre-mRNAs; the PA endonuclease then cleaves the snatched pre-mRNA 10–13 nucleotides downstream of the cap. The resulting oligonucleotide is used as a primer to initiate polymerization by the PB1 subunit. The crystal structure and functions of PB2cap have been carefully studied (19, 45, 46). The m⁷GTP binding site consists of a hydrophobic side chain (residue His357 in influenza A virus or Trp359 in influenza B virus) and a cluster of four aromatic residues (**Figure 2-1A**). The purine moiety is sandwiched between the two hydrophobic side chains (His357 and Phe404). In addition, two hydrogen bonds are formed by the purine moiety to the side chains of Glu361 and Lys376, respectively. The ribose and the triphosphate may interact with PB2cap as well, but their contribution to m⁷GTP binding appears to be less critical (45). It has been shown that the structure of PB2cap is different from the cap-binding domains of other proteins, including human proteins, and that the cap-binding site of PB2cap is conserved among different strains of influenza virus (21). For instance, the cap moiety is sandwiched between a His residue and a cluster of hydrophobic residues (Phe) in a deep pocket in PB2cap of influenza A virus. In the human cap-binding protein, however, it is sandwiched between two Tyr residues and its phosphate moiety is buried inside the binding site. PB2cap is therefore a viable target for broad-spectrum antiviral agents against influenza virus (47).

The structures of inhibitors in complex with PB2cap have previously been reported (47, 48). A series of m⁷GTP derivatives have been synthesized with modifications at the N-2, N-7 and N-9 positions of the guanine moiety (47). The compounds that showed good activities in blocking m⁷GTP binding mostly have a methyl group at the N-7 position, indicating a limited

space for accommodating large groups at this position. The main modification that yielded good inhibitory activities is at the N-9 position, where an aromatic moiety was added to replace the pyranose in the cap structure. This aromatic moiety shows hydrophobic interactions with the side chain of Phe323 (47). Recently, a new structure showed that a cyclohexyl carboxylic acid analogue (named VX-787) binds in the cap-binding site with the azaindole moiety replacing the 7-methylguanidine (**Figure 2-1B**). The azaindole moiety recapitulates both the hydrophobic interactions and hydrogen bonds of the purine moiety in the cap. In addition, the cyclohexyl carboxylic group is coupled to azaindole via fluoropyrimidine. The cyclohexyl group provides more hydrophobic interactions with the cluster of four aromatic residues. The carboxyl group may form hydrogen bonds to two ordered water molecules. VX-787 reportedly has a strong potency to inhibit a number of influenza A virus strains (48), but showed negligible activity against influenza B virus (49). To explore the potential of designing a potent inhibitor that is effective against multiple strains of both type A and type B influenza viruses, we solved the crystal structures of PB2cap from pH1N1 A/California/07/2009 and its complex with m⁷GTP at resolutions of 1.54 and 1.40 Å, respectively. By comparing with different structures, we show that the cap-binding site has a measurable flexibility to accommodate a compound that fits in the hydrophobic pocket. The hydrogen bonds to the side chains of Glu361 and Lys376 should be established because the inhibitor needs to displace a few ordered water molecules that interact with these side chains. It is also necessary to establish a hydrogen bond to the side chain of Arg334 in influenza B virus (Arg332 in influenza A virus) if the same inhibitor is to also bind influenza B virus PB2cap (50).

2.3 Materials and methods

2.3.1 Protein expression and purification

The DNA coding sequence for PB2cap from A/California/07/2009 H1N1 (CA09-PB2cap) was inserted into pET-28a vector between NdeI and XhoI sites to generate a His₆-tag fusion protein with a thrombin cleavage site. Protein expression was carried out in *Escherichia coli* BL21 (DE3) cells and expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 18°C for 18 h. The recombinant protein was purified using a nickel-affinity column (Ni²⁺-charged HiTrap chelating HP column from GE Healthcare). The His₆ tag was removed by incubation with thrombin protease for 4 h at 25°C and the sample was applied onto an Ni²⁺-charged HiTrap column again to remove uncleaved protein. The protein was further purified by gel-filtration chromatography using a Superdex 75 column (GE Healthcare) previously equilibrated with 10 mM Tris pH 8.0, 200 mM NaCl. The fractions corresponding to the PB2cap protein were pooled and concentrated to 4–10 mg ml⁻¹ for crystallization. The expression vector for the PB2cap truncation mutant was generated using a QuikChange kit (Agilent Technologies, California, USA) and the protein was purified following the same procedure as used for native PB2cap.

2.3.2 Crystallization and structure analyses

The native and mutant PB2cap proteins were subjected to crystal screens (Index, Matrix, PEG/Ion, Crystal Screen and Crystal Screen 2; Hampton Research, California, USA). After optimization, crystals were grown by vapor diffusion with a hanging drop consisting of 10 mg ml⁻¹ native PB2cap protein mixed with a reservoir solution composed of 0.2 M magnesium nitrate 0.1 M HEPES pH 8.0, 20%(w/v) PEG 3350 and a hanging drop consisting of 4 mg ml⁻¹ mutant PB2cap protein mixed with a reservoir solution composed of 0.1 M bis-tris pH 6.5,

15% (w/v) PEG 3350 at 20°C. Cocrystallization was carried out by adding 0.2 mM m⁷GTP to the PB2cap solution and concentrating it to a final protein concentration of 4 mg ml⁻¹. The cocrystals were grown under the same condition as used for the protein without m⁷GTP.

Protein crystals were transferred to a solution of mother liquor containing 20% glycerol and were flash-cooled in liquid nitrogen. X-ray diffraction data were collected at Shanghai Synchrotron Radiation Facility, China and SER-CAT at the Advanced Photon Source, USA. Data processing was carried out with the HKL-2000 program suite. Molecular replacement was performed with MOLREP using the coordinates of PDB entry 4enf (45) as the search model. The structure of the mutant PB2cap was solved with the native structure following the same protocol. Structural refinement was carried out with REFMAC5. MOLREP and REFMAC5 are part of the CCP4 crystallographic package (51). X-ray crystallographic analyses are summarized in Table 1. The coordinates of the reported structures were downloaded from the RCSB PDB. Structure superposition and figure preparation were carried out using PyMOL (v.1.3; Schrödinger).

2.4 Results and discussion

2.4.1 *The cap-binding site in PB2 of A/California/07/2009*

The deeper end of the cap-binding site contains two critical residues: Glu361 and Lys376. The side chains of these two residues form hydrogen bonds to the guanine moiety when the cap binds. The side chain of Glu361 has a stable conformation, whereas that of Lys376 may assume two different conformations (**Figure 2-2A**), one of which is suitable for forming the hydrogen bond to the guanine moiety. There are also two important residues with aromatic side chains: His357 and Phe404. These side chains sandwich the guanine moiety by π - π stacking interactions. The side chain of His357 appears to rotate when the cap-binding site is empty (**Figure 2-2B**). The orientation of the His357 side chain is different in A/California/07/2009 (H1N1) PB2cap

compared with that in A/PR/8/34 (H1N1) PB2cap (**Figure 2-2B**). Next, the side chain of Phe323 makes hydrophobic interactions with the pyranose in the cap. Moreover, the side chain of Lys339 is in a position to form a hydrogen bond to the 2'-OH of the pyranose, but it rotates away when the cap-binding site is empty. Further out, the side chain of His432 may be involved in charge interactions with the α -phosphate group in the cap, but its orientation is quite flexible. The side chain of Asn429 is also in position to form a hydrogen bond to the α -phosphate group, but it shows two orientations when the cap-binding site is empty.

There is an ordered water molecule that forms a hydrogen bond to the carbonyl group of Phe404 (**Figure 2-2A**). This water molecule is displaced when the cap binds. A thiocyanate group from the crystallization solution is modeled in a pocket lined by the side chains of Gln406 and Met431. In the structure of A/PR/8/34 PB2cap, a nitrate group from the crystallization buffer is modeled in a nearby position. These observations suggest that there is a large space at this location.

The most flexible regions are located in the large loops (**Figure 2-2C**). However, the main-chain conformation near the cap-binding site also showed some flexibility. The major differences are in the region from Thr333 to Lys339 and the region from Ala413 to Lys440 (**Figure 2-2D**), Thr333 is at the end of a β -strand and the polypeptide makes a sharp turn into the next β -strand ending at Lys340. In A/PR/8/34 PB2cap, this region opens more widely so that the side chain of Lys339 is no longer in a position to interact with the cap. Ala413 is at the start of an α -helix, followed by a large flexible loop that ends at Lys440. This loop has a large influence in crystallization. In the native PB2cap of A/California/07/2009 the side chain of Arg423 in the loop occupies the cap-binding site, bridging strong interactions between two molecules in the asymmetric unit. Attempts to soak m⁷GTP into the crystal or to cocrystallize with PB2cap of

A/California/07/2009 both failed. Since the Arg432 loop blocks the cap-binding site, we therefore replaced residues Val421–Arg427 with a GSG linker to cut the loop short. The shortened loop resulted in the loss of one turn in the α -helix. On the other hand, this helix has one extra turn in the PB2cap of A/PR/8/34, leading to a different conformation of the following loop. Overall, the cap-binding site in PB2cap showed considerable flexibility observed as alternative side-chain orientations of the active-site residues and an alternative conformation of the secondary-structure elements near the cap-binding site. These factors should be taken into account when an inhibitor is designed to tightly bind the active site in PB2cap of all influenza virus strains.

2.4.2 Interactions between m^7 GTP and the cap-binding site

As mentioned above, efforts to soak m^7 GTP into the crystal of the native PB2cap of A/California/07/2009 were unsuccessful. When cocrystallization was attempted, the crystals obtained did not contain m^7 GTP. We therefore carried out mutagenesis to shorten the loop that occupies the cap-binding site of the neighboring molecule in the asymmetric unit. The mutant PB2cap was subjected to crystal screens, and protein crystals were grown using 0.1M bis-tris pH 6.5, 15%(w/v) PEG 3350 as the reservoir solution. However, the crystals dissolved when soaked in mother liquor containing 1mM m^7 GTP, indicating that a large conformational change was induced by cap binding. Cocrystals of m^7 GTP were grown with the mutant PB2cap using the same reservoir solution. Comparisons between the native and the complex structures revealed the conformational changes when m^7 GTP binds to the cap-binding site (**Figure 2-3A**). All structures used in the comparisons have been determined at 1.5 Å resolution or better. The truncated mutant PB2cap has essentially the same structure as the native PB2cap (r.m.s.d. of 0.379 Å for 905 aligned atoms), except that the side chain of His357 has a different orientation, indicating a

very flexible conformation for this side chain. The truncated mutant PB2cap was not included in the comparisons.

In the complex of m⁷GTP with the mutant PB2cap, it appears that the main chain around Glu361 has moved away by about 0.6 Å to allow the formation of proper hydrogen bonds between the side chain of Glu361 and the guanine moiety (**Figure 2-3A**). The distances of the Glu361 carboxyl O atoms to N1 and NH2 at position 2 of the guanine moiety are now 2.7 and 2.9 Å, respectively. At the same time, the side chain of Lys376 assumed one conformation instead of the two possible conformations found in the native PB2cap in order to form a proper hydrogen bond to the carbonyl O atom at position 6. The ordered water molecule at the carbonyl group of Phe404 was clearly displaced. More visibly, the side chain of His357 was rotated so that it could make the most π - π stacking with the guanine moiety (**Figure 2-3A**). The side chain of Phe404 is usually identified as contributing to π - π stacking on the other side of the guanine moiety (21). However, the side chain of Phe323 is also in an appropriate position to form π - π stacking with the guanine moiety. Moreover, the side chain of Phe363 is within a suitable distance to make a T-shaped interaction with the guanine moiety. It is likely that the three aromatic side chains jointly make strong hydrophobic interactions with the guanine moiety. The methyl group at position 7 is located in a hydrophobic pocket formed by Phe404, Met431 and part of the Gln406 side chain. The side chain of Met431 actually retracted so that the distance between the two methyl groups changes from 3.0 Å (closest distance in van der Waals interactions) to 4.5 Å.

The 2'-OH group forms hydrogen bonds to the side chains of Lys339 and His357 in PB2cap of A/California/07/2009. The pyranose ring is roughly orthogonal to the guanine moiety. This ring may have some degree of hydrophobic interaction with the side chain of Phe323, but

their orientation is not parallel and the distance is long. It may not make a significant contribution to cap binding. The α -phosphate group clearly makes salt-bridge and hydrogen-bond interactions with the side chains of His432 and Asn429. His432 assumes two conformations in the interactions, one of which overlaps with one of the two conformations in the native PB2cap. The Asn429 side chain assumes only one conformation that overlaps with one of the two conformations in the native PB2cap. The β -phosphate group could make a salt bridge with the side chain of Lys339, which moved from a distance of 2.2 to 2.7 Å. However, the conformation of the β -phosphate group as well as that of the γ -phosphate group change dramatically in different PB2cap structures. How the phosphate groups interact with PB2cap may only become clearer when the cap is placed in the context of the full-length viral polymerase complex. It is clear, however, that a positively charged residue, Lys or Arg, is conserved at position 339. Its side chain is highly likely to form a salt bridge to one of the phosphate groups in the cap.

The cap binding is also compared with PB2cap of A/Duck/Shantou/4610/2003 (H5N1; Pautus et al., 2013; **Figure 2-3B**) and A/Hong Kong/1/68 (H3N2; Liu et al., 2013; **Figure 2-3C**). In PB2cap of H5N1, m⁷GTP is moved outwards from the cap-binding site by 0.5 Å. Accordingly, the side chain of His357 also moves to maintain the π - π stacking, accompanied by movement of the main chain. The pyranose ring is rotated by about 22.6° so that it has less inclination relative to the guanine moiety. As a result, the 2'-OH group no longer forms a hydrogen bond to the side chain of His357. The side chain of Lys339, on the other hand, still forms a hydrogen bond to the 2'-OH group, but it has to move towards the OH group accompanied by movement of the main chain. The conformation of all three phosphate groups is different in the three structures, and no consensus interactions may be discerned. In PB2cap of

H3N2 all interactions are similar to those found in A/California/07/2009, except that the side chain of Phe323 is shifted owing to movement of the main chain. Since Phe323 is close to the terminus of PB2cap, the crystal contacts may have an effect on its side-chain conformation.

2.4.3 Comparisons with PB2cap inhibitors

The structures of two classes of inhibitors in complex with PB2cap are available (**Figure 2-4**). The complex of a representative compound (8f) with PB2cap of H5N1 A/duck/Shantou/4610/2003 PB2cap is used for comparison with the first class of inhibitors. Compound 8f is a derivative of m⁷-guanine and showed modest inhibition of cap binding to PB2cap (47). The structure shows that the guanine moiety of compound 8f retains essentially the same interactions as those in m⁷GTP (**Figure 2-4A**). The phenyl group coupled to the guanine moiety should make some hydrophobic interactions with the side chain of Phe303, but is not close enough for π - π stacking. The hydroxyl group on the phenyl ring may form a hydrogen bond to the main chain via a water molecule. This compound has only limited interactions with PB2cap, consistent with its modest inhibitory activity.

On the other hand, the cyclohexyl carboxylic acid analogue VX-787 is a potent inhibitor (48). In its complex with PB2cap of A/Victoria/3/1975 H3N2 (**Figure 2-4B**), the N atom at position 4 of the azaindole moiety forms a hydrogen bond to the side chain of Glu361 and the N atom at position 9 forms a hydrogen bond to the side chain of Lys376. More importantly, the azaindole pyrimidinyl moiety is in a perfect position to make π - π stacking interactions with the side chains of Phe323 and Phe404 simultaneously. Furthermore, the bicyclooctane coupled to the pyrimidinyl moiety perfectly occupies the hydrophobic pocket formed by the side chains of Phe325 and Met431. Finally, the carboxyl group linked to the bicyclooctane forms a salt bridge with the side chain of Arg355 and hydrogen bonds to the side chains of His357 and

Gln406 via water molecules. The side chain of Lys339 (Arg339) is close by. Its side chain could also form a salt bridge with the carboxyl group in different influenza A virus strains. VX-787 capitalizes on all available interactions with PB2cap and showed potent inhibitory activities against a number of influenza A virus strains (48).

2.4.4 Differences from PB2cap of influenza B virus

VX-787 is not active against influenza B virus (49). Since influenza B virus is also a serious health threat, it is more desirable to have an inhibitor that is effective against both influenza A and B viruses. Comparison of the VX-787 complex structure with the GDP complex of influenza B PB2cap (50) may suggest why VX-787 does not effectively inhibit influenza B viruses (**Figure 2-5A**). The guanine moiety in PB2cap of influenza B virus B/Jiangxi/BV/2006 is rotated 180° relative to that in PB2cap of influenza A virus (**Figure 2-5B**). As a result, the N atom at position 1 and the amino group linked to position 2 still form hydrogen bonds to the side chain of Glu361, but in a reversed orientation. The carbonyl at position 6 now forms a hydrogen bond to the side chain of Arg332 (Arg334 in influenza B virus) instead of that of Lys376. His357 is replaced by Trp, which makes more π - π stacking interactions with the guanine moiety. However, π - π stacking interactions on the other side are only retained for the side chain of Phe404. Phe323 is replaced by Gln, which can no longer provide π - π stacking. Moreover, His432 is replaced by Tyr, which has a bulkier side chain that provides steric hindrance to bicyclooctane. It seems that many of the favorable interactions between VX-787 and PB2cap of influenza A virus do not exist in PB2cap of influenza B virus, and more importantly some residues in PB2cap of influenza B virus could potentially pose steric hindrance to VX-787, such as Tyr432. The characteristics of both cap-binding sites need to be considered simultaneously if one inhibitor is to be designed against both influenza A and B viruses.

2.4.5 Comparisons with human cap-binding proteins

Binding of an inhibitor to host cap-binding proteins needs to be avoided in order to design specific inhibitors of PB2cap. The structures of a cap-specific mRNA methyltransferase (52) and the translation initiation factor eIF4E (53) were used as examples (**Figure 2-6**). In the methyltransferase, the carbonyl O atom in the guanine moiety forms a hydrogen bond to the main-chain amide N atom of Trp102 (**Figure 2-6A**). The side chain of Glu103 forms two hydrogen bonds to the N atom at position 1 and the amino group linked to position 2, respectively. There are π - π stacking interactions made by Trp56 and Trp102 on each side of the guanine moiety. The methyl group linked to position 7 does not seem to be in a pocket but is near the side chain of Trp166. There seems to be no direct interactions between the protein and the pyranose ring. On the other hand, strong interactions are present between the α - and β -phosphate groups and the side chain of Arg157, and between the β - and γ -phosphate groups and the side chain of Lys162.

In eIF4E, the side chain of Asp207 forms a hydrogen bond to the N atom at position 1 and that of Asn374 forms a hydrogen bond to the amino group linked to position 2 (**Figure 2-6B**). There are no aromatic π - π stacking interactions with the guanine moiety, but the side chain of Glu373 shows an anion-aromatic stacking interaction with the guanine moiety. Again, there is no direction interaction with the pyranose ring. Charge or hydrogen-bond interactions are present between the β -phosphate group and the side chain of Asn439 and between the α - and γ -phosphate groups and the side chain of Arg218.

Overall, the interactions between the cap and the host proteins include those with the guanine and preferentially the phosphate groups. The binding site is not as deep as that in

PB2cap. There is no interaction with the pyranose ring, which seems to serve only as a linker between the guanine moiety and the triphosphate.

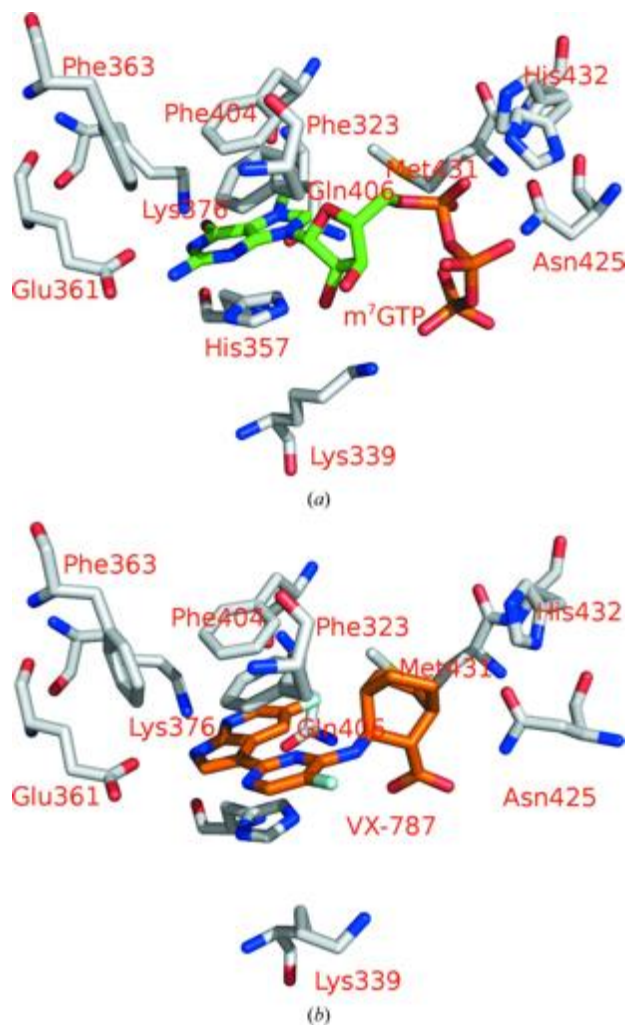


Figure 2-1 The cap-binding site in PB2cap

(a) A/California/07/2009 PB2cap bound to m⁷GTP (PDB entry 5eg7), (b) A/Victoria/3/1975 PB2cap bound to VX-787 (PDB entry 4p1u ; Clark et al., 2014). The residues are labeled according to the numbering of A/California/07/2009 PB2cap in this and the following figures. Each ligand in the PB2cap binding site is also labeled.

Table 2-1 Data-collection and refinement statistics

Values in parentheses are for the highest resolution shell

Protein	CA09-PB2cap	m ⁷ GTP-mutant CA09-PB2cap	Mutant CA09- PB2cap
PDB code	5eg8	5eg7	5eg9
Data Collection			
Space group	P3 ₁ 21	P4 ₃ 2 ₁ 2	P12 ₁ 1
Unit-cell parameters			
<i>a</i> (Å)	54.61	45.26	37.13
<i>b</i> (Å)	54.61	45.26	109.01
<i>c</i> (Å)	196.49	156.40	39.66
α (°)	90	90	90
β (°)	90	90	90.54
γ (°)	120	90	90
Molecules in asymmetric unit	2	1	2
Wavelength (Å)	0.9795	1.0000	1.0000
Resolution (Å)	15-1.54 (1.59- 1.54)	25-1.40 (1.42-1.40)	50-2.30(2.34- 2.30)
R _{merge} [†] (%)	7.0 (21.4)	7.1 (61.1)	7.1(19.4)
<i>I</i> /σ(<i>I</i>)	57.9 (14.2)	33.2 (1.8)	21.2(6.2)
CC _{1/2}	0.997 (0.992)	0.912 (0.581)	0.893 (0.906)
Completeness (%)	97.2 (96.9)	99.8 (97.0)	87.7(92.6)
Multiplicity	15.1 (16.7)	10.4 (4.0)	1.8(1.7)
Refinement			
Resolution (Å)	14.63-1.54	25-1.40	19.83-2.30
No. reflections	47547	31406	11663
R _{work} /R _{free} [‡] (%)	12.2/17.7	12.4/16.9	19.0/25.6
Average <i>B</i> factor (Å ²)	24.9	16.1	29.3
R.m.s. deviations			
Bond lengths (Å)	0.019	0.020	0.014
Bond angles (°)	1.93	2.35	1.65
Ramachandran plot (%)			
Favored region	97.5	98.1	96.8
Allowed region	2.2	1.9	2.6
Outlier region	0.3	-	0.6

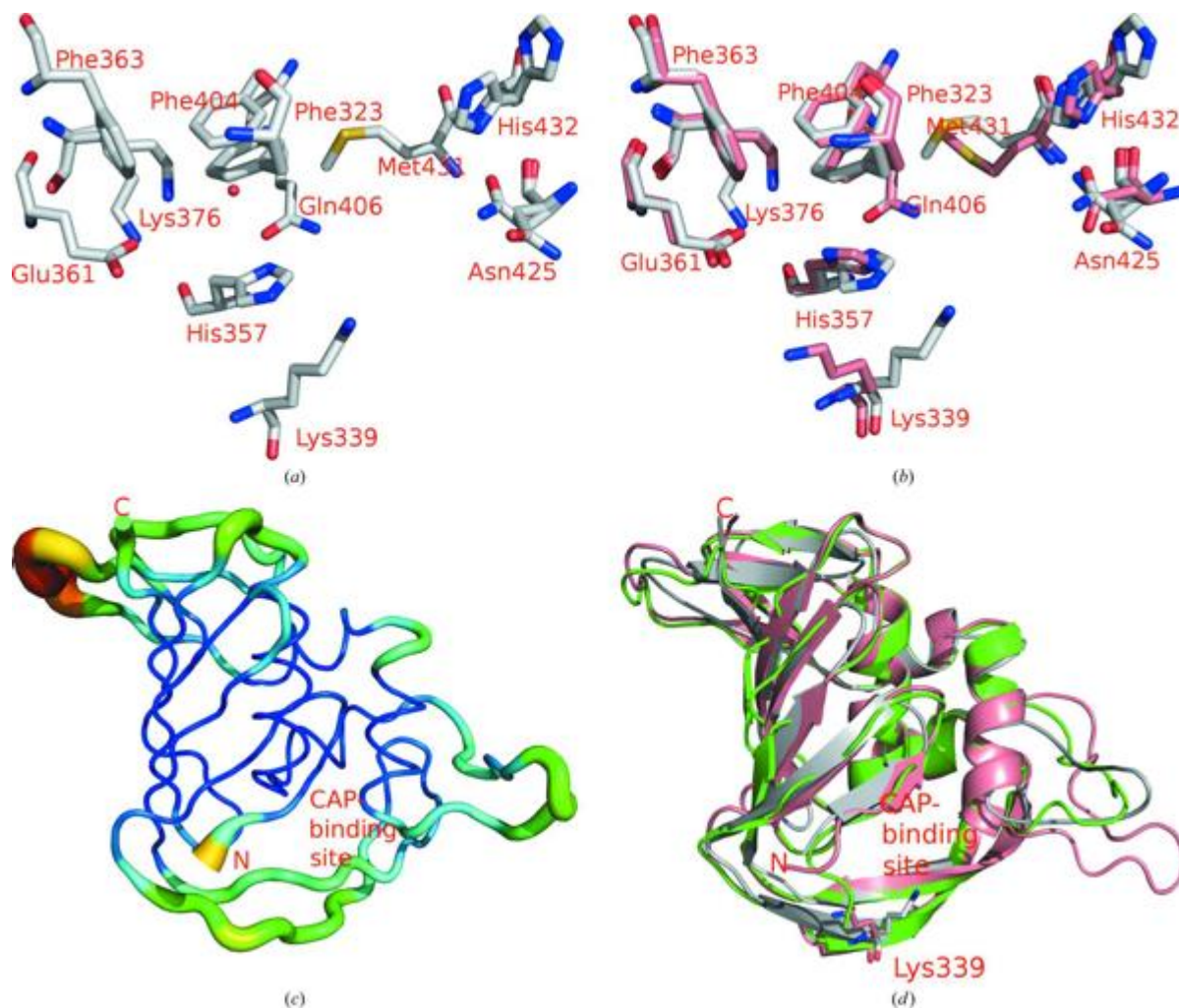


Figure 2-2 Structure of Influenza A cap binding site

(a) The cap-binding site in PB2cap of A/California/07/2009 (PDB entry 5eg8); 11 residues (labeled) and one water molecule (sphere) are shown. Lys376, Asn429 and His432 are represented by two alternative side-chain conformations. (b) Comparisons of the cap-binding site in A/California/07/2009 PB2cap (gray) and A/PR/8/34 PB2cap (PDB entry 4enf ; pink; Liu et al., 2013). Residues Lys376, Asn429 and His432 in A/PR/8/34 PB2cap do not show alternative conformations. (c) A main-chain diagram of A/California/07/2009 PB2cap; the B factor is used to puff the trace (PyMOL). The N- and C-termini are labeled N and C, respectively. The location of the cap-binding site is indicated. (d) Comparison of three PB2cap structures: native (gray), mutant (green) and A/PR/8/34 (pink); residue Lys339 is labeled. The N- and C-termini are labeled N and C, respectively. The location of the cap-binding site is indicated.

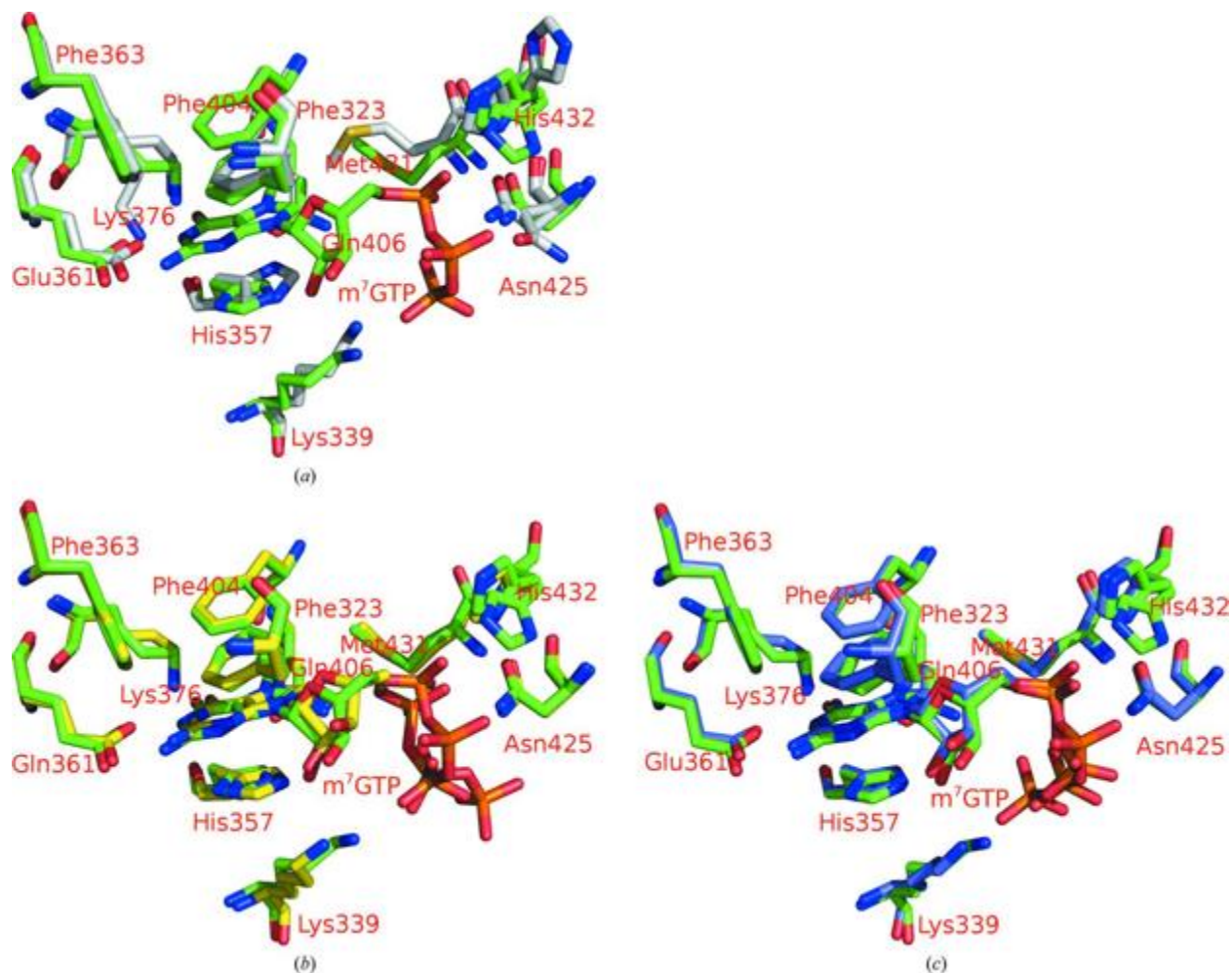


Figure 2-3 Comparing bound and unbound PB2cap

(a) Comparison of the native (gray) and m⁷GTP-bound (green) cap-binding site in A/California/07/2009 PB2cap. Lys376 and Asn425 showed only one conformation in the complex, whereas His357, Met431 and His432 have alternative conformations in the complex.

(b) Comparison of the cap-binding site with m⁷GTP bound in A/California/07/2009 PB2cap (green) and A/Duck/Shantou/4610/2003 (H5N1) PB2cap (PDB entry 4cb4 ; yellow; Pautus et al., 2013). These two structures have the closest superposition.

(c) Comparison of the cap-binding site in m⁷GTP-bound A/California/07/2009 PB2cap (green) and A/Hong Kong/1/68 PB2cap (H3N2; PDB entry 4eqk ; blue; Liu et al., 2013). The two structures are also superimpose well.

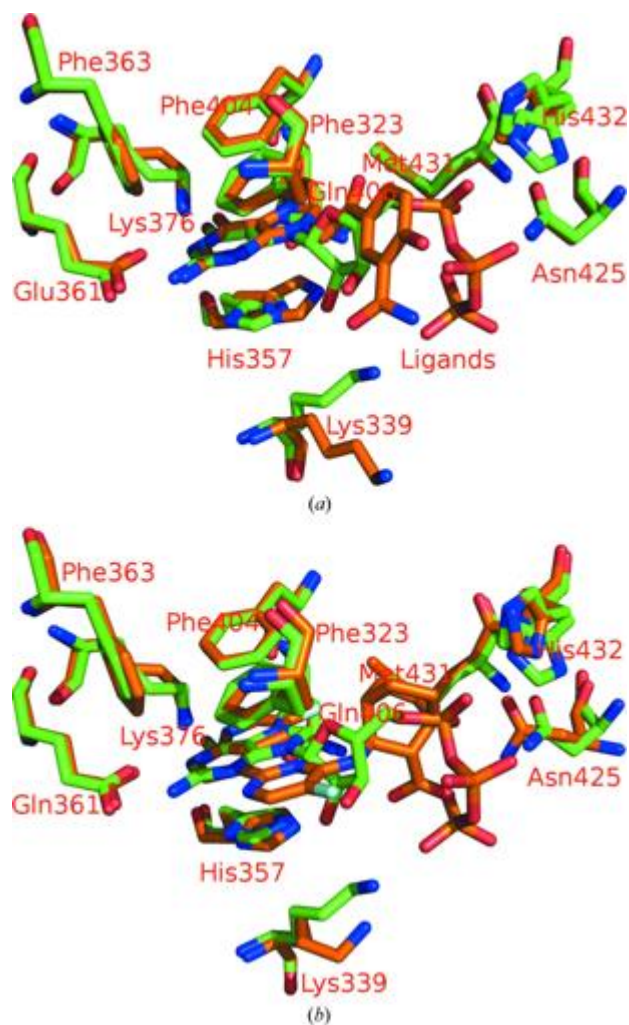


Figure 2-4 Comparison of the PB2cap cap-binding site.

(a) A/California/07/2009 PB2cap bound to m7GTP (green) and A/Duck/Shantou/4610/2003 PB2cap bound to compound 8f (PDB entry 4cb5 ; orange; Pautus et al., 2013). `Ligands' indicate the locations of m7GTP and 8f. (b) A/California/07/2009 PB2cap bound to m7GTP (green) and A/Victoria/3/1975 PB2cap bound to VX-787 (PDB entry 4p1u ; orange; Clark et al., 2014).

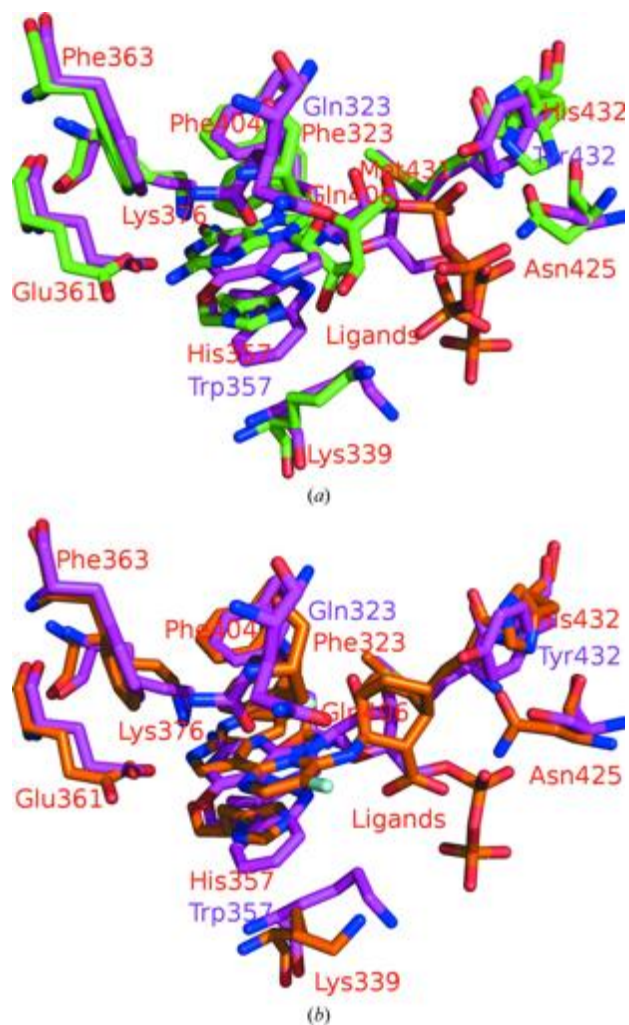


Figure 2-5 Comparing Influenza A and B cap binding

(a) Comparison of the GDP-bound cap-binding site in *B/Jiangxi/BV/2006 PB2cap* (PDB entry 4q46 ; magenta; Liu et al., 2015) and the m7GTP-bound cap-binding site in *A/California/07/2009 PB2cap* (green). Residues in *B/Jiangxi/BV/2006 PB2cap* are labeled in purple if they are different from those in *A/California/07/2009 PB2cap*. 'Ligands' indicates the locations of GDP and m7GTP, respectively. (b) Comparison of the GDP-bound cap-binding site in *B/Jiangxi/BV/2006 PB2cap* (PDB entry 4q46; magenta) and the VX-787-bound cap-binding site in *A/Victoria/3/1975 PB2cap* (PDB entry 4p1u; orange; Clark et al., 2014). Residues in *B/Jiangxi/BV/2006 PB2cap* are labeled in purple if they are different from those in *A/California/07/2009 PB2cap*. 'Ligands' indicates the locations of GDP and VX-787, respectively.

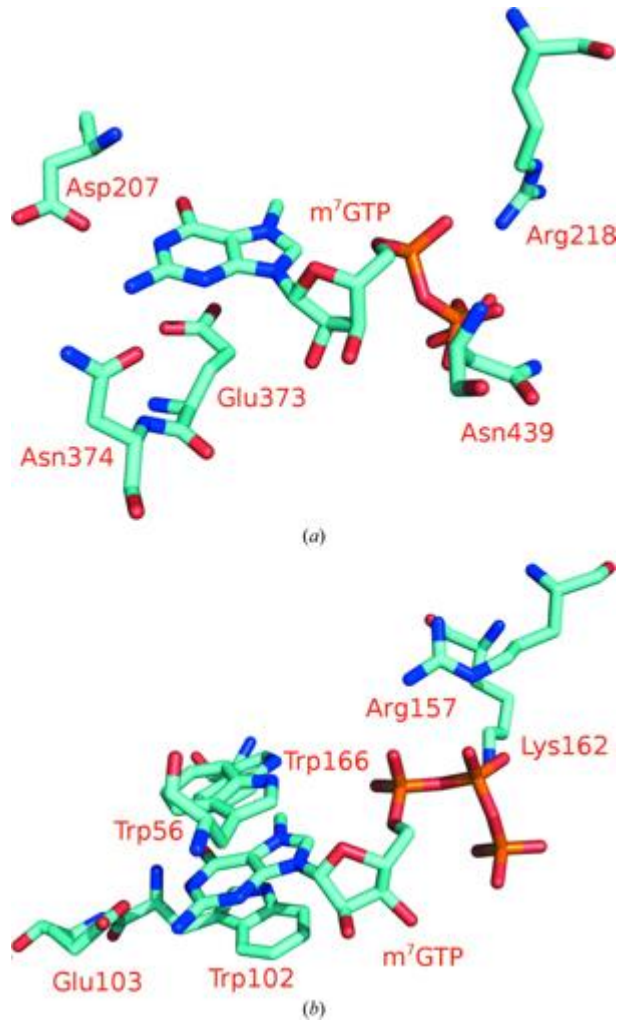


Figure 2-6 Human cap binding proteins

The cap-binding site (a) in the cap-specific mRNA methyltransferase bound to m⁷GTP (PDB entry 4n49 ; Smietanski et al., 2014) and (b) in the translation initiation factor eIF4E bound to m⁷GTP (PDB entry 4tqb ; Papodopolous et al., 2014).

3. RELEASING THE GENOMIC RNA SEQUESTERED IN THE MUMPS VIRUS NUCLEOCAPSID

Chelsea Severin, James R. Terrell, James R. Zengel, Robert Cox, Richard K. Plemper, Biao He,
Ming Luo

Severin et al. *Journal of Virology*. **2016**. DOI: 10.1128/JVI.01422-16

Copyright © 2016 American Society for Microbiology

3.1 Abstract

In a negative-strand RNA virus, the genomic RNA is sequestered inside the nucleocapsid when the viral RNA-dependent RNA polymerase uses it as the template for viral RNA synthesis. It must require a conformational change in the nucleocapsid protein (N) to make the RNA accessible to the viral polymerase during this process. The structure of an empty mumps virus (MuV) nucleocapsid-like particle was determined to 10.4-Å resolution by cryo-electron microscopy (cryo-EM) image reconstruction. By modeling the crystal structure of parainfluenza virus 5 into the density, it was shown that the α -helix close to the RNA became flexible when RNA was removed. Point mutations in this helix resulted in loss of polymerase activities. Since the core of N is rigid in the nucleocapsid, we suggest that interactions between this region of the

mumps virus N and its polymerase, instead of large N domain rotations, lead to exposure of the sequestered genomic RNA.

3.1.1 Importance

Mumps virus (MuV) infection may cause serious diseases, including hearing loss, orchitis, oophoritis, mastitis, and pancreatitis. MuV is a negative-strand RNA virus, similar to rabies virus or Ebola virus, that has a unique mechanism of viral RNA synthesis. They all make their own RNA-dependent RNA polymerase (RdRp). The viral RdRp uses the genomic RNA inside the viral nucleocapsid as the template to synthesize viral RNAs. Since the template RNA is always sequestered in the nucleocapsid, the viral RdRp must find a way to open it up in order to gain access to the covered template. Our work reported here shows that a helix structural element in the MuV nucleocapsid protein becomes open when the sequestered RNA is released. The amino acids related to this helix are required for RdRp to synthesize viral RNA. We propose that the viral RdRp pulls this helix open to release the genomic RNA.

3.2 Introduction

Many negative-strand RNA viruses (NSVs) are important human pathogens that frequently cause outbreaks. The Ebola virus outbreak in West Africa in 2014 (54) and the pandemic influenza A virus H1N1 outbreak in 2009 (55) are two recent examples. Some pathogens appear to reemerge in spite of available vaccines, such as mumps virus and measles virus (56–58). Effective controls are needed to combat these pathogens. In order to develop more effective countermeasures, the mechanism of NSV replication should be better understood. One of the unique features in NSVs is that the genomic RNA is sequestered in the nucleocapsid (59). During transcription and replication, the viral RNA-dependent RNA polymerase (vRdRp) must be able to gain access to the sequestered genomic RNA in order to use it as the template.

For *Rhabdoviridae* and *Paramyxoviridae*, the virus encodes a single nucleocapsid protein (N) that polymerizes as a linear capsid to encapsidate the genomic RNA (60). The viral polymerase complex consists of the large protein (L) and the phosphoprotein (P).

The structure of the nucleocapsid or a nucleocapsid-like particle has been solved for several members of *Rhabdoviridae* and *Paramyxoviridae* by X-ray crystallography or cryo-electron microscopy (cryo-EM) three-dimensional (3D) reconstruction (4, 61–64). The common features among various structures are that the N protein has an N-terminal domain and C-terminal domain in its core, composed mostly of α -helices. When the N subunits assemble into a polymeric capsid, they are aligned in parallel in a linear fashion (65). There are extensive side-by-side interactions between the neighboring domains and domain swaps of extended loops and long termini. The genomic RNA is encapsidated in a cavity formed between the two core domains. Most of the RNA bases are stacked, some of which face the exterior and some the interior of the N protein core. The tight assembly of the nucleocapsid clearly suggests that vRdRp must open the N protein core in order to unveil the genomic RNA. How this action is carried out remains to be discovered. The interaction of the polymerase cofactor P with the nucleocapsid may provide some insights on this subject. The C-terminal domain of vesicular stomatitis virus (VSV) P protein binds between the extended loops in the C-terminal domains of two neighboring parallel N subunits (66). Since the P protein is a part of the vRdRp complex, this binding will place the polymerase in a close proximity to the “gate” covering the genomic RNA. However, the binding of VSV P protein does not seem to induce a significant conformational change in the N protein. It has also been shown that an N-terminal fragment of VSV P protein binds in a truncated empty capsid with an α -helix that sits in the RNA cavity and an extended N-terminal polypeptide that occupies the space vacated by the deletion of the N-

terminal arm of VSV N protein (67). However, this fragment of the VSV P protein could not bind the nucleocapsid or release the genomic RNA. It seems that the P protein can bind the nucleocapsid but is not able to unveil the genomic RNA alone.

Mumps virus (MuV) and parainfluenza virus 5 (PIV5) are members of *Rubulavirus*, a genus of *Paramyxoviridae*. The nucleocapsids of these two viruses tend to coil into a helical structure even when packaged inside the virion. There are on average 13 subunits per turn in the helical structure (68). A unique characteristic of viruses like rubulaviruses is that the length of the RNA genome should be an integer with 6 as a divisor, the so called “rule of six,” suggesting that the single N subunit repeat corresponds to a repeat of 6 nucleotides in the encapsidated genomic RNA (69). In previous studies, it was shown that the MuV N protein forms a ring of 13 subunits when it is coexpressed with the P protein in *Escherichia coli* (68). A 78-nucleotide piece of RNA presumably having random sequences was found encapsidated in the ring structure. The packaged RNA could be easily removed by high salt concentrations, low pH, or RNase A, in contrast to the case for VSV nucleocapsid. The C-terminal domain of MuV P protein binds between the two neighboring N subunits in the nucleocapsid, having a similar stoichiometry as VSV P protein. However, the binding site for MuV P protein is closer to the N-terminal domain of its N protein, whereas the binding site for VSV P protein is closer to the C-terminal domain of its N protein. Moreover, the N-terminal domain of MuV P protein also binds and uncoils the nucleocapsid (40). The P N-terminal domain alone can enhance viral RNA synthesis, which has not been reported for other NSVs. Recently, the crystal structure of a truncated PIV5 N-RNA ring was reported (64). The N protein of PIV5 has the same typical two-domain fold as the N protein of other NSVs and is most homologous to that of Nipah virus (NiV) (70). In each N subunit, six nucleotides were covered, with three stacked bases facing the interior

and three facing the exterior of the N protein. By comparing the structure of the PIV5 N protein with that of the monomeric NiV N protein, which seems to have a more open conformation between the two N domains, it was hypothesized that the C-terminal domain of PIV5 N protein needs to rotate out in order for the viral polymerase complex to unveil the sequestered RNA (64). In this report, we show that a loop-helix $\alpha 7$ region in MuV N protein is the most flexible region when the sequestered RNA is released. Mutation of a few residues in this region also diminished or reduced viral RNA synthesis. These data suggest an alternative mechanism, namely, that the viral polymerase complex needs only to induce a local conformational change of the loop-helix $\alpha 7$ to unveil the sequestered genomic RNA and does not need to bend open a very stable N protein core.

3.3 Materials and methods

3.3.1 Expression and purification of recombinant MuV N protein

All plasmid sequences were based on MuV isolated during an outbreak in Iowa in 2006 (GenBank accession no. JN012242). The coding sequence corresponding to residues 1 to 379 of MuV N protein was coexpressed with the His tagged P protein using plasmid pET28b. After purification of the protein complex with an Ni column, the N₃₇₉ protein was purified by ion-exchange chromatography (HiTrap Q HP; GE Healthcare). The purified N₃₇₉ protein sample that contains random RNA was dialyzed against 20 mM HEPES, pH 7.5. The packaged RNA was removed from the purified N₃₇₉ protein by treatment with 1 mg/ml of RNase A overnight at room temperature.

3.3.2 Cryo-EM structure of the empty MuV N assembly.

Cryo-EM images of the N₃₇₉-RNA complex and the empty N₃₇₉ complex were collected at the National Resource for Automated Molecular Microscopy (NRAMM) in the Scripps Research

Institute. Data were acquired on a Tecnai F20 electron microscope operating at 200 kV, with a Gatan 4kx4k charge-coupled device (CCD) camera to record the images at a pixel size of 1.21 Å. A total of 191 images were included in the final data set for the empty N₃₇₉ complex. Segments of the empty capsid helices were selected with heliboxer from the SPARX/EMAN2 package, using 10% overlap between particles. Particles were aligned using a mask. Contrast transfer function (CTF) corrections were performed with EMAN2. The IHRSR method was used to refine the 3D reconstruction using a noisy cylinder of 220 Å in diameter as the initial model. The initial parameters used for the helical refinement were from the cryo-EM structure of the authentic nucleocapsid (40). The final structure was refined with 5,578 particles, and the resolution was determined to be 10.4 Å using FSC = 0.5. The helical structural model was constructed using PyMol (71) and PIV5 N coordinates from PDB code 4XJN. NiV N coordinates were from PDB code 4CO6. Fitting of the model coordinates into the density was carried out with Chimera (72). Segmentation and superposition of densities were also performed with Chimera.

3.3.3 *Minigenome assays.*

The point mutations in the loop-helix $\alpha 7$ region of MuV N were generated by introducing point mutations into the MuV N gene previously cloned into the pCAGGS expression vector (73). Mutations were introduced by splicing by overlap extension (SOE) mutagenesis using Phusion polymerase (Thermo Scientific), as previously described (74). All constructs were confirmed by sequencing at Genewiz.

The minigenome assay was performed as previously described (74). In short, BSR-T7 cells (1 day, 60 to 80% confluent, 24-well plate) were transfected with pCAGGS-P (80 ng), pCAGGS-L (500 ng), pT7-MG-RLuc (100 ng), and pFF-Luc (1 ng), along with various amounts

of pCAGGS-N (wild type [wt], Tyr185Pro, Ala197Gln, or Gln200Arg at 25, 50, 100, or 200 ng) using jetPRIME (Polyplus) according to the manufacturer's specifications. After 48 h, cells were lysed and a dual-luciferase assay (Promega) was performed using a portion of the lysate.

Luminescence was measuring using a GloMax 96 microplate luminometer (Promega). The ratio of *Renilla* to firefly luciferase was reported for 4 experimental replicates.

Expression levels of N were determined by Western blotting using a portion of the minigenome lysate. All four experimental replicates were combined and mixed with 2× Laemmli sample buffer (Bio-Rad) containing β-mercaptoethanol. Samples were heated at 95°C, resolved on 10% mini-Protean TGX protein gels (Bio-Rad) by SDS-PAGE, and transferred to Amersham Hybond LFP polyvinylidene difluoride (PVDF) membranes (GE Healthcare Life Sciences).

Immunoblotting was performed with an anti-N monoclonal antibody (MAb), followed by incubation with a Cy3-conjugated goat anti-mouse IgG (Jackson ImmunoResearch). The blot was visualized on a Typhoon FLA 7000 instrument (GE Healthcare Life Sciences), and densitometry was performed using ImageQuant TL (GE Healthcare). Values were normalized to wt N at 50 ng/well.

3.4 Results

3.4.1 *Structure of a truncated empty capsid.*

In a previous study, the MuV N protein was coexpressed with the P protein, and a ring of 13 subunits that packages random RNA inside was isolated (68). When the purified ring was stored for a few weeks at 4°C, it was found that the N protein was truncated after residue 379 (N₃₇₉) (75). The same truncation could be generated by trypsin treatment. Here, a vector was constructed to coexpress N₃₇₉ with the P protein that has a His₆ tag at the N terminus. The N₃₇₉-P complex was purified using an Ni column, and the N₃₇₉ protein was further purified with an ion-

exchange column. N₃₇₉ still forms a ring of 13 subunits with random RNA packaged. However, the rings of N₃₇₉-RNA can stack and transform into a nucleocapsid-like helical structure (**Figure 3-1A**). The random RNA sequence could be removed with RNase A, and long helical empty capsids were formed by N₃₇₉ (**Figure 3-1B**). The structure of the empty capsid formed by N₃₇₉ was determined to a 10.4-Å resolution by cryo-EM 3D image reconstruction (**Figure 3-1C**). The diameter of the left-handed truncated empty capsid is about 218 Å, similar to that of the authentic nucleocapsid purified from mumps virions (40). The pitch height, however, is much lower, at about 49 Å (a rise of 3.7 Å per subunit), compared to 67 Å for the authentic nucleocapsid. The rotation of one subunit to the next is 27° about the central axis, making 13.3 subunits per turn, compared to 12.7 subunits per turn in the authentic nucleocapsid. The crystal structure of the PIV5 N-RNA ring complex contains only the fragment of N₄₀₁ (64). Since the sequence of MuV N₃₇₉ is highly homologous to the sequence of N₄₀₁, the coordinates of residues 3 to 379 without RNA from this crystal structure were used to construct an atomic model by rotating a subunit by 27° counterclockwise about the central axis and downshifting by 3.7 Å. The model as a rigid body fits our density map well (**Figure 3-1D**). The empty space between the N- and C-terminal domains of the N protein density is consistent with removal of RNA by RNase A treatment. It appears that transition from a ring structure to a helical structure does not require significant conformational changes in the subunits. In this model, no rotation was introduced in either the N- or C-terminal domain of the N protein.

3.4.2 Comparison with the authentic nucleocapsid.

The structure of the truncated empty capsid was compared with that of the authentic nucleocapsid determined at an 18-Å resolution (40). The density of the two structures was segmented at about the 2σ contour level. As shown in **Figure 3-2A**, the two segments can be

superimposed fairly well except for two regions. The authentic nucleocapsid has more density between the N- and C-terminal domains, consistent with having genomic RNA encapsidated in the nucleocapsid. As shown in **Figure 3-2B**, there is no density corresponding to the location of RNA when the coordinates of PIV5 N are superimposed in the segment. There is also a piece of extra density in the authentic nucleocapsid near the C-terminal end of N₃₇₉. When the density of an N₃₇₉ segment is superimposed onto the helical structure of the authentic nucleocapsid (**Figure 3-2C**), this extra piece of density is involved in the contact between the successive turns. The extra density seems to be responsible for increasing the pitch height of the authentic nucleocapsid. Part of residues 380 to 549, missing from N₃₇₉ (also known as the N-tail), corresponds to this piece of density, but the size of the density is too small to account for all residues. While some of the N-tail residues make contacts between the successive turns, the rest of the N-tail is likely to point to the exterior of the helical structure, where it was shown to interact with the nucleocapsid binding domain of the P protein (40).

3.4.3 *How is the encapsidated genomic RNA unveiled by the viral polymerase?*

As shown for a number of negative-strand RNA virus nucleocapsid-like structures, the genomic RNA is sequestered in the nucleocapsid with some of the stacked nucleotide bases facing the interior of the N protein (76). In order to use the sequestered genomic RNA as a template for viral synthesis, a conformational change must be induced by the viral polymerase complex to temporarily release the RNA from the N protein. It has been suggested that one of the two N protein domains surrounding the genomic RNA can swing open so the template RNA becomes accessible by vRdRp (64). The N structure from the PIV5 N₄₀₁-RNA complex was compared with that of a truncated N protein (N₃₂₋₃₈₃) of NiV in complex with a fragment of the P protein. There is no RNA in the NiV N₃₂₋₃₈₃ structure. The comparison showed that the N- or C-

terminal domain of the two N proteins can be superimposed separately. If the N-terminal domains were superimposed, it would require a 20° rotation to bring the C-terminal domain of PIV5 N to overlap that of NiV N. This observation was the basis for the hypothesis that the C-terminal domain of PIV5 N is the domain that rotates upon polymerase binding, not the N-terminal domain (64). To examine this hypothesis, the coordinates of PIV5 N and NiV N₃₂₋₃₈₃ were superimposed onto the density of the truncated empty capsid of MuV. The PIV5 N structure can be superimposed well without any conformational changes in the two domains. On the other hand, only one of the two domains in NiV N₃₂₋₃₈₃ may be properly superimposed in the density each time. If the N-terminal domain is superimposed, the C-terminal domain of NiV N₃₂₋₃₈₃ will stick out of the density toward the interior of the helical empty capsid. If the C-terminal domain is superimposed, the N-terminal domain of NiV N₃₂₋₃₈₃ will be outside density. However, the motion that may bring the N-terminal domain back into the density requires mostly a rotation about the helical axis (**Figure 3-3C**). The two sets of coordinates were also mapped based on the B factor, a factor that correlates with structural stability (**Figure 3-3A and D**). The core of the PIV5 N-RNA complex has very low B factors, suggesting a high structural stability. This will make it very hard to open either the C- or N-terminal domain because of the high energy requirement. Similarly, the core of NiV N₃₂₋₃₈₃ also has lower B factors. However, the surface residues in the N-terminal domain NiV N₃₂₋₃₈₃ have relatively higher B factors. This is consistent with the fact that the truncated monomeric NiV N protein has neither neighboring subunits nor RNA to stabilize it. It was also observed that when the RNA was removed from the authentic nucleocapsid, the MuV empty capsid became more flexible (40). We noticed, however, that residues in helix $\alpha 7$ and the prior loop of PIV5 N also have very high B factors comparable to the residues in the N and C termini, among which residues 183 to 186 were actually

disordered (**Figure 3-3B**). The homologous region in the NiV N has similar B factors and disordered residues. When the coordinates of PIV5 N are superimposed with the structure of the MuV truncated empty capsid, there is no density corresponding to the loop-helix $\alpha 7$ (**Figure 3-2B**). The same observation was made with superposition of the NiV N_{32–383}, in which there is no corresponding density for the homologous loop-helix. We propose that the viral polymerase complex is required only to open this loop-helix $\alpha 7$ region in order to gain access to the sequestered RNA, instead of bending open a very stable protein core by rotating either the N- or C-terminal domain.

3.4.4 Residues in the loop-helix $\alpha 7$ are critical.

We compared the sequence of the MuV N protein with that of the PIV5 N protein and found that in the loop-helix $\alpha 7$, three residues are different between the two proteins: residues Tyr185, Ala197, and Gln200 (**Figure 3-3A**). Since vRdRp of one virus could not work with other nucleocapsids, we argue that changing these three amino acids of MuV to those of PIV5 would compromise viral RNA synthesis if they are required for vRdRp interactions with the N protein. Three mutant N proteins were therefore generated, corresponding to the changes Tyr185Pro, Ala197Gln, and Gln200Arg. The minigenome activity assay and Western blotting for quantitating the expression of mutant MuV N proteins were carried out, and the results are summarized in **Figure 3-4**. According to Western blot analysis, the mutant N proteins have levels of expression similar to or even higher than that of the wt sequence (**Figure 3-4A**). However, the minigenome activities using the mutant N proteins were significantly reduced compared to that using the wt N protein (**Figure 3-4B**). Gln200Arg had almost no activity, consistent with the fact that mutation from Gln to Arg represents a large side chain change from a polar residue to a positively charged residue.

3.5 Discussion

Transition of the MuV N-RNA complex to a helical structure suggests that the helical nucleocapsid is a very stable structure. By changing to a helical structure, most of the lateral interactions observed in the N-RNA ring are likely to be preserved, especially if the N subunit is allowed to rotate itself. The interactions observed in the crystal structure of NSV N-RNA rings are therefore valid for interpreting N subunit interactions in the authentic nucleocapsid. The protein-protein contact between the neighboring N subunits is massive in the nucleocapsid, provided by side-by-side interactions and domain swaps. Deletion of these interactions will result in disassembly of the capsid and loss of RNA encapsidation (77). In the structure of empty capsids shown here and reported previously (77), the same interactions are retained, suggesting that release of sequestered RNA may not require global conformational changes in the N protein. As shown in the PIV5 N-RNA structure, each C-terminal domain has an interface of 327 \AA^2 with both sides and domain swapping of its C-terminal arm (residues 373 to 401) with another neighboring N subunit. The interactions of the N-terminal domain are even more extensive (883 \AA^2). A global conformational change of either domain will cost a large amount of destabilization energy. In addition, the integrity of the nucleocapsid must be restored when the viral polymerase finishes RNA synthesis at the region where RNA is unveiled. It may not be reversible if large conformational changes are induced in the N protein during viral RNA synthesis.

The conformation of the two domains in NiV N appears to be more open than that in the N protein of respiratory syncytial virus (70). However, the two N domains may become more closed in other NSV N proteins, such as bunyaviruses (78–81). The degree of openness of the two domains may not necessarily be related to the mode of RNA encapsidation by the N protein or the flexibility of the N protein core. If a hinge is present between the two N domains, it is

likely that a more flexible linker is present in the N protein core. Based on the B factors, there is no such flexible linker between the two domains in either PIV5 or NiV N. Moreover, the crystal structure of a measles virus Ncore-P complex shows that the RNA-free monomeric N protein has a more collapsed conformation than that in the nucleocapsid (82). The conformation of the RNA-free N protein is therefore not related to how the sequestered RNA is unveiled during viral RNA synthesis.

An N-terminal fragment of NiV P could bind the RNA-free monomeric truncated N₃₂₋₃₈₃, but no evidence supports that this fragment could bind the nucleocapsid (70). The P protein functions as a chaperone to keep the N protein monomeric before nucleocapsid assembly. The published structures showed that the N-terminal regions can bind at the sites that are involved in stabilizing the nucleocapsid, such as interactions for domain swapping or RNA binding (70, 82). Once the monomeric N subunit is incorporated in the nucleocapsid, the P protein must be dissociated, and the interactions of the N subunits are established. There is no evidence to suggest that the P protein can compete with such cooperative interactions. Our results showed that the loop-helix $\alpha 7$ is the most flexible region when sequestered RNA is released. When RNA is removed, this region becomes more flexible. Single-amino-acid mutations of the MuV N sequence to those of the PIV5 N sequence significantly reduced the minigenome activity (**Figure 3-4**), reaffirming the involvement of this region in viral RNA synthesis. We suggest that specific interactions of this MuV N region with the polymerase are required for unveiling the RNA for viral RNA synthesis. These observations are consistent with the proposed mechanism that the viral polymerase complex can unveil the sequestered RNA by inducing a local conformational change of the loop-helix $\alpha 7$. The loop-helix $\alpha 7$ should be able to readily restore the structure of the nucleocapsid after viral RNA synthesis because it is only a local conformational change.

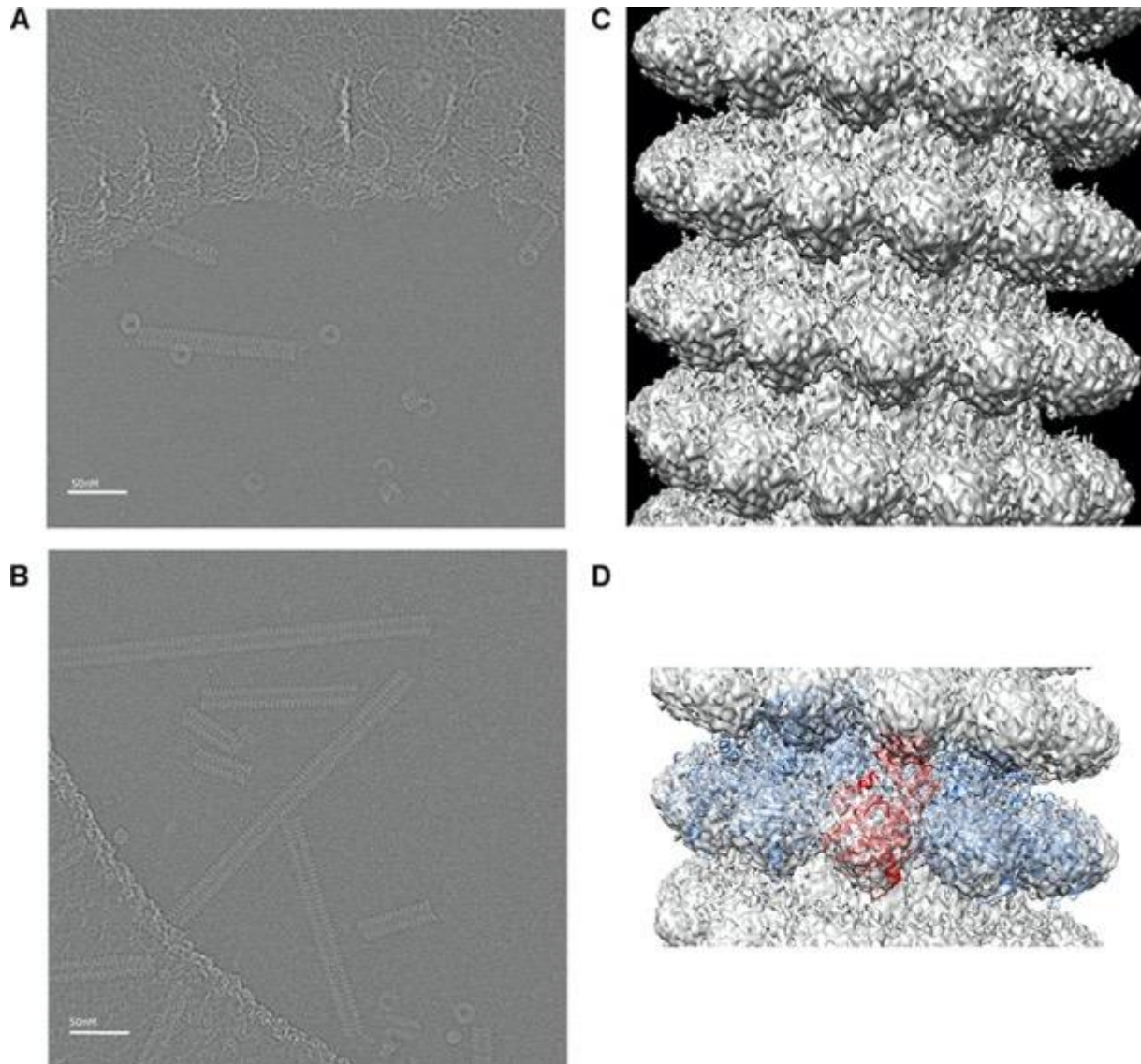


Figure 3-1 CryoEM Structure of N₃₇₉

(A) Superposition (central image) of the segmented density of the empty N₃₇₉ helical structure (transparent gray) with that of the authentic nucleocapsid (cyan). (B) The coordinates of PIV5 N without RNA (red ribbon) were superimposed with the segmented density. Half of the segmented density was removed by slicing through the center. The loop-helix $\alpha 7$ region is labeled. The cleft in the MuV density corresponds to the location where the RNA would be released. (C) Superposition of the segmented density of the empty N₃₇₉ helical structure (gray) with the helical density of the authentic nucleocapsid (cyan). The density of the authentic nucleocapsid is clipped through the center.

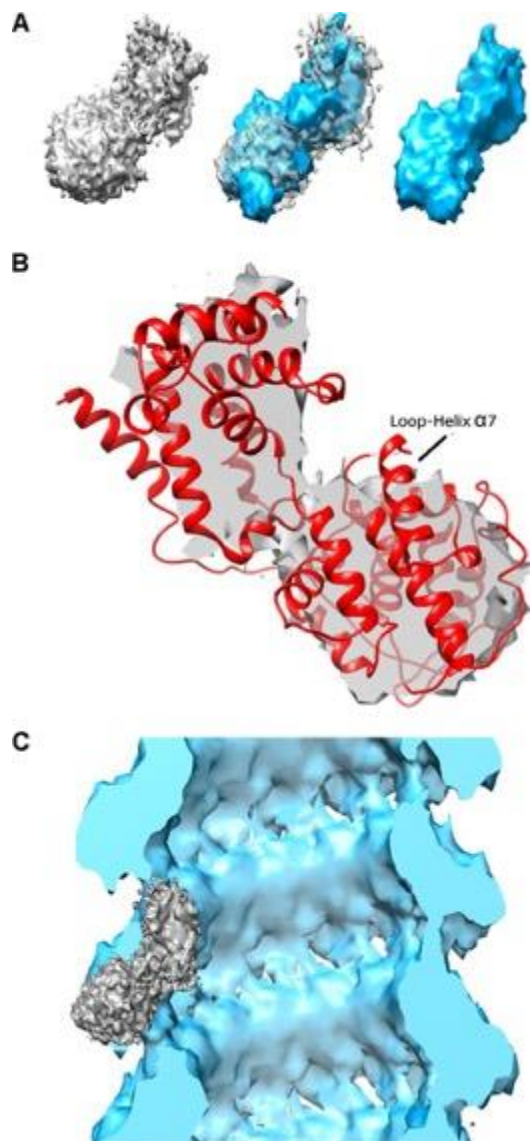


Figure 3-2 Comparing empty capsid with that of the authentic

(A) Superposition (central image) of the segmented density of the empty N_{379} helical structure (transparent gray) with that of the authentic nucleocapsid (cyan). (B) The coordinates of PIV5 N without RNA (red ribbon) were superimposed with the segmented density. Half of the segmented density was removed by slicing through the center. The loop-helix $\alpha 7$ region is labeled. The cleft in the MuV density corresponds to the location where the RNA would be released. (C) Superposition of the segmented density of the empty N_{379} helical structure (gray) with the helical density of the authentic nucleocapsid (cyan). The density of the authentic nucleocapsid is clipped through the center.

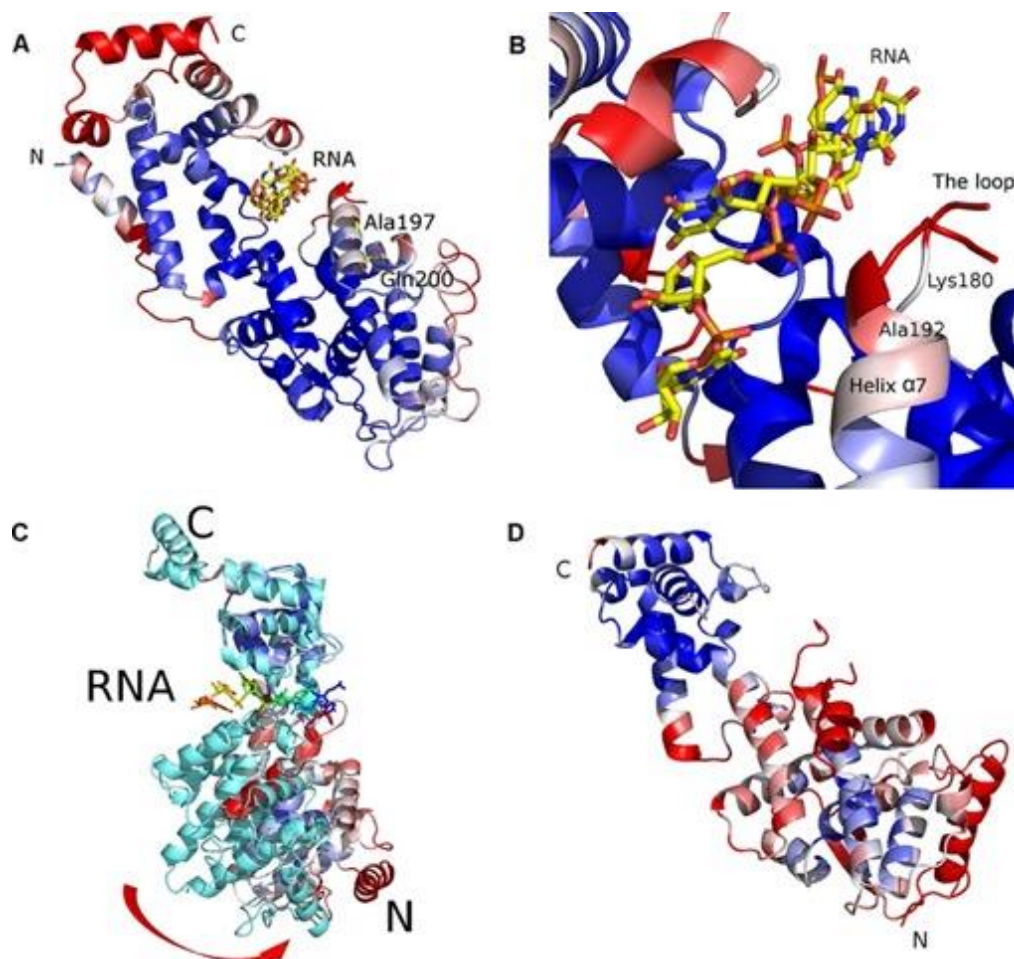


Figure 3-3 RNA binding per N subunit

(A) Ribbon drawing of the PIV5 N structure (4XJN) colored by the B factor (side view). The blue color corresponds to a low B factor, whereas the red color corresponds to a high B factor. N and C, N terminus and C terminus, respectively, of the PIV5 N-RNA complex. The encapsidated RNA is shown for one N subunit as a stick model. Residues Ala197 and Gln200 are displayed as sticks and labeled. Residue Tyr185 is not present in the crystal structure. (B) A close-up view of helix $\alpha 7$ and the loop prior to helix $\alpha 7$ in the PIV5 N structure. Residues Lys180 and Ala192 are at each end of this flexible region. Disordered residues 183 to 186 are not present. (C) Ribbon drawings to illustrate the motion required to fit the coordinates of NiV N (PDB code 4CO6) with the density of the empty N₃₇₉ helical complex. The structure of NiV N is represented by a ribbon in cyan. The structure of PIV5 N as fitted in the density of the empty N₃₇₉ helical complex is represented by a ribbon colored from blue to red by B factors. The view is approximately down the axis of the empty N₃₇₉ helical complex. The C-terminal domains of the two structures were superimposed together. The red arrow indicates the rotational motion required for the N-terminal domain of NiV N to be superimposed with that of PIV5 N. N and C, N and C termini of PIV5 N, respectively. RNA is encapsidated in the center of PIV5 N. (D) Ribbon drawing of the NiV N structure colored by the B factor (side view). The blue color corresponds to a low B factor, whereas the red color corresponds to a high B factor. N and C, N terminus and C terminus, respectively.

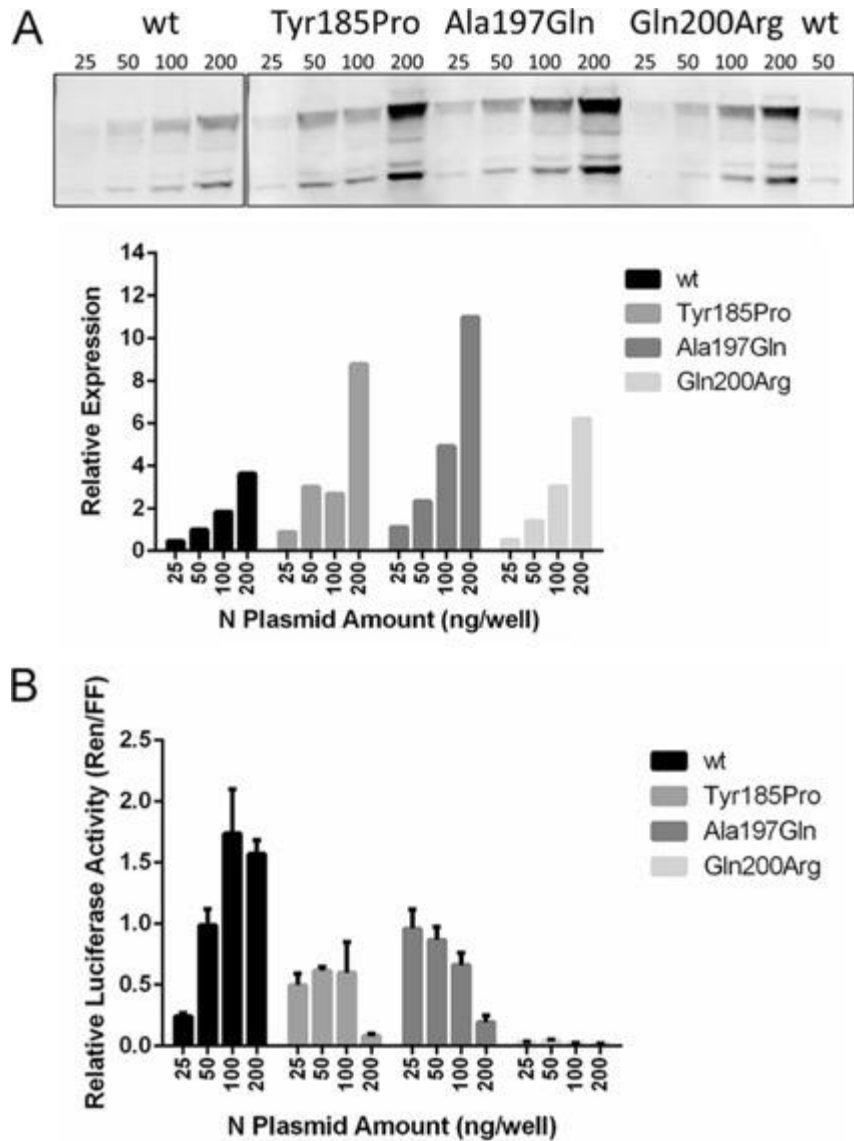


Figure 3-4 Minigenome assay of mutants

(A) Expression levels of the mutated MuV N proteins detected by Western blotting, compared with that of wt MuV N. The full-length N protein was used for quantitation. (B) Activities of the minigenome reporter gene when the mutated MuV N proteins were used in the system. wt MuV N was used as the positive control.

4. PROBING THE ROLE OF THE PHOSPHOPROTEIN AMINO TERMINAL END IN THE UNCOILING OF THE NUCLEOCAPSID

Chelsea Severin and Ming Luo

4.1 Abstract

Paramyxovirus genome synthesis involves specific interactions between the large protein, the phosphoprotein (P) and the helical viral ribonucleoprotein (vRNP) complex. The binding activity of P is necessary to dock the nucleocapsid (N) protein coated RNA to the large protein for viral replication. Previously, we have shown in mumps virus (MuV) that the amino terminal domain of P binds to a nucleocapsid-like particle (NLP) and induces uncoiling of the authentic nucleocapsid (40, 75). However, the actual region of the P N-terminal domain (P_{NTD}) involved in this interaction has not been determined. Here, we sought to delineate the specific residues of P_{NTD} involved in the uncoiling of the viral genome. We have developed an *in silico* model of the MuV NLP-P_{NTD} complex and have identified a region within P_{NTD} which is necessary for interaction with the NLP. Using site directed mutagenesis and binding assays we have identified residues which are necessary for binding to the NLP. This work provides insights into the mechanism of the polymerase activity of P.

4.2 Introduction

The *Paramyxoviridae* family of viruses include many pathogens which are infamous for their toll on public health and the agricultural economy. This group includes many notable

viruses such as mumps virus (MuV), parainfluenza viruses, measles virus (MeV), and Nipah virus (NiV). Due to their single strand negative polarity RNA genome, paramyxoviruses can be further categorized into the broad group of negative stranded RNA viruses (NSVs). The paramyxovirus RNA genome is non-segmented and about 15 kb in length. The viral genome is always coated by the N protein to form a helical viral ribonucleoprotein (vRNP) and this complex is the active transcript for viral replication and transcription. For efficient infection in host cells, paramyxoviruses require interaction between the viral RNA dependent RNA polymerase (vRdRp) and the helical vRNP. The vRdRp is a complex formed between the large (L) protein and the phosphoprotein (P). Catalytic activity of the polymerase resides in L while P is necessary for recognizing the N-RNA complex for transcription and replication.

During viral RNA synthesis, P must recognize the nucleocapsid and remain associated with the N protein as the polymerase moves across the viral genome. The location of the nucleocapsid binding domain on several paramyxoviruses and other NSVs have been mapped. In vesicular stomatitis virus (VSV) of the *Rhabdoviridae* family, the C-terminal domain of P binds to the C-terminal domain of two adjacent N protomers within the nucleocapsid like particle (66). As a chaperone, P also tethers to soluble, monomeric N to prevent illegitimate interactions with cellular RNA and nascent N-N oligomerization. Upon a switch from transcriptase to replicase mode which is believed to be modulated by the concentration of free N protein, a conformational change in the interaction between N and P allows the monomeric N protein (N⁰) to encapsidate the newly formed viral RNA. In paramyxoviruses, the N-terminal domain has been purported to play a role in N⁰-P binding. High resolution structures of MeV, PIV5 and NiV have indicated that the amino terminal end of the phosphoprotein binds to the C-terminal end of the

nucleocapsid (70, 82, 83). Examination of these structures further suggest that binding of P to the nucleocapsid like particles induces conformational changes in the RNA binding groove.

The P protein of MuV is a highly phosphorylated tetramer composed of an N-terminal region (aa 1-194), a central core region which has an oligomerization domain (aa 213-277) and a C-terminal domain (aa 277-391) (75). The amino terminus contains the binding site of polo-like kinase I which phosphorylates S292/S294 in the C-terminal domain of P(84). In other paramyxoviruses, the N-terminal domain of P is responsible for preventing N from binding to cellular RNA. The nucleocapsid binding site has been mapped to the C-terminal end of the protein. The structure of the oligomerization domain reveals a novel structure with two pairs of anti-parallel alpha helices which position two amino terminal and two carboxy terminal ends at each end of the tetramer(75). Studies have shown that the oligomeric state of the protein is necessary for viral replication which suggests that MuV might possess a unique method of replication.

During the replication cycle, the viral genome is always associated with the nucleocapsid protein. This encapsidated RNA is the active template for transcription which P recognizes. The N protein of MuV is comprised of two domains: an N-terminal core and a C-terminal tail(68). The N-terminal domain is responsible for binding to the polymerase and RNA encapsidation. The carboxy end of P binds to the N-terminal domain of the nucleocapsid core in MuV and other paramyxoviruses. When expressed in bacterial systems, nucleocapsid proteins tend to bind to non-specific RNA and form ring structures which represent one turn of the helical nucleocapsid. Structures of nucleocapsid-like particles (NLPs) have been solved for various negative strand RNA viruses such as respiratory syncytial virus (RSV), VSV and more recently Parainfluenza virus 5 (61, 63, 85). Parainfluenza virus 5 (PIV5) is a member of the *Paramyxoviridae* family

which shares about 42% sequence homology with MuV. The structure of a PIV5 NLP lacking the C-terminal tail revealed that the nucleocapsid-like structure is a 13-mer ring which consists of a core region flanked by a N-arm and C-arm at its N-terminal and C-terminal ends respectively. The RNA is sequestered between the N-terminal and C-terminal domains of the N-core where each monomer binds exactly six nucleotides.

Several studies have sought to determine the mechanism of interaction between the nucleocapsid and phosphoprotein domains in paramyxoviruses. To date, the body of work available on the N-P complex focuses primarily on the interaction between the monomeric form of N and N-terminal peptides of P. However, the actual residues involved in docking the nucleocapsid to the polymerase is still unclear. Here, we utilize homology models and rational mutagenesis to identify important residues in an amino terminal fragment of the phosphoprotein which are important for interaction with a NLP of MuV. The data implicates critical residues in the N-terminal domain of the phosphoprotein are necessary for binding the NLP. Furthermore, these new insights will allow the rational design for antivirals that target the polymerase complex.

4.3 Materials and Methods

4.3.1 Molecular Biology

The genes encoding N and P_{NTD} were previously cloned into pet28B vector (68, 75). Briefly, the NLP was expressed using a pet28b vector which was designed to coexpress the N and P proteins in equimolar ratios with an N-terminal His tag residing on the P gene. Constructs of the N-terminal fragment of P and its alanine mutants were generated by site directed mutagenesis of the P NTD plasmid using Phusion site directed mutagenesis kit (ThermoFisher)

according to the manufacturer's protocol. All plasmids were verified by sequencing (Eurofins Genomics).

4.3.2 Computer aided studies

Predictive 3D structures of the NLP and P110 were generated using MODELLER 9.20 (86). The structures of PIV5 NLP and PIV5 V were used as templates for the models of MuV N_{CORE} (1-401) and P110 (aa 1-110) respectively (39, 87). Alignment of the target and template sequences were generated using python script salign.py with MODELLER. The alignment file was used to generate 10 models of the target protein, of which the structure with the lowest discrete optimized protein energy (DOPE) score was selected for further model validation. The selected model was evaluated using PROCHECK(88) and the Ramachandran plot was used to examine the dihedral angles of the residues to assess the quality of the models. The best models of N_{CORE} and P110 were used to perform docking studies of the proteins using HADDOCK2.2 (89) server. The active and passive residues were defined based on the structure of the PIV5 N⁰-P₅₀ structure. The highest ranked models were visually inspected, and the best model was selected.

4.3.3 Protein expression and purification

Plasmids harboring genes encoding either the N/P genes or P110 wild type and mutant proteins were transformed into *Escherichia coli* Rosetta DE3 cells. A starter culture was grown overnight in 2x YT medium supplemented with 50 µg/µl of kanamycin. This starter culture was used to inoculate 1 L of medium containing 50 µg/µl of kanamycin and the cells were grown at 37 °C until the OD was between 0.6 and 0.8. Protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside at 25 °C for 16 h or 18 °C for 18 h for the N and P110

mutants respectively. The bacteria were harvested by centrifugation at 4000 rpm and stored at -80 °C.

For protein purification of the P110 wild type and variant proteins, bacterial cells expressing the protein of interest were resuspended in binding buffer A containing 50 mM Tris (pH 8), 500 mM NaCl, 5 mM Imidazole, 5 mM β -mercaptoethanol (β -ME) and 5% glycerol. The cells were lysed by sonication after which the insoluble cell debris was separated by centrifugation at 18000 rpm for 45 min. The cleared lysate was loaded onto a 5 ml HisTrap HP Ni affinity column (GE Healthcare) which had been equilibrated with binding buffer A. The column was washed to UV baseline with binding buffer. Non-specifically bound proteins were eluted in 5 column volumes (CV) of washing buffer (50 mM Tris pH 8, 500 mM NaCl, 50 mM imidazole, 5 mM β -ME, 5% glycerol). The protein was eluted from the column in 5 CV of elution buffer containing 50 mM tris pH 8, 500 mM NaCl, 500 mM imidazole, 5 mM β -ME and 5% glycerol. Size exclusion chromatography (Superdex 75; GE Healthcare) was used as a polishing step and the protein of interest was eluted in buffer containing 20 mM Tris (pH 8) and 150 mM NaCl.

The N protein was purified by Ni affinity column using the same protocol described above for the P110 wild type and mutants. To separate the N and P proteins the fractions from the elution step were pooled together and dialyzed overnight at 4 °C against buffer A8.3 containing 50 mM tris (pH 8.3) and 150 mM NaCl for ion exchange chromatography. The dialyzed protein was loaded onto a 5 ml Hitrap Q HP column (GE Healthcare) that had been equilibrated with buffer A8.3 and the column was washed to UV baseline with binding buffer A8.3. The protein was eluted in a step gradient of 20%, 40%, 66%, 75% and 100% B, where buffer B8.3 contained 50 mM Tris (pH 8.3) and 750 mM NaCl. The fractions from the 66% B

step were pooled together and concentrated with an Amicon centrifugal filter. The concentrated protein was loaded onto a Superose 6 16/600 (GE Healthcare) that had been equilibrated with 20 mM Tris (pH 8) and 150 mM NaCl.

4.3.4 Protein binding assay

To probe the interactions between the NLP and P110 or NLP and P110 alanine mutants a His6 pulldown assay was performed. To 50 μ l of HisPur NiNTA resin (ThermoFisher), 50 μ g of P110 or P110 variants was added and incubated at room temperature for 30 min. The beads were spun down at 4000 rpm for 5 min and the supernatant discarded. The beads were washed with buffer containing 20 mM Tris, 50 mM NaCl and 5 mM imidazole. The NLP (50 μ g) was added to the beads and incubated at room temperature for 30 min under rotation. The beads were washed with binding buffer B and the proteins were eluted with binding buffer B containing 20 mM Tris pH8, 50 mM NaCl and 500 mM Imidazole. The samples were electrophoresed on a 12% SDS-PAGE gel and visualized with Coomassie blue staining. Non-specific binding of the NLP was assessed in the absence of P110 protein.

4.3.5 Thermal shift assay

Thermal scanning was used to further investigate the effect of the mutations on the NLP-P110 complex. NLP and P110 wild type or mutant proteins were incubated at room temperature for 30 min at a 3:1 P to N molar ratio to ensure complete binding. Solutions of 20 μ l of protein sample per well were prepared by mixing 2 μ l of 100X Sypro Orange or Sybr Safe with the protein mixture. The samples were set up in triplicates in a Microamp fast 96 well plate (Thermo Fisher) sealed with optical quality sealing film. The thermal shift assay (TSA) was performed on a Step One Real Time PCR instrument where the temperature was increased with a gradient from 25 to 95 $^{\circ}$ C with a ramp rate of 1%. Raw data was analyzed in MATLAB and smoothed using

Savitzky-Golay method. The normalized data was used to plot first derivative graphs of fluorescence vs temperature and to identify the melting temperature of the protein.

4.4 Results

4.4.1 Identification of N-terminal residues of P involved in the potential NLP binding site

The full-length P protein consists of three independent domains: the amino terminal domain, the core oligomerization domain and the C-terminal domain. We were the first to illustrate that the N-terminal end of the phosphoprotein interacts with the NLP and the authentic viral nucleocapsid of a negative stranded RNA virus. Here, we utilized computer simulations to identify the important residues of the first 110 amino acids of P. The only available 3D structure with significant sequence homology to MuV N and P110 is that of PIV5. Residues 3 to 401 of MuV N were aligned with residues 3 to 401 of PIV5 N_{CORE} structure (PDB 4XJN). MODELLER 9.20 was used to generate 10 models of a N homotrimer of which the model with the lowest DOPE score was selected. The DOPE score represents the highest ranked model. The model was further subjected to validation using PROCHECK. The model had approximated 92% of amino acids falling in the core region, which indicates that the structure is a good quality model. PIV5 V protein was aligned with the protein sequence P110 and showed a sequence identity of 50%. The P110 model was also generated by MODELLER 9.20 and assessed by PROCHECK. The Ramachandran plot indicates that 85.3% of the residues are in most favored regions, 13.3% are in additional allowed regions and 1.3% are in generously allowed regions.

Utilizing these computer generated structures, we docked the homology models of N_{CORE} and P110 using HADDOCK. The structure of PIV5 monomeric N bound to a 50-residue long N-terminal P fragment was recently determined. The 3D structure indicates that the contacts between the PIV5 protein are both hydrophobic and polar. We examined the structure of the

PIV5 N_{CORE}-P₅₀ complex and our docked structure and identified region consisting of amino acids 13-19 (except G17, M18) as necessary for the P110-NLP interaction (**Figure 4-1**). P110 variants were expressed and purified using the same methods as the wt protein. The yield of protein per liter varied for three of the variants. Proteins harboring the L13A and I14A mutations had a greater yield than that of the wt; whereas, N19A had a decrease in expression.

4.4.2 Residues in loop are necessary for NLP-P110 interaction in pulldown assay

To determine the effects of the identified residues on the interaction between P110 and NLP, alanine mutants were analyzed in a pulldown assay (**Figure 4-2**). The proteins were expressed in *E coli* with the 6X N-terminal His tag residing on the P110 protein. Purified P110 wild type or alanine mutants were added to a NiNTA beads and then incubated with the NLP. The bound proteins were eluted and visualized. The experiment confirmed that the P110 wt protein binds to NLP. We also determined that residues L13 and I14 were necessary for interactions with the NLP. Point mutations of these residues to alanine disrupted binding to the NLP.

4.4.3 Effects of targeted mutations on thermal stability of NLP-P110 interaction

To further investigate the effects of the alanine mutations on the NLP-P110 interactions we used a thermal shift assay to determine the impact of P110 on the thermal stability of the NLP. Previously, we have used stability shift analyses to quantify the interactions between VSV NLP and a polyamide (90). Here, a similar approach was used to assess the interactions between the MuV NLP and P110. The protein samples were incubated with a fluorescent dye, Sypro Orange. The melting curve plot shows changes in the fluorescence of the Sypro Orange dye as temperature increases. In this study, we define T_m as the temperature at which the change rate of

fluorescence is at maximum, indicating that increasing hydrophobic sites are exposed. Under our experimental conditions, the NLP exhibits a T_m of 70.3 °C (**Figure 4-3**).

Previously, we have shown that the N-terminal domain of the phosphoprotein can unwind the authentic nucleocapsid and NLP. It follows then that upon binding to the NLP the N-terminal domain of P should destabilize the protein and decrease the T_m . Upon binding to the NLP, the P110 protein decreased the T_m by 0.5 °C (**Figure 4-3**). Proteins harboring mutations L13A and I14A did not have statistically significant effect on the T_m . Interestingly, the N19A mutation decreased the T_m by 1°C to 69.2°C suggesting that this mutation enhanced the destabilization of the NLP.

4.4.4 Alanine variants affect the thermal release of RNA from the NLP

The nucleocapsid protein effectively protects the viral RNA from degradation by formation of helical structures. However, it was shown that RNA loss in the NLP is dependent on pH, temperature and ionic strength. Under thermal denaturation conditions, the NLP undergoes changes to reveal the hydrophobic core; therefore, it is expected that at a point prior to unfolding a conformational change in the protein will encourage the release of the encapsidated RNA. The accessibility of the encapsidated RNA might be therefore monitored using a fluorescent dye which intercalates with nucleic acid. We define the approximated temperature at which the RNA is released from the NLP (T_R) as the peak in the second derivative plot at which the change of the change rate of fluorescence is at maximum.

Analysis of the NLP melting curve when monitored by Sybr Safe has a high initial fluorescence most likely due to tight binding of the dye intercalating between the stacked bases of RNA within the NLP. The thermal release of RNA, T_R , observed from the experiment was measured as 68.1 °C (**Figure 4-3;Table 4-1**). This result was consistent with previous studies

(68). In the presence of the amino terminal fragment of P, the observed T_R decreased by 0.7 °C to 67.5 °C, suggesting that P110 binding of NLP facilitates RNA release. Variants L13A and I14A had no effect on the thermal release of RNA from the NLP; while, N19A decreased the T_R by 2.0°C to 66.1°C.

4.5 Discussion

During viral replication, only the encapsidated viral genome can be used for viral RNA synthesis. The intimate interaction between the N protein which coats the viral genome and the polymerase cofactor P is critical for this process. To effectively design therapeutics that target these viruses identification of the binding mode of this N-P interaction is necessary. However, obtaining high resolution structures of the N-P complex has proven to be difficult due to flexibility in the proteins. The MuV nucleocapsid binding domains has been mapped to the N-terminal and C-terminal domains of MuV P(75). When visualized by negative stain, in the presence of the N-terminal domain of P, the authentic nucleocapsid was unwound indicating that this domain is necessary for unveiling the viral RNA (40). Here, we sought to delineate the residues in the N-terminal domain of P which are necessary for binding the nucleocapsid.

Our model of the nucleocapsid-phosphoprotein complex illustrates that a loop consisting of the first 19 aa sits between the interface of the lobes/arms of the nucleocapsid protein. These arms have been illustrated to be necessary for oligomerization of the nucleocapsid protein subunits in PIV5 and VSV(39, 61). Therefore, it is plausible that the N-terminal domain of P might sit in the interface and disrupt the interactions between N and N+1. To identify amino acid residues in this domain which are involved in the interaction with the NLP binding to facilitate RNA release studies were carried out using recombinantly expressed proteins. We designed alanine mutants of residues D12, L13, I14, E15, T16 and N19 in P110 to assess the effects of

these mutations on NLP binding. In our interaction assay, proteins harboring the L13A and I14A mutations did not pull down the NLP (**Figure 4-2**).

Binding of P110 to the NLP should cause a structural change in the protein and affect its thermal stability. To assay this change, we monitored the interactions of the proteins using the fluorescent dye Sypro Orange. In the presence of P110 the melting temperature of the NLP decreased by 0.5 °C. The decrease in the T_m of the NLP suggests that P destabilizes the NLP. This data is consistent with previous studies which show that the N-terminal domain of P uncoils the nucleocapsid protein. Proteins harboring either L13A or I14A had no effect on the thermal stability of the NLP most likely due to the loss of interaction between the proteins.

To further characterize the effect on RNA release from the NLP we used a thermal shift assay to monitor RNA release using Sybr Safe. When intercalated between bases in nucleic acids, Sybr Safe dye has fluorescence excitation maxima at 280 and 502 nm and an emission maximum at 530 nm. We applied this property to the TSA to determine the temperature at which RNA is released from the NLP. The crystal structure of PIV5 NLP indicates that the RNA strand sits in a polar groove in which the bases are stacked. In the NLP, the environmentally sensitive dye, Sybr Safe, intercalates between the stacked bases of the encapsidated RNA. Since the concentration of RNA is high within the ring, it is expected that the initial fluorescence of Sybr Safe would be high (**Figure 4-3**). The first derivative plot of Sybr Safe fluorescence quantifies the rate of fluorescence change as a function of temperature. Fluorescence emission initially decreases with increasing temperature which represents the release of dye molecules. The fluorescence intensity begins to increase later because the dye is able to intercalate more efficiently into the sequestered RNA at a more appropriate temperature. At a turning point, the fluorescence signal is decreased due to release of the single stranded RNA which has no

structure and thus the dye is unable to intercalate between the stacked bases. The point at which fluorescence change rate is maximum is denoted as T_R , as observed in the second derivative graphs

The release of nucleic acid (T_R) and protein melting (T_m) temperature of the NLP were found to be distinct at 68.1°C and 70.3 °C, respectively. This observation is in agreement with structural data which indicates that the sequestered RNA sits within a polar pocket in the NLP. From the second derivative graph, it was determined that the release of the RNA from the NLP in the presence of P110 was reduced by 0.7°C. However, in the presence of L13A and I14A the release of sequestered RNA was unaffected, which confirms our above results which indicate that these residues abrogate binding of P110 to the NLP.

The MuV P protein is transcribed from the V/P/I gene from a shift in the ORF by insertion of two guanine residues. The V protein is the faithful transcript of the V/P/I gene and V is thought to be necessary for evasion of the host immune responses. P and V share an N-terminal domain; therefore, it is possible that these residues serve as regulation of the viral replication by the V protein. Previously it was shown in PIV5 that the same residues (L16 and I17 in PIV5) are necessary for the V-NP interaction. Furthermore, it was illustrated that these two residues were critical for the inhibitory activity of V in the minigenome system.

The N-terminal domain of P relaxes the nucleocapsid of MuV. We believe that this interaction is necessary to facilitate the release of the encapsidated viral RNA from the nucleocapsid complex. Here, we have identified residues in the amino terminus of P which are important for binding to the NLP. To our knowledge, this is the first study in which the critical residues in the NP-P interaction were identified.

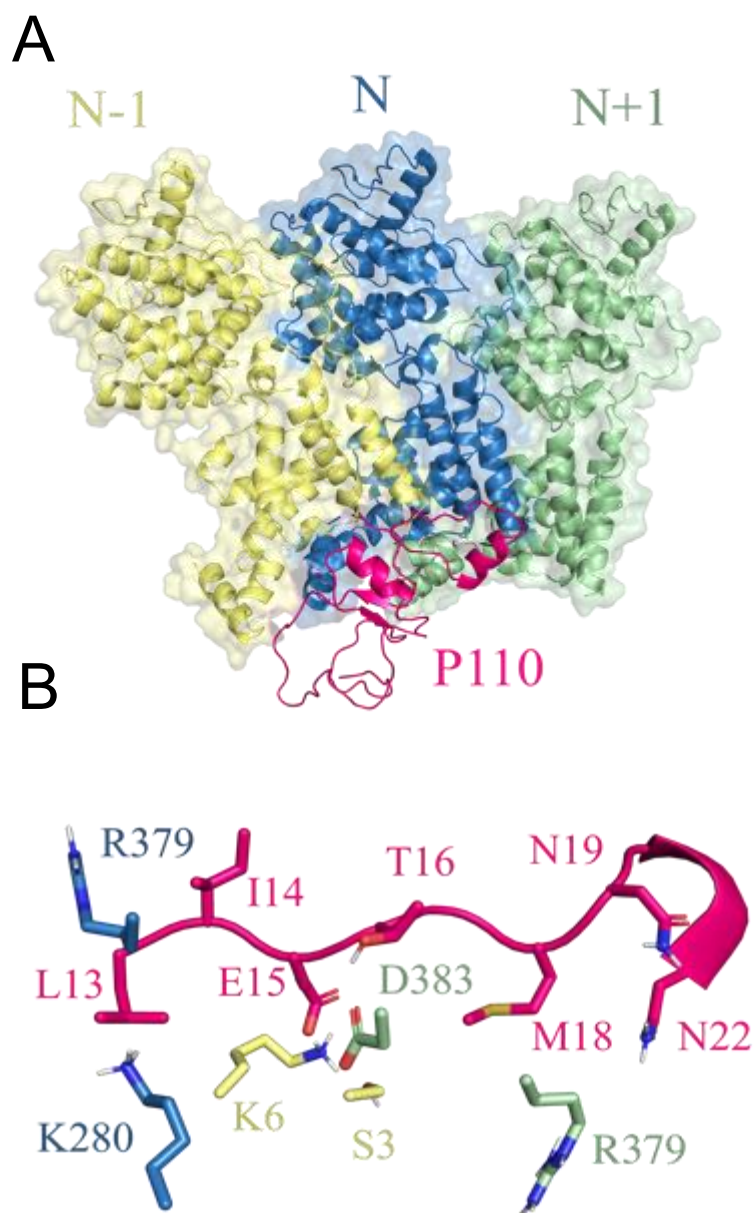


Figure 4-1 Homology Model of N_{CORE}-P110 complex.

(A) MuV NLP-P110 model as determined by HADDOCK. (B) Potential site of interaction between NLP core and P110 according to in silico model.

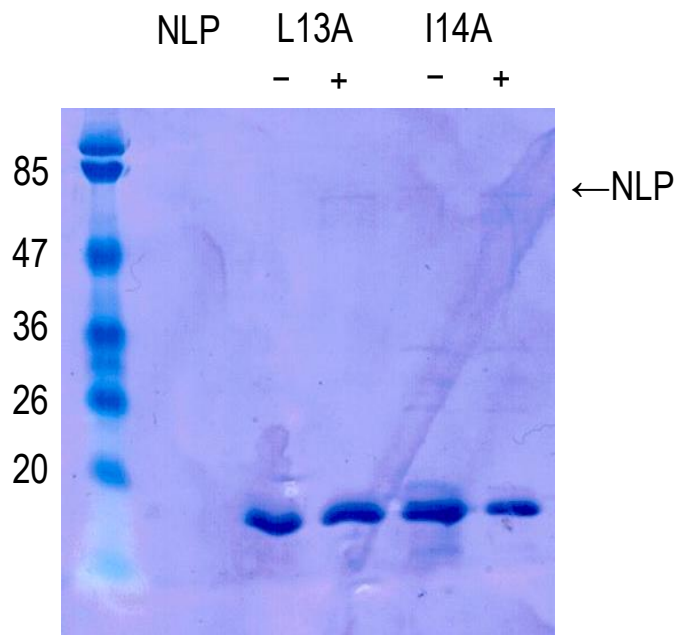
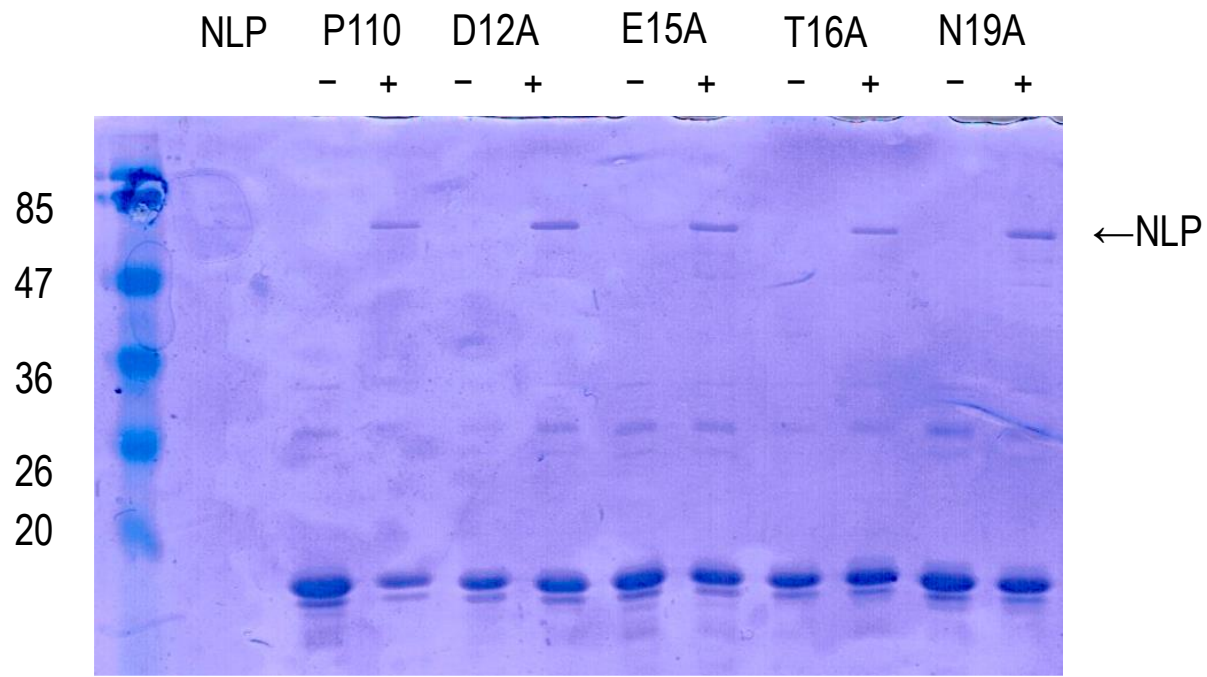


Figure 4-2 Pulldown assay
Interaction of NLP with P110 variants in a pulldown assay. Proteins harboring L13A and I14A were unable to pulldown NLP.

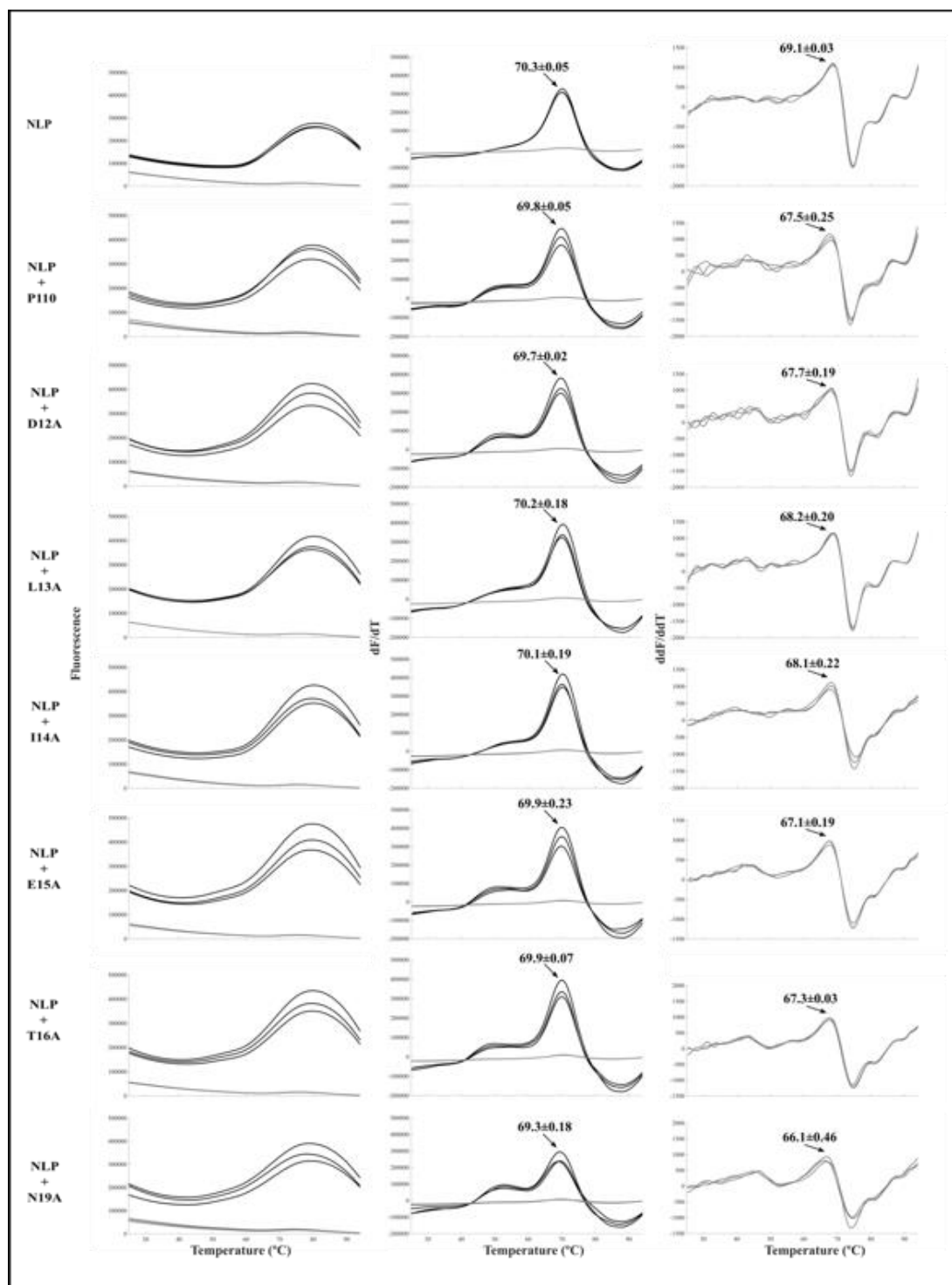


Figure 4-3 Thermal Stability Assay

Melt curve analyses of NLP alone and together with P110 alanine mutants as monitored by Sypro Orange (black) and Sybr Safe (gray) dyes (column one). The first derivatives of curves in column one were plotted for Sypro Orange (column two) and second derivatives were plotted for Sybr Safe (column three). The T_m was determined as the maxima from the plots in column two, whereas the T_R was determined as the maxima from the plots in column three. Errors are shown as standard deviations from triplet experiments.

Table 4-1 Summary of Affinities and Thermal Release of RNA

Sample	T_R^a	ΔT_R	Disruption of binding in pulldown assay?
NLP	68.13 (0.03)	-	-
P110	67.46 (0.25)	-0.67	-
D12A	67.65 (0.19)	-0.48	No
L13A	68.24 (0.20)	0.11	Yes
I14A	68.13 (0.22)	0.00	Yes
E15A	67.11 (0.19)	-1.02	No
T16A	67.31 (0.03)	-0.82	No
N19A	66.14 (0.46)	-1.99	No

^aStandard deviations indicated in parentheses

5. RESEARCH CHALLENGES AND IMPACT

5.1 Crystallization of PB2 cap binding protein

Structural determination of the PB2cap binding domain in complex with the m7 GTP substrate is necessary for rational drug design of inhibitors which target the protein. Crystal soaking and co-crystallization are common techniques used for structural determination of a protein-ligand complex. The apo crystals of PB2cap binding protein were grown in buffer containing 0.1 M HEPES pH 8.0, 20% (w/v) PEG 3350 and 0.2 M magnesium nitrate. However, when we attempted cocrystallization of the protein-ligand complex together no protein crystals were obtained from screening. Alternatively, soaking the wild type crystals with the substrate caused dissolution of the protein crystals. It is possible that crystal packing in the apo crystal form inhibits the functional binding site of the protein. The structure of the unbound PB2cap protein alone includes two molecules in the asymmetric unit. In the crystal structure the cap-binding site is blocked by a long, flexible loop consisting of aa 421-427 (denoted here as 424-loop). To circumvent this, the 424-loop was truncated and replaced by the linker GSG. Utilizing this loop truncation mutant, cocrystals were obtained of the PB2cap binding protein alone and in complex with the cap analogue.

5.2 Implications for Design of inhibitors that target influenza viruses A & B

Antigenic variation in influenza A viruses can cause the emergence of novel viruses and lack of innate immunity against these viruses may cause pandemics. Influenza B viruses also pose a significant threat since these viruses circulate primarily in humans. In the United States, the 2019-2020 flu season was characterized by two waves, with influenza B viruses dominating early in the season and then the A(H1N1)pdm09 virus (91). This confirms that a potent inhibitor of influenza virus infection should target both types A and B viruses.

When the structure of the PB2cap of B/Jiangxi/BV/2006 was compared to that of the 2009 pandemic strain, differences in the binding pocket and important interactions were evident. Firstly, the cap analog, m7 GTP binds further into the binding pocket of PB2cap binding domain of influenza virus H1N1 A/California/ 07/2009. This suggests that a potent inhibitor against influenza A and B viruses should fit the pocket of influenza B viruses. Additionally, the structure of the PB2cap of B/Jiangxi/BV/2006 reveals that His357 of pdm09 strain is replaced by Trp, which makes more π - π stacking interactions with the guanine moiety. Also, Phe323 of influenza A is replaced by Gln, which can no longer provide π - π stacking and His432 is replaced by the bulkier Tyr residue. Our structural analyses of both type A and B PB2cap have laid the foundation for design of a novel anti-influenza virus therapies.

5.3 Viral RNA sequestered within the Paramyxovirus Nucleocapsid

The viral RdRp of paramyxoviruses uses the encapsidated viral genome for viral replication and transcription. The viral RNA is always sequestered within the helical nucleocapsid; therefore, the polymerase complex must somehow open the nucleocapsid. Utilizing the structure of PIV5 NLP and monomeric NiV N it was proposed that the sequestered RNA within the nucleocapsid is revealed once the C-terminal domain of N rotates outwards and allows the polymerase access to the viral genome. Here, cryo-EM was used to analyze the structure of an empty MuV capsid and determine the most likely model for RNA release. The coordinates of the PIV5 NLP structure were mapped onto the empty capsid of MuV. The temperature factor or B factor is a term which reflects the flexibility of different parts of the molecule. When the coordinates of PIV5 were mapped onto the structure of the NLP it revealed that it is unlikely that the C-terminal domain would rotate outward to reveal the viral RNA since this region has low B-factors which suggest high structural stability. Analysis of the structures

suggest that the loop-helix $\alpha 7$ of MuV N may be required for unveiling the genome due to the high flexibility in this region. It is likely that the polymerase induces a local conformational change in the helix $\alpha 7$ region and therefore, N does not need to open a very stable core to reveal the sequestered RNA.

5.4 Uncoiling the MuV Nucleocapsid

MuV P is necessary for the L protein to gain access to the nucleocapsid for viral replication. The mechanism by which the vRdRp gains access to the sequestered RNA is not clearly understood. Previous paramyxovirus models suggest that rotation in the C-terminal domain would expose the viral genome for the viral polymerase. Crystal structure analysis of MuV P show that the protein is a tetramer in which P molecules orient in an antiparallel manner. Additionally, it was shown that the N-terminal domain of P was binds to the authentic nucleocapsid and induces uncoiling. Therefore, it is plausible that the N- and C- terminal domains of P cooperatively induce local conformational changes in the nucleocapsid to reveal the viral genome. The C-terminal domain recognizes the helical nucleocapsid, while the N-terminal domain disrupts the N-N interactions and opens the “gate” to the vRNA. It would be very difficult for the polymerase alone to unveil the genome.

5.5 Implications for design of antivirals and vaccine development of MuV

The nucleocapsid-phosphoprotein interaction constitutes a reasonable target for design of antivirals that inhibit viral replication. In RSV, inhibitors which bind to N and compete with P have been identified. Structural information of the MuV N-P binding will provide an opportunity to design inhibitors which disrupt this interaction. Furthermore, knowledge about the replication mechanism will aid in the development of more effective vaccine candidates that are selected for efficient replication.

6. CONCLUSIONS

The coronavirus pandemic has exemplified the effects of viruses on public health and everyday life when unprepared with therapeutics. We need to identify current and emerging viruses which have the potential to cause pandemics and epidemics. Therapeutics and/or vaccines that target these pathogens are necessary to combat these viruses. Negative stranded RNA viruses possess a single stranded viral genome of negative polarity as their genome. The polymerase is packaged within the virion and is a rational target for drug design. Here, the polymerase complex of a segmented and unsegmented NSV was structurally and biochemically probed.

For over 100 years influenza viruses have caused pandemics. Implementation of the annual flu vaccine has significantly curbed the spread of the virus, but its effectiveness is only as good as the prediction of viruses that could be circulating during the season. Influenza viruses are highly mutagenic and consequently there is not a standalone vaccine that can confer immunity against each virus strain. Current influenza inhibitors on the market target the glycoproteins that are necessary for viral entry. However, due to the constant antigenic drift in the stalk proteins, these drugs are losing effectiveness. This suggests that we need to alter our strategy to target proteins within the virus which are unlikely to develop mutations which would render the antiviral useless.

In aim 1, the polymerase complex of influenza virus was identified as a viable target for design of influenza inhibitors. The PB2cap protein of the polymerase complex, utilizes a cap snatching mechanism to steal a m⁷-methyl cap from the host pre-mRNA to cap its viral mRNA. The structure of the PB2cap binding domain alone and in complex with a cap analog revealed that it does not make the same interactions as human capping proteins and thus is a viable target.

Furthermore, the structure of the cap analog of PB2cap from an influenza A and B virus indicated important interactions for drug design that will confer resistance to both strains.

Mumps virus has been largely controlled by the MMRII vaccine, however, sporadic outbreaks on university campus highlight the need for therapeutics that target the virus. The polymerase complex of mumps virus is a complex formed between the Large protein and the phosphoprotein. The phosphoprotein cofactor acts to home the polymerase onto the helical nucleocapsid for viral RNA synthesis. Due to the fact that the polymerase complex is not highly mutagenic, disrupting viral replication and transcription is a viable target for drug design.

Chapter 2 sought to assess the RNA sequester of the nucleocapsid. Using cryo-EM, the structure of the nucleocapsid-like particle lacking the C-terminal tail was solved to 10.4 Å. The structure of PIV5 nucleocapsid like particle was superimposed upon the cryo-EM structure to examine the regions of the nucleocapsid involved in RNA release. The alpha7-helix of the nucleocapsid protein was identified as the most likely region that allows RNA sequester.

The interactions involved in the uncoiling of the nucleocapsid like particle of MuV were evaluated in specific aim 3. An *in silico* model of MuV NLP and an amino terminal fragment of PNTD protein was generated. Residues L13 and I14 were identified as important for this interaction. Furthermore, the results suggest that L13 and I14 are particularly important due to their effects on thermal release.

The polymerase complex of NSVs is a viable target for drug development. This body of work unveiled the polymerase complex of influenza virus a sNSV and mumps virus a nsNSV. This work furthers the available information with regards to the viral proteins involved in viral replication. Furthermore, emergence of drug-resistant virus strains to polymerase complex inhibitors will be unlikely due to high sequence conservation in the proteins.

REFERENCES

1. Baltimore D. 1971. Expression of animal virus genomes. *Bacteriol Rev* 35:235–241.
2. Weller SK, Coen DM. 2012. Herpes simplex viruses: mechanisms of DNA replication. *Cold Spring Harb Perspect Biol* 4:a013011–a013011.
3. Zerboni L, Sen N, Oliver SL, Arvin AM. 2014. Molecular mechanisms of varicella zoster virus pathogenesis. *Nat Rev Microbiol* 2014/02/10. 12:197–210.
4. Gutsche I, Desfosses A, Effantin G, Ling WL, Haupt M, Ruigrok RWH, Sachse C, Schoehn G. 2015. Near-atomic cryo-EM structure of the helical measles virus nucleocapsid. *Science* (80-) 348:704–707.
5. Zhou H, Sun Y, Guo Y, Lou Z. 2013. Structural perspective on the formation of ribonucleoprotein complex in negative-sense single-stranded RNA viruses. *Trends Microbiol* 21:475–484.
6. Noda T, Sugita Y, Aoyama K, Hirase A, Kawakami E, Miyazawa A, Sagara H, Kawaoka Y. 2012. Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. *Nat Commun* 3:639.
7. Ruigrok RWH, Crépin T. 2010. Nucleoproteins of negative strand RNA viruses; RNA binding, Oligomerisation and binding to polymerase co-factor. *Viruses* 2:27–32.
8. Ortín J, Martín-Benito J. 2015. The RNA synthesis machinery of negative-stranded RNA viruses. *Virology* 479–480:532–544.
9. Reguera J, Gerlach P, Rosenthal M, Gaudon S, Coscia F, Günther S, Cusack S. 2016. Comparative Structural and Functional Analysis of Bunyavirus and Arenavirus Cap-Snatching Endonucleases. *PLOS Pathog* 12:e1005636.
10. Reguera J, Gerlach P, Cusack S. 2016. Towards a structural understanding of RNA

- synthesis by negative strand RNA viral polymerases. *Curr Opin Struct Biol* 36:75–84.
11. Reguera J, Cusack S, Kolakofsky D. 2014. Segmented negative strand RNA virus nucleoprotein structure. *Curr Opin Virol* 5:7–15.
 12. Morin B, Kranzusch PJ, Rahmeh AA, Whelan SPJ. 2013. The polymerase of negative-stranded RNA viruses. *Curr Opin Virol* 2013/04/18. 3:103–110.
 13. Fearn R, Plemper RK. 2017. Polymerases of paramyxoviruses and pneumoviruses. *Virus Res* 2017/01/16. 234:87–102.
 14. Johnson NPAS, Mueller J. 2002. Updating the Accounts: Global Mortality of the 1918-1920 “Spanish” Influenza Pandemic. *Bull Hist Med* 76:105–115.
 15. Jester B, Uyeki TM, Jernigan DB, Tumpey TM. 2019. Historical and clinical aspects of the 1918 H1N1 pandemic in the United States. *Virology* 527:32–37.
 16. Dawood FS, Iuliano AD, Reed C, Meltzer MI, Shay DK, Cheng P-Y, Bandaranayake D, Breiman RF, Brooks WA, Buchy P, Feikin DR, Fowler KB, Gordon A, Hien NT, Horby P, Huang QS, Katz MA, Krishnan A, Lal R, Montgomery JM, Mølbak K, Pebody R, Presanis AM, Razuri H, Steens A, Tinoco YO, Wallinga J, Yu H, Vong S, Bresee J, Widdowson M-A. 2012. Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. *Lancet Infect Dis* 12:687–695.
 17. Noda T, Sagara H, Yen A, Takada A, Kida H, Cheng RH, Kawaoka Y. 2006. Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature* 439:490–492.
 18. Shaw M, Palese P. 2007. Orthomyxoviridae: The viruses and their replication. *Fields Virol* 1647–1689.
 19. Tarendeau F, Boudet J, Guilligay D, Mas PJ, Bougault CM, Boulo S, Baudin F, Ruigrok

- RWH, Daigle N, Ellenberg J, Cusack S, Simorre J-P, Hart DJ. 2007. Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. *Nat Struct Mol Biol* 14:229–33.
20. Pflug A, Guilligay D, Reich S, Cusack S. 2014. Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* 2.
21. Guilligay D, Tarendeau F, Resa-Infante P, Coloma R, Crepin T, Sehr P, Lewis J, Ruigrok RWH, Ortin J, Hart DJ, Cusack S. 2008. The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat Struct Mol Biol* 15:500–6.
22. Einfeld AJ, Neumann G, Kawaoka Y. 2015. At the centre: influenza A virus ribonucleoproteins. *Nat Rev Microbiol* 2014/11/24. 13:28–41.
23. Sakaguchi A, Hirayama E, Hiraki A, Ishida Y o-ich., Kim J. 2003. Nuclear export of influenza viral ribonucleoprotein is temperature-dependently inhibited by dissociation of viral matrix protein. *Virology* 306:244–253.
24. Paterson D, Fodor E. 2012. Emerging roles for the influenza A virus nuclear export protein (NEP). *PLoS Pathog* 2012/12/06. 8:e1003019–e1003019.
25. Grohskopf LA, Alyanak E, Broder KR, Walter EB, Fry AM, Jernigan DB. 2019. Prevention and Control of Seasonal Influenza with Vaccines: Recommendations of the Advisory Committee on Immunization Practices — United States, 2019–20 Influenza Season. *MMWR Recomm Reports* 68:1–21.
26. Itzstein M von. 2007. The war against influenza: discovery and development of sialidase inhibitors. *Nat Rev Drug Discov* 6:967–974.
27. Davis AM, Chabolla BJ, Newcomb LL. 2014. Emerging antiviral resistant strains of Influenza A and the potential therapeutic targets within the viral ribonucleoprotein

- (vRNP) complex. *Virology* 11:167.
28. McKimm-Breschkin JL. 2013. Influenza neuraminidase inhibitors: antiviral action and mechanisms of resistance. *Influenza Other Respi Viruses* 7 Suppl 1:25–36.
 29. Collins PJ, Haire LF, Lin YP, Liu J, Russell RJ, Walker PA, Skehel JJ, Martin SR, Hay AJ, Gamblin SJ. 2008. Crystal structures of oseltamivir-resistant influenza virus neuraminidase mutants. *Nature* 453:1258–1261.
 30. Schnell JR, Chou JJ. 2008. Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* 451:591–595.
 31. Stouffer AL, Acharya R, Salom D, Levine AS, Di Costanzo L, Soto CS, Tereshko V, Nanda V, Stayrook S, DeGrado WF. 2008. Structural basis for the function and inhibition of an influenza virus proton channel. *Nature* 451:596–599.
 32. Pielak RM, Schnell JR, Chou JJ. 2009. Mechanism of drug inhibition and drug resistance of influenza A M2 channel. *Proc Natl Acad Sci U S A* 2009/04/21. 106:7379–7384.
 33. Omoto S, Speranzini V, Hashimoto T, Noshi T, Yamaguchi H, Kawai M, Kawaguchi K, Uehara T, Shishido T, Naito A, Cusack S. 2018. Characterization of influenza virus variants induced by treatment with the endonuclease inhibitor baloxavir marboxil. *Sci Rep* 8:9633.
 34. Rubin S, Eckhaus M, Rennick LJ, Bamford CGG, Duprex WP. 2015. Molecular biology, pathogenesis and pathology of mumps virus. *J Pathol* 235:242–252.
 35. Johnson CD, Goodpasture EW. 1934. AN INVESTIGATION OF THE ETIOLOGY OF MUMPS. *J Exp Med* 59:1–19.
 36. Lamb RA, Parks GD. 2007. Paramyxoviridae: The Viruses and their Replication, p. . *In* Knipe, D., Howley, P. (eds.), *Fields Virology*. Lippincott Williams & Wilkins.

37. Plumet S, Duprex WP, Gerlier D. 2005. Dynamics of Viral RNA Synthesis during Measles Virus Infection. *J Virol* 79:6900 LP – 6908.
38. Kolakofsky D. 2016. Paramyxovirus RNA synthesis, mRNA editing, and genome hexamer phase: A review. *Virology* 498:94–98.
39. Alayyoubi M, Leser GP, Kors C a., Lamb R a. 2015. Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein–RNA complex. *Proc Natl Acad Sci* 112:E1792–E1799.
40. Cox R, Pickar A, Qiu S, Tsao J, Rodenburg C, Dokland T, Elson A, He B, Luo M. 2014. Structural studies on the authentic mumps virus nucleocapsid showing uncoiling by the phosphoprotein. *Proc Natl Acad Sci* 111:15208–15213.
41. Klepser ME. 2014. Socioeconomic impact of seasonal (epidemic) influenza and the role of over-the-counter medicines. *Drugs* 74:1467–1479.
42. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, Griffin J, Baggaley RF, Jenkins HE, Lyons EJ, Jombart T, Hinsley WR, Grassly NC, Balloux F, Ghani AC, Ferguson NM, Rambaut A, Pybus OG, Lopez-Gatell H, Alpuche-Aranda CM, Chapela IB, Zavala EP, Guevara DME, Checchi F, Garcia E, Hugonnet S, Roth C. 2009. Pandemic Potential of a Strain of Influenza A (H1N1): Early Findings. *Science* (80-) 324:1557 LP – 1561.
43. Nitsch-Osuch A, Brydak LB. 2014. Influenza viruses resistant to neuraminidase inhibitors. *Acta Biochim Pol* 61:505–508.
44. Ruigrok RWH, Crépin T, Hart DJ, Cusack S. 2010. Towards an atomic resolution understanding of the influenza virus replication machinery. *Curr Opin Struct Biol* 20:104–113.

45. Liu Y, Qin K, Meng G, Zhang J, Zhou J, Zhao G, Luo M, Zheng X. 2013. Structural and functional characterization of K339T substitution identified in the PB2 subunit cap-binding pocket of influenza A virus. *J Biol Chem* 288:11013–23.
46. Tsurumura T, Qiu H, Yoshida T, Tsumori Y, Hatakeyama D, Kuzuhara T, Tsuge H. 2013. Conformational polymorphism of m7GTP in crystal structure of the PB2 middle domain from human influenza A virus. *PLoS One* 8:e82020.
47. Pautus S, Sehr P, Lewis J, Fortuné A, Wolkerstorfer A, Szolar O, Guilligay D, Lunardi T, Décout J-L, Cusack S. 2013. New 7-methylguanine derivatives targeting the influenza polymerase PB2 cap-binding domain. *J Med Chem* 56:8915–30.
48. Clark MP, Ledebner MW, Davies I, Byrn R a, Jones SM, Perola E, Tsai A, Jacobs M, Nti-Addae K, Bandarage UK, Boyd MJ, Bethiel RS, Court JJ, Deng H, Duffy JP, Dorsch W a, Farmer LJ, Gao H, Gu W, Jackson K, Jacobs DH, Kennedy JM, Ledford B, Liang J, Maltais F, Murcko M, Wang T, Wannamaker MW, Bennett HB, Leeman JR, McNeil C, Taylor WP, Memmott C, Jiang M, Rijnbrand R, Bral C, Germann U, Nezami A, Zhang Y, Salituro FG, Bennani YL, Charifson PS. 2014. Discovery of a Novel, First-in-Class, Orally Bioavailable Azaindole Inhibitor (VX-787) of Influenza PB2. *J Med Chem* 57:6668–78.
49. Byrn R a, Jones SM, Bennett HB, Bral C, Clark MP, Jacobs MD, Kwong AD, Ledebner MW, Leeman JR, McNeil CF, Murcko M a, Nezami A, Perola E, Rijnbrand R, Saxena K, Tsai AW, Zhou Y, Charifson PS. 2015. Preclinical Activity of VX-787, a First-in-Class, Orally Bioavailable Inhibitor of the Influenza Virus Polymerase PB2 Subunit. *Antimicrob Agents Chemother* 59:1569–82.
50. Liu Y, Yang Y, Fan J, He R, Luo M, Zheng X. 2015. The crystal structure of the PB2 cap-

- binding domain of influenza B virus reveals a novel cap recognition mechanism. *J Biol Chem* 2015/02/17. 290:9141–9149.
51. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 2011/03/18. 67:235–242.
 52. Smietanski M, Werner M, Purta E, Kaminska KH, Stepinski J, Darzynkiewicz E, Nowotny M, Bujnicki JM. 2014. Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. *Nat Commun* 5:3004.
 53. Papadopoulos E, Jenni S, Kabha E, Takrouri KJ, Yi T, Salvi N, Luna RE, Gavathiotis E, Mahalingam P, Arthanari H, Rodriguez-Mias R, Yefidoff-Freedman R, Aktas BH, Chorev M, Halperin JA, Wagner G. 2014. Structure of the eukaryotic translation initiation factor eIF4E in complex with 4EGI-1 reveals an allosteric mechanism for dissociating eIF4G. *Proc Natl Acad Sci* 111:E3187–E3195.
 54. Green A. 2014. West Africa struggles to contain Ebola outbreak. *Lancet* 383:1196.
 55. Shinde V, Bridges CB, Uyeki TM, Shu B, Balish A, Xu X, Lindstrom S, Gubareva L V, Deyde V, Garten RJ, Harris M, Gerber S, Vagasky S, Smith F, Pascoe N, Martin K, Dufficy D, Ritger K, Conover C, Quinlisk P, Klimov A, Bresee JS, Finelli L. 2009. Triple-Reassortant Swine Influenza A (H1) in Humans in the United States, 2005–2009. *N Engl J Med* 360:2616–2625.
 56. Centers for Disease Control and Prevention (CDC). 2006. Mumps outbreak at a summer camp--New York, 2005. *MMWR Morb Mortal Wkly Rep* 55:175–177.

57. Rota JS, Rosen JB, Doll MK, McNall RJ, McGrew M, Williams N, Lopareva EN, Barskey AE, Punsalang Jr A, Rota PA, Oleszko WR, Hickman CJ, Zimmerman CM, Bellini WJ. 2013. Comparison of the sensitivity of laboratory diagnostic methods from a well-characterized outbreak of mumps in New York city in 2009. *Clin Vaccine Immunol* 2013/01/16. 20:391–396.
58. Zipprich J, Winter K, Hacker J, Xia D, Watt J, Harriman K. 2015. Measles Outbreak - California, December 2014-February 2015. *MMWR Morb Mortal Wkly Rep* 64:153–154.
59. Luo M. 2011. *Negative Strand RNA Virus*. World Scientific, Singapore.
60. Ivanov I, Yabukarski F, Ruigrok RWH, Jamin M. 2011. Structural insights into the rhabdovirus transcription/replication complex. *Virus Res*. Elsevier B.V.
61. Green TJ, Zhang X, Wertz GW, Luo M. 2006. Structure of the Vesicular Stomatitis Virus Nucleoprotein-RNA Complex. *Science* (80-) 313:357 LP – 360.
62. Albertini AA V, Wernimont AK, Muziol T, Ravelli RBG, Clapier CR, Schoehn G, Weissenhorn W, Ruigrok RWH. 2006. Crystal Structure of the Rabies Virus Nucleoprotein-RNA Complex. *Science* (80-) 313:360 LP – 363.
63. Tawar RG, Duquerroy S, Vornrhein C, Varela PF, Damier-Piolle L, Castagné N, MacLellan K, Bedouelle H, Bricogne G, Bhella D, Eléouët J-F, Rey FA. 2009. Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. *Science* 326:1279–83.
64. Alayyoubi M, Leser GP, Kors CA, Lamb RA. 2015. Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein-RNA complex. *Proc Natl Acad Sci U S A* 112:E1792-9.
65. Green TJ, Cox R, Tsao J, Rowse M, Qiu S, Luo M. 2014. Common Mechanism for RNA

- Encapsidation by Negative-Strand RNA Viruses. *J Virol* 88:3766–3775.
66. Green TJ, Luo M. 2009. Structure of the vesicular stomatitis virus nucleocapsid in complex with the nucleocapsid-binding domain of the small polymerase cofactor, P. *Proc Natl Acad Sci U S A* 106:11713–11718.
 67. Leyrat C, Yabukarski F, Tarbouriech N, Ribeiro EA, Jensen MR, Blackledge M, Ruigrok RWH, Jamin M. 2011. Structure of the Vesicular Stomatitis Virus N0-P Complex. *PLoS Pathog* 7:e1002248.
 68. Cox R, Green TJ, Qiu S, Kang J, Tsao J, Prevelige PE, He B, Luo M. 2009. Characterization of a Mumps Virus Nucleocapsidlike Particle. *J Virol* 83:11402–11406.
 69. Vulliémoz D, Roux L. 2001. “Rule of Six”: How Does the Sendai Virus RNA Polymerase Keep Count? *J Virol* 75:4506 LP – 4518.
 70. Yabukarski F, Lawrence P, Tarbouriech N, Bourhis J-M, Delaforge E, Jensen MR, Ruigrok RWH, Blackledge M, Volchkov V, Jamin M. 2014. Structure of Nipah virus unassembled nucleoprotein in complex with its viral chaperone. *Nat Struct Mol Biol* 21:754–759.
 71. Delano WL. 2002. The PyMOL Molecular Graphics System. 1.3. Schrödinger LLC. Cambridge, MA.
 72. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera -- A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.
 73. Xu P, Li Z, Sun D, Lin Y, Wu J, Rota PA, He B. 2011. Rescue of wild-type mumps virus from a strain associated with recent outbreaks helps to define the role of the SH ORF in the pathogenesis of mumps virus. *Virology* 2011/06/14. 417:126–136.

74. Pickar A, Xu P, Elson A, Li Z, Zengel J, He B. 2014. Roles of serine and threonine residues of mumps virus P protein in viral transcription and replication. *J Virol* 88:4414–22.
75. Cox R, Green TJ, Purushotham S, Deivanayagam C, Bedwell GJ, Prevelige PE, Luo M. 2013. Structural and functional characterization of the mumps virus phosphoprotein. *J Virol* 87:7558–68.
76. Green TJ, Rowse M, Tsao J, Kang J, Ge P, Zhou ZH, Luo M. 2011. Access to RNA Encapsidated in the Nucleocapsid of Vesicular Stomatitis Virus. *J Virol* 85:2714–2722.
77. Zhang X, Green TJ, Tsao J, Qiu S, Luo M. 2008. Role of intermolecular interactions of vesicular stomatitis virus nucleoprotein in RNA encapsidation. *J Virol* 82:674–682.
78. Dong H, Li P, Böttcher B, Elliott RM, Dong C. 2013. Crystal structure of Schmallenberg orthobunyavirus nucleoprotein-RNA complex reveals a novel RNA sequestration mechanism. *RNA* 2013/06/24. 19:1129–1136.
79. Niu F, Shaw N, Wang YE, Jiao L, Ding W, Li X, Zhu P, Upur H, Ouyang S, Cheng G, Liu Z-J. 2013. Structure of the Leanyer orthobunyavirus nucleoprotein–RNA complex reveals unique architecture for RNA encapsidation. *Proc Natl Acad Sci* 110:9054–9059.
80. Reguera J, Malet H, Weber F, Cusack S. 2013. Structural basis for encapsidation of genomic RNA by La Crosse Orthobunyavirus nucleoprotein. *Proc Natl Acad Sci U S A* 110:7246–51.
81. Ariza A, Tanner SJ, Walter CT, Dent KC, Shepherd DA, Wu W, Matthews S V, Hiscox JA, Green TJ, Luo M, Elliott RM, Fooks AR, Ashcroft AE, Stonehouse NJ, Ranson NA, Barr JN, Edwards TA. 2013. Nucleocapsid protein structures from orthobunyaviruses reveal insight into ribonucleoprotein architecture and RNA polymerization. *Nucleic Acids*

- Res2013/04/17. 41:5912–5926.
82. Guryanov SG, Liljeroos L, Kasaragod P, Kajander T, Butcher SJ. 2015. Crystal Structure of the Measles Virus Nucleoprotein Core in complex with an N-terminal Region of Phosphoprotein. *J Virol* 90:JVI.02865-15.
 83. Aggarwal M, Leser GP, Kors CA, Lamb RA. 2018. Structure of the Paramyxovirus Parainfluenza Virus 5 Nucleoprotein in Complex with an Amino-Terminal Peptide of the Phosphoprotein. *J Virol* 92:e01304-17.
 84. Pickar A, Zengel J, Xu P, Li Z, He B. 2015. Mumps Virus Nucleoprotein Enhances Phosphorylation of the Phosphoprotein By Polo-Like Kinase 1. *J Virol* 90:JVI.02160-15.
 85. Alayyoubi M, Leser GP, Kors CA, Lamb RA. 2015. Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein–RNA complex. *Proc Natl Acad Sci* 112:E1792–E1799.
 86. Šali A, Blundell TL. 1993. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J Mol Biol* 234:779–815.
 87. Li T, Chen X, Garbutt KC, Zhou P, Zheng N. 2006. Structure of DDB1 in Complex with a Paramyxovirus V Protein: Viral Hijack of a Propeller Cluster in Ubiquitin Ligase. *Cell* 124:105–117.
 88. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291.
 89. van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastiris PL, Karaca E, Melquiond ASJ, van Dijk M, de Vries SJ, Bonvin AMJJ. 2016. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol* 428:720–725.

90. Gumper RH, Li W, Castañeda CH, Scuderi MJ, Bashkin JK, Luo M. 2018. A Polyamide Inhibits Replication of Vesicular Stomatitis Virus by Targeting RNA in the Nucleocapsid. *J Virol* 92:e00146-18.
91. Owusu D, Hand J, Tenforde MW, Feldstein LR, DaSilva J, Barnes J, Lee G, Tran J, Sokol T, Fry AM, Brammer L, Rolfes MA. 2020. Early Season Pediatric Influenza B/Victoria Virus Infections Associated with a Recently Emerged Virus Subclade — Louisiana, 2019. *MMWR Morb Mortal Wkly Rep* 69:40–43.