UNIVERSITY OF CALIFORNIA
RIVERSIDE

qFRET Technologies Towards the Development of Novel Influenza Virus Therapies

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Bioengineering

by

George S. Way

December 2020

Dissertation Committee:
Dr. Jiayu Liao, Chairperson
Dr. Victor Rodgers
Dr. John Jefferson Perry

The Dissertation of George S. Way is approved:

Committee Chairperson

University of California, Riverside

Acknowledgements

I would like to thank Dr. Jiayu Liao for his support and guidance throughout my Ph.D. study. I would also like to thank Dr. Victor Rodgers and Dr. John Jefferson Perry for their helpful advice and insight. I also am thankful for all the members in Dr. Jiayu Liao's group for their contributions and discussions that have helped further my studies. I would like to give a special thank you to the UCI GHTF for providing their expertise with sequencing technologies. Dr. Jiayu Liao supervised this research which laid the foundation for this dissertation.

Dedication

I would like to dedicate my thesis work to my friends, family, and wife Linda Nguyen.

ABSTRACT OF THE DISSERTATION

qFRET Technologies Towards the Development of Novel Influenza Virus Therapies

by

George S. Way

Doctor of Philosophy, Graduate Program in Bioengineering
University of California, Riverside, December 2020
Dr. Jiayu Liao, Chairperson

The influenza virus infects many people causing localized epidemics with the occasional pandemic. Influenza is known to have higher rates of death and morbidity in immunocompromised individuals, infants, and elderly people. Current methods to control the spread of the influenza virus are vaccinations and antiviral drugs. Vaccines are the most effective method to protect the public but can lack the necessary efficacy to protect the public every flu season. Antiviral drugs alleviate the burden of the flu, but the influenza virus mutates rapidly and has developed resistances to the current antiviral drugs. Genome-wide screenings have identified SUMOylation as an important host factor in the influenza virus lifecycle. Small Ubiquitin-like Modifier (SUMO) is a reversible post-translational modifier that uses a multi-step enzymatic cascade to alter protein function and stability. The dysregulation of SUMOylation has been known to associated with many diseases such as; carcinogenesis, neurodegenerative disease, and viral infections. Förster Resonance

Energy Transfer (FRET) is a non-radiative energy transfer that occurs between two fluorophores that are within 1-10 nm and have overlapping spectra. Our group has developed a quantitative FRET (qFRET) platform to measure protein interactions and enzyme kinetics. I have developed another application for our qFRET platform to identify the SUMOylated lysine residue (K131) of the non-structural protein 1 of the influenza A virus which is important for the influenza virus replication but not essential for the influenza virus. I have tested the efficacy of our SUMOylation inhibitor (STE) on the influenza A virus and found a novel SUMOylated lysine residue that is essential for the influenza A virus. Finally, I have discerned the acquisition of drug-resistant mutations by targeting host factors compared to viral proteins of the influenza A virus.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

FRET: Förster Resonance Energy Transfer

qFRET: quantitative Forster Resonance Energy Transfer

Ubls: Ubiquitin-like Modifiers

SUMO: Small Ubiquitin Like Modifier

SENP: Sentrin-specific protease

UBA2: Ubiquitin Activating Enzyme 2

UBC9: Ubiquitin Conjugating Enzyme 9

PIAS: Protein Inhibitor of Activated STAT

RANGAP: RanGTPase-Activating Protein

CyPet/YPet: Cyan/Yellow Fluorescent Protein

NS1: non-structural protein 1

HA: hemagglutinin

NA: neuraminidase

M1: Matrix Protein 1

PR8: A/Puerto Rico/8/1934 (H1N1)

IAV: Influenza A virus

IBV: Influenza B virus

Chapter 1: Introduction

FRET-based techniques can be used as powerful tools to study protein-protein interactions

FRET is a physical phenomenon that occurs between two fluorophores within 1 – 10 nm and have significant overlap between the emission spectrum of the donor and excitation spectrum of the acceptor. The energy transfer from the donor to the acceptor results in a quenching of the donor and the excitation of the acceptor (Figure 1.1). FRET efficiency is highly dependent on the distance between the donor and the acceptor (Figure 1.2). FRET is used extensively in biological research to elucidate protein interactions, gene expression, monitoring real-time intracellular signaling activities, and high-throughput screening of bioactive molecules[1–5]. In comparison to common techniques to study protein-protein interactions – co-immunoprecipitation and yeast two-hybrid – FRET-based methods can provide real-time monitoring in live cells and can be adapted for high-throughput screening. In FRET-based protein assays, proteins are tagged with fluorophores from a known FRET pair. The interaction between the two proteins leads to the fluorophores coming within close proximity, resulting in the energy transfer from the donor to the acceptor fluorophore. Any deviations in the protein-protein interactions by small molecules will be detected by observing an increase or a decrease in the FRET of the system.

Previous FRET-based assays took advantage of organic fluorophores to study proteins but at a cost of limiting the applications in functionality assays. Fluorescent proteins are widely used because they are easy to use and can label most proteins. Our

FRET-based assays use fluorescently fused recombinant proteins which enables a wider range of assays to study protein-protein interactions but fails to completely emulate the native environment of the proteins in the context of living cells. This research utilized an engineered FRET pair – CyPet and YPet – derived from the CFP and YFP proteins but with a higher fluorescence quantum yield and FRET efficiency[6]. There has been a push to develop FRET into a more quantitative method to determine kinetic parameters and protein interaction affinities[7,8].

The importance of Post-Translational Modifications

Ubiquitin is a relatively small protein (8.5 kDa) that utilizes an ATP-dependent enzymatic cascade resulting in the covalent attachment of ubiquitin to a target protein, which can ultimately degrade the target protein[9]. Since the discovery of ubiquitin, many ubiquitin-like modifiers (Ubls) have been discovered and they utilize enzymatic cascades that resemble ubiquitin. Small Ubiquitin-like Modifier (SUMO) is a reversible posttranslational protein modification that has more recently garnered the attention of the scientific community as an important protein modifier. SUMO proteins are composed of about 100 amino acids and undergoes a reversible covalent interaction to specific lysine residues in target proteins through a multi-step enzymatic cascade. Although the structural motifs of SUMO and ubiquitin are related, they only share an 18% identity, their attachment to target proteins have drastically different outcomes. SUMOylation is essential for many cellular processes in eukaryotes such as protein stability, cell cycle progression, gene expression, and nuclear-cytosolic trafficking[10]. Interestingly, SUMOylation and ubiquitination may also compete for the same lysine residues in target protein which can

alter the stability of the target proteins. Dysregulation and hijacking of SUMOylation has been linked to many human diseases such as neurodegenerative disease, carcinogenesis, and viral infections[11–15].

SUMOylation cascade

The process of covalently attaching a SUMO protein to a target substrate is termed SUMOylation (Figure 1.2). SUMOylation begins with the maturation of a SUMO precursor protein via SENP cleavage, unveiling a diglycine motif. After maturation, the SUMO protein will become activated through an ATP-dependent reaction whereby the active-site cysteine residue of the heterodimeric SUMO Activating Enzyme (E1) and the carboxy terminus of SUMO form a thioester bond. The thioester bond between the SUMO protein and E1 is then transferred to the E2, UBC9. The SUMO protein is then conjugated

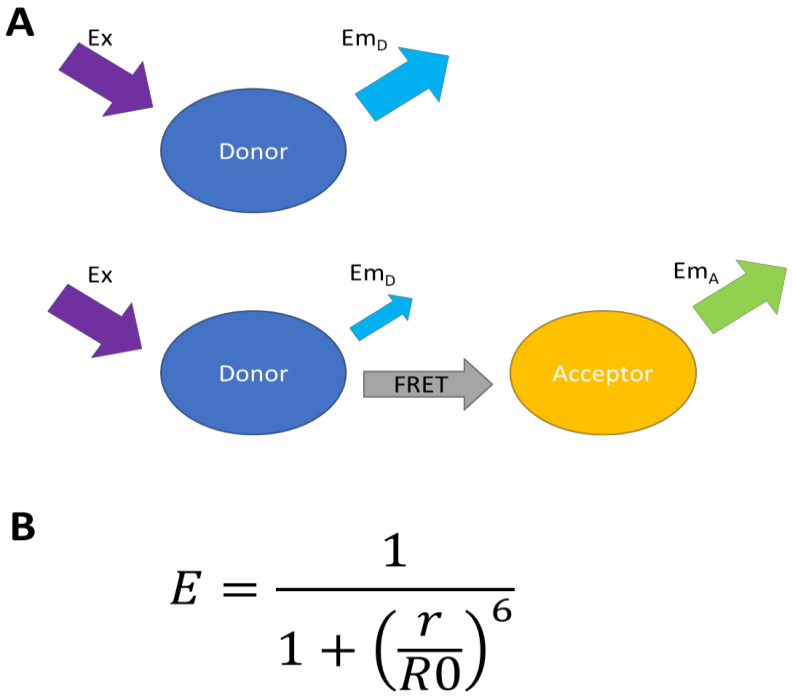


Figure 1.1. Principle of FRET. (A) General FRET diagram. (B) FRET efficiency (E) is dependent on the radial distance (r) between the donor and acceptor to the sixth power.

to a target substrate protein in the presence or absence of a SUMO E3 ligase, forming an isopeptide bond between a target lysine residue on the protein substrate and the carboxy terminus of the SUMO protein.

Currently, there are four known SUMO isoforms, only one SUMO E1 and SUMO E2, several SUMO E3 ligase families, and many substrates of SUMOylation. SUMO1 is a distinct isoform compared to SUMO2 and SUMO3 which share a 86% sequence identity with each other[16]. SUMO4 is not ubiquitously expressed in all tissues and only shares a 41% sequence identity with SUMO1 and an 85% sequence identity with SUMO2 and SUMO3. Each of these SUMO isoforms have varying affinities for different target proteins and have the ability to form a single SUMOylation event or polysumoylation. SUMO2 and

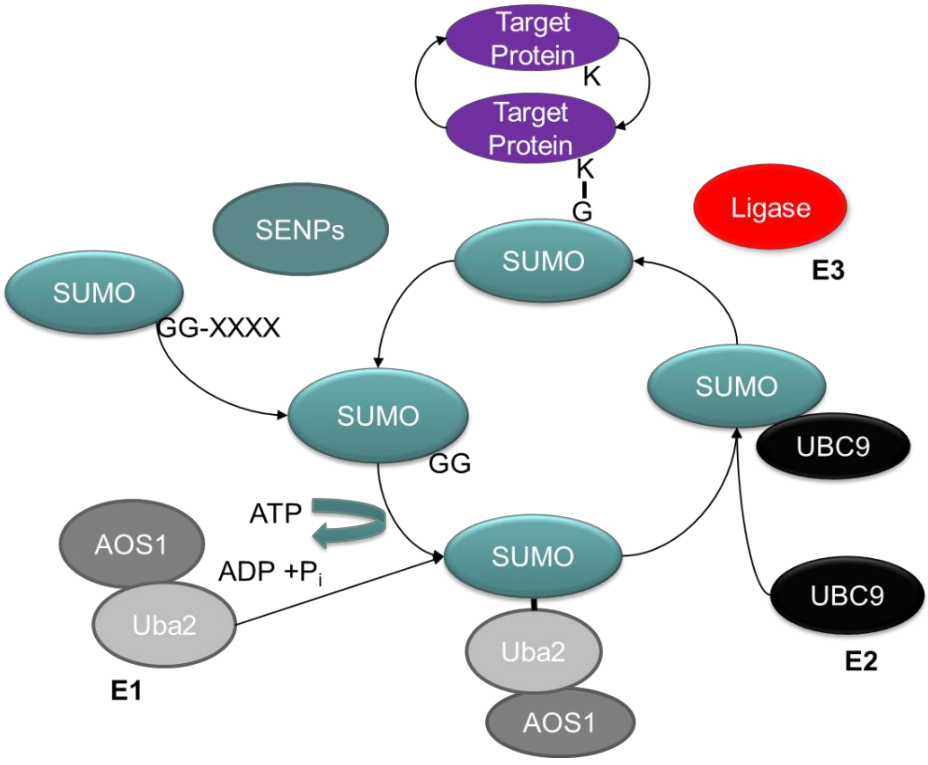


Figure 1.2. Overview of the SUMOylation cascade

SUMO3 have been noted to polysumoylate but SUMO1 lacks the needed lysine residue to polysumoylate[17,18]. Since there is only one E1 and E2, these are essential for SUMOylation to occur. The SUMO E3 ligases are the most diverse group of proteins in the cascade. Currently, there are more than 15 known SUMO E3 ligases which contribute to the specificity of the SUMO protein conjugation[19–28]. The biological importance of these differences allows for the SUMO proteins to achieve a higher specificity for its target protein and alter their function in a more controlled manner.

Influenza Virus

The influenza virus is responsible for 650,000 deaths around the world every year. The influenza virus is a segmented negative-sense single strand RNA virus consisting of eight segments encoding more than 12 proteins. There are currently 4 known types of influenza: A, B, C, and D. Influenza A and B viruses have a history of causing seasonal epidemics and occasional pandemics. Influenza C infections are mild and not known to have a high transmissibility between humans and influenza D is not known to infect people[29]. Humans hosts for these viruses are the primary concern for public health but the influenza A virus (IAV) is known to infect many different types of animals which poses a much greater public health threat than influenza B virus (IBV) which humans are its only known reservoir.

The IAV subtypes are characterized by the hemagglutinin (H) and neuraminidase (N) combination found on the surface of the virus. There are 18 hemagglutinin subtypes and 11 neuraminidase subtypes which presents a total of 198 different possible IAV subtypes. Currently, there are two main IAV subtypes that circulate every year: H1N1 and

H3N2. These subtypes can be further classified into clades and subclades which can be genetically different but not always antigenically distinct. Between the two dominant circulating strains, the H3N2 viruses have a higher propensity to undergo genetic changes which has resulted in the formation of many separate clades that co-circulate[30].

The IBV is categorized into two distinct lineages: Victoria and Yamagata. These lineages can be further classified into distinct clades and sub-clades, however, IBV has not been observed to change its genetics or antigenic properties nearly as rapid as IAV[31]. During each flu season, IAV and IBV circulate widely among human populations, with IAV being the predominate virus and IBV representing approximately 25% of total flu

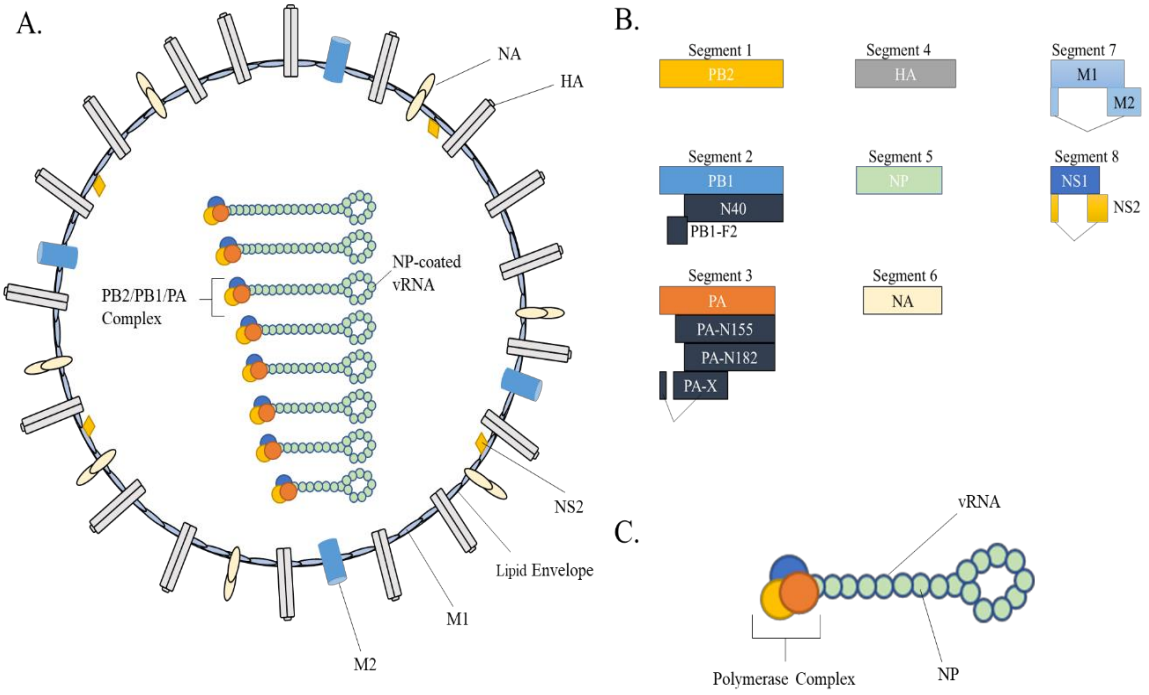


Figure 1.3. The anatomy of an influenza A virion. (A) Influenza A virion with all of the proteins and vRNP complexes. (B) The different segments and the proteins they encode. Segments 2, 3,7, and 8 encode for more than one protein. (C) The vRNP structure is composed of the (-)ssRNA wrapped around NP and capped with the RdRp complex.

cases. The disparity between the two shifts the focus of health organizations to target the IAV but the IBV still carries a significant risk to public health. The evolutionary mechanisms between IAV and IBV are still not well understood and are an ongoing area of research.

Influenza Virus Lifecycle

The influenza virion is a simple, roughly spherical, enveloped virus. The outer layer of the virus is a lipid membrane taken from its host cell. There are transmembrane glycoproteins inserted in the membrane- hemagglutinin (HA) and neuraminidase (NA)- and transmembrane ion channels (M2). Inside of the lipid membrane is a lining of matrix protein (M1) which helps provide rigidity and strength to the viral envelope. At the center of the virion are the eight negative-sense-single-strand RNA segments encoding the genes of the influenza virus and another protein- NS2 (Figure 1.3A). The strands are ordered based on the length, from longest to shortest; PB2, PB1, PA, HA, NP, NA, M, and NS (Figure 1.3B). PB2, PB1 and PA are the three polymerase proteins that make up the RNA-dependent RNA Polymerase (RdRp) complex (Figure 1.3C). NP is the nucleoprotein which the viral RNAs (vRNAs) are wrapped around, forming viral riboucleoproteins (vRNPs) which are anchored to the M1 proteins in the inner leaflet of the virion[32].

The influenza virus begins its journey as its HA proteins bind to sialylated receptors on the host cell's surface. Viruses from avian and equine origins recognize the $\alpha(2,3)$ linkages whereas human influenza viruses recognize the $\alpha(2,6)$ linkage from sialic acid[33]. Viruses from swine can recognize both types. After multiple binding events between the HA proteins and the sialylated receptors, the virus is engulfed in the membrane and

trafficked into the cell through an endosome. The acidic environment of the mature endosome causes two major changes in the viral proteins. The first major change caused by the acidic environment of the endosome shuttles protons through the M2 ion channels, causing the M1 shell to disassemble. The second major change is a result of the acidified environment which causes conformational changes in the HA subunits, resulting in the fusion of the viral membrane with the endocytic membrane, releasing the vRNPs and vRNAs into the cell[34,35].

Upon entry into the cytoplasm, the vRNPs need to enter the nucleus to facilitate the transcription and replication of the influenza virus' genetic material. All of the vRNPs have nuclear localization signals (NLSs) allowing the vRNPs to shuttle into the nucleus via interactions with the nuclear import proteins. After entering the nucleus, the negative sense viral genome must be converted into a positive sense. The RdRp initiates RNA synthesis on the vRNA without the need for primers due to the extreme 5' and 3' ends of the vRNA having partial complementarity to one another[36]. Interestingly, the newly synthesized mRNAs have poly-A tails but still lack another critical piece in the translational puzzle, the mRNA needs a 5' methylated cap. The PB2 protein has endonuclease activity which lead to the discovery of the "cap-snatching" mechanism where the RdRp complex cuts the nascent mRNAs 10-15 nucleotides downstream of the 5' methylated cap and uses the resulting RNA fragment to prime viral transcription[37–39].

The vRNPs are later exported out of the nucleus and are shuttled to the plasma membrane to form new viral particles. The influenza virus utilizes the host cell's plasma membrane to form viral particles since it is an enveloped virus. vRNPs are essential to form

a virion but all of the transmembrane viral proteins must be present to form a viral particle. Once the vRNPs, M1, and the transmembrane proteins have been trafficked to their correct locations the viral particle begins budding off. Before the newly formed viral particle can completely bud off of the plasma membrane, the NA protein cleaves the sialic acid residues from glycoproteins and glycolipids, finally allowing the viral particle to be released and infect a new cell[40].

Influenza A Virus Accessory Proteins

Since the first ten proteins of the influenza A virus have been described, more proteins have been discovered. These accessory proteins contribute to the virus' ability to adapt to new hosts, suppress host processes, regulate viral replication, and induce cell death.

The PB1 segment can encode two accessory proteins PB1-F2 and PB1 N40. PB1-F2 is encoded by an alternative ORF of the PB1 segment and can induce apoptosis, aggravating the inflammatory response, and promoting secondary bacterial infection[41–43] . PB1-N40 is an N-terminally truncated form of PB1 that lacks the polymerase activity but helps regulate virus replication in addition to PB1-F2[44].

The PA segment encodes up to three accessory proteins in addition to the PA protein. PA-X is caused by a frameshift in translation, PA-N155 and PA-N182 are N-terminally truncated versions of PA. PA-X has been shown decrease the pathogenicity of the 1918 H1N1 pandemic strain and a highly pathogenic H5N1 strain in mouse models[45,46]. PA-N155 and PA-N182 are important virulence factors demonstrated in a mouse study and can interact with PB1, translocating into the nucleus.

Since the discovery and description of the M1 and M2 proteins in the M segment, another M2-like ion channel protein has been discovered- M42. M42 functions like an M2 ion channel but with an antigenically distinct ectodomain. Although this protein is not essential for the virus, it has been found in a highly pathogenic H5N2 strain of avian influenza virus[47].

The NS segment is responsible for encoding three proteins- NS1, NS2 (NEP), and NS3. NS3 is the most recently discovered protein encoded by the NS segment, it is encoded by a novel splice site by a single nucleotide transition A374G. The presence of the NS3 protein has been linked to 33 influenza viruses that are associated with avian to mammalian transmission which links this protein with host adaptation[48].

Current Influenza Virus Treatments

Due to the severity of the influenza virus every year, governmental organizations and the World Health Organization (WHO) have spearheaded efforts to curb the impact influenza has on the world. Although vaccines and drugs have been developed to combat the influenza, there are still shortcomings that can be improved upon.

Vaccines are an efficient and cost-effective method to control the spread of the seasonal influenza virus outbreaks and potential pandemics. Current seasonal influenza vaccines efficacies can range from 40 – 60% according to the CDC. Traditionally, vaccines have been made from inactivated viruses, live attenuated viruses, or subunits from a virus. Trivalent inactivated viruses have been made with formulations containing three different types on influenza virus – H1N1, H3N2, and a B type (Yamagata or Victoria lineage). These vaccines were created from viruses cultured in embryonated chicken eggs which are

later inactivated as a whole or split with a detergent. Quadrivalent vaccines have been created to protect against H1N1, H3N2, and both IBV lineages which can provide a better range of efficacy than the older trivalent vaccines. Live attenuated virus vaccines are a relatively new FDA-approved vaccine technology which leverages temperature-sensitive influenza viruses which cannot tolerate temperatures higher than 35°C. This allows the for the attenuated viruses to replicate in the nasal passage but not the respiratory tract which can provide more time for the vaccine to stimulate the immune system[49].

In the event when vaccines do not protect the individuals from the population, anti-influenza drugs have been developed. In the late 1960s, the first class of FDA-approved drugs used to treat the influenza came to the market- amantadine and rimantadine (Table 1.1). These drugs targeted the transmembrane M2 ion channel protein which is essential for fusion of the viral and the endosome membrane, releasing the vRNPs and other internal protein into the cytoplasm of the infected cell[50]. Later, the FDA approved the use of neuraminidase inhibitors as a therapy for influenza infections with the discovery of oseltamivir and zanamivir[51–54]. Since then, two more neuraminidase inhibitors have been discovered and approved by regulatory agencies – zanamivir and lananimivir (Table 1.1). Neuraminidase inhibitors prevent the cleavage of the hemagglutinin bound to sialic acid, stopping the newly formed virions from budding off of the host cell. Recently, there has been a push for drugs that target much more conserved regions of the virus- the vRNP complex. Favipiravir, ribavirin, pimodivir, and baloxavir marboxil are novel influenza inhibitors which target the vRNP subunits. Currently, ribavirin and baloxavir marboxil are FDA-approved drugs in this group and pimodivir is currently undergoing phase II clinical

trials while favipiravir has yet to be approved by the FDA (Table 1.1)[55–57]. There are on-going efforts to create inhibitors for the HA protein. There are two main groups- small molecule inhibitor and antibodies. The small molecule inhibitors can be further broken down into two groups- one group which prevents the attachment of the HA protein to the cell surface and another group which prevents the maturation of the virus inside the cell. These drug candidates have shown efficacy in basic research and clinical trials, highlighting their therapeutic efficacies[58,59].

Although there have been extensive efforts to develop effective vaccines and drug therapies to treat influenza virus infections, pitfalls still exist. Vaccines are reliant on a strict influenza surveillance by the World Health Organization (WHO) and state-run health agencies to accurately predict influenza subtypes that may circulate in the next year, but vaccines are limited by the resources and technologies available. On the other hand, drug therapies have been demonstrated to be an effective method to combat the virus at the onset of clinical symptoms. However, the influenza is able to overcome these drugs with point mutations and widespread use of antiviral drugs has selected for drug-resistant strains which can be found circulating in different areas of the world as shown in Table 1.1[56,60–67]. These two methods to control and treat the virus are effective but they are still lacking the ability to effectively protect people from a broad-spectrum of influenza viruses.

Table 1.1. Anti-influenza drugs and mutations known to decrease the efficacy of drugs.

Anti-Influenza Drug	Target	Drug-Resistant Mutations
Amantadine	M2 ion channel	V27I, S31A/N
Rimantadine	M2 ion channel	L26I, S31A
Oseltamivir	Neuraminidase	E119V, N142S, R152K, D199N, I223T/R/V, S246N, S250G, H273Y, H274Y, H275Y, R292K, N294S, N295S, G402S
Zanamivir	Neuraminidase	E119V, Q136K/R, N142S, R152L/K, D198N, R292K
Peramivir	Neuraminidase	Q136K/R, D198N, I222R/V, S246N, H274Y, H275Y, R292K, H273Y, R152K
Lananimivir	Neuraminidase	G147E, K133E, E119G, D197E
Baloxavir marboxil	PA	I38T/M/F
Favipiravir	PB1	K229R
Ribavirin	PB1	D27N
Pimodivir	PB2	S324K/N/R, F325L, S337P, K376N/R, T378S, N510K

Chapter 2: Development of a highly sensitive FRET methodology to identify SUMOylation site(s) of influenza virus proteins and elucidate the significance of SUMOylation in the influenza virus lifecycle.

Abstract

SUMOylation is an important PTM for eukaryotic cells due to its role in the cell cycle, protein translocation, and protein stability. Nonstructural protein 1 of the influenza A virus is a major contributor to its virulence because it interferes with the host viral defense mechanisms. Most viruses do not come equipped with everything they need for their lifecycle and as a result, they have developed an extensive network of interactions with their host's cellular machinery to create progeny viruses. The SUMOylation of NS1 has been shown to regulate its activity. Using traditional biochemical approaches, two SUMOylated lysine residues have been identified. In this study, we developed a novel FRET assay to identify SUMOylation sites. We have demonstrated that the lysine residue K131 in the effector domain of NS1 is a novel SUMO acceptor site. A recombinant influenza A virus (H1N1/PR/8/34) expressing a NS1 K131A mutant had a significantly lower growth rate than the wild-type virus. These results suggest SUMOylation of NS1 on K131 is required for the rapid replication of H1N1 influenza viruses. The SUMO-NS1 interaction may serve as a novel target for drug development of anti-influenza A drugs. We have also identified novel SUMOylated lysine residues in the Matrix 1 protein of the influenza A virus which may have a critical role in the influenza A virus lifecycle.

Introduction

Influenza viruses have posed a major public health threat since the turn of the century with the 1918 Spanish influenza, since then we have experienced three more major influenza pandemics – 1957 (H2N2), 1968 (H3N2), and 2009 (H1N1). These pandemics have caused a significant loss of human life, economic burdens, and societal reformations. Annual influenza epidemics are responsible for 3 to 5 million severe cases with a death toll of 650,000 people annually according to the World Health Organization[68,69]. Localized epidemics are most commonly associated with human influenza A viruses. Aquatic fowl are the primary reservoirs of multiple strains of influenza viruses, some of which cross the species barriers resulting in a highly infectious disease in poultry and swine. These influenza viruses that cross the species boundaries are the main source of novel mammalian influenza A viruses, including the past pandemic strains with high rates of morbidity and mortality[70].

The influenza A virus is a negative-sense single-stranded RNA virus composed of eight segments encoding upwards of 12 proteins. Each protein has its own unique role in the influenza virus lifecycle ranging from binding to the cell surface, replicating viral RNA, and creating new viruses. These proteins can function on their own, but they also rely on certain host factors to properly function. One of the host factors that has shown to extensively interact with the influenza virus proteins is SUMOylation[71–75]. SUMOylation is a post-translational modification similar involved in many cellular processes such as gene expression, cell cycle progression, protein stability, and protein trafficking[76]. The relationship between SUMOylation and the influenza A virus have been documented but

the exact role SUMOylation plays in respect to the overall lifecycle of the influenza A virus is not completely understood. The nonstructural protein 1 (NS1) protein of the influenza A virus is a reported target of SUMOyation[77]. NS1 is involved with a variety of virus-host interactions, including antagonizing the interferon response during influenza infection[78,79]. Currently, there are conflicting results regarding the lysine residues in the NS1 protein that are SUMOylated. Xu et al. found lysine residues 219 and 221 were responsible for the SUMOylation of the NS1 protein, whereas Santos et al. identified K70 and K219 to be the lysine residues targeted by SUMOylation[73,75]. Moreover, Zhao et al. have documented K227 to be an important SUMOylation site in the NS1 protein but many strains of the influenza A virus have a C-terminal truncation and may lack the terminal lysine residues[80]. Furthermore, the matrix protein 1 (M1) has been shown to be SUMOylated at lysine 242 by CY Wu et al[74]. The M1 K242E mutant in this study attenuated the growth of the influenza A virus and caused malformed viral particles. However, there may be other SUMOylation sites on M1 that are more important for the influenza A virus.

Due to the different results regarding the exact lysine residues responsible for the SUMOylation of the NS1 protein, we revisited the identification of these sites in the NS1 of a common laboratory strain – A/Puerto Rico/8/1934 (H1N1) (PR8, GenBank accession number AAM75163). This NS1 contains the previously described SUMOylated lysine residues K70 and K219 but it lacks the K221 and K227 residues. To examine the SUMOylation of NS1 more precisely, we have developed a quantitative Forster resonance energy transfer (qFRET) biochemical approach to identify the lysine residue responsible for the SUMOylation of NS1[2,81]. We utilized SUMOylation site prediction tools based on

the NS1 peptide sequence to focus on specific lysine residues. We then conducted a site-directed mutagenesis on each lysine residue in NS1 and determined that K131, in the effector domain of the PR8 NS1, is the main target of SUMOylation[82–84]. Also, we developed a fluorescence-based influenza A virus assay and determined the growth rate of the PR8 mutant virus with a SUMOylation-deficient NS1 protein. These data suggest that NS1 SUMOylation promotes rapid replication of the influenza A virus.

Materials and Methods

Plasmid Constructs

The pET28b (+) constructs for CyPet-SUMO1, UBA2, AOS1, and UBC9 were cloned as outlined in our previous study[5]. The pET28b (+) YPet-Linker2 construct was made by amplifying the open reading frame (ORF) of YPet with primers containing NheI and Linker 2 (gtcacctctggttctccgggtctgcaggaatttggtacc) SalI and ligating the amplified ORF into a linearized pET28b (+) vector (Millipore Corporation Billerica, MA)[85]. After the sequence was verified, the ORF of PR8 NS1 was amplified via PCR with primers containing SalI (N-terminal) and NotI (C-terminal) and was ligated into the linearized pET28b (+) vector containing YPet-Linker 2. The mutagenesis of NS1 was performed via PCR with Phusion polymerase (New England Biolabs, Ipswich, MA) and tail-to-tail primers designed to introduce site-specific mutations and amplify the full plasmid[86]. The lysine-to-alanine NS1mutants were sequenced to verify the correct mutations were introduced. All plasmid DNA constructs were amplified in TOP10 DH5a *E. coli* cells.

The plasmids used to generate the recombinant influenza A virus were first described by Fodor et al.[87]. In this study, we used ambisense plasmids (pDZ) encoding the eight segments of the influenza PR8 virus. The plasmid containing the YPet-YPet insert in

the pDZ-HA was prepared by adding a NotI restriction site at both ends of the 3' (45 nucleotides of the HA gene coding sequence) and 5' (80 nucleotides of the HA gene coding sequence) as outlined by Marsh et al.[88]. The YPet-YPet insert with a Linker 2 between the two YPet ORFs was amplified by PCR with a NotI restriction site incorporated at both ends and cloned into the pDZ-HA packaging plasmid with the two packaging regions flanking it.

The pcDNA™3.1/Hygro(+) construct of the H1N1 HA gene was cloned by inserting the PCR-amplified HA gene from the pDZ plasmid into the pcDNA™3.1/Hygro(+) vector, with KpnI and NotI restriction enzyme sites on the gene's amino terminus and carboxy terminus, respectively.

Protein Expression and Purification

BL21 DE3 *E. coli* cells were transformed with the pET28b (+) constructs encoding CyPet-SUMO1, AOS1, UBA2, UBC9, YPet-Linker2-NS1, and YPet-Linker2-NS1 mutants. The transformed *E. coli* were plated onto LB agar plates containing 50 μg/mL kanamycin. Single colonies were inoculated into 5 mL liquid LB with 50 μg/mL kanamycin starter culture. Each starter culture was inoculated into 1 L of 2x YT medium with 50 μg/mL kanamycin and grown at 37°C, 180 RPM for 3 hours. Expression of recombinant proteins were induced with 0.6 mM IPTG at 25°C, 150 RPM overnight. The bacterial cells were harvested the next day at 4°C, 8,000 RPM. The bacterial cell pellet was resuspended in 30 mL of 20 mM Tris-HCl pH 7.4, 500 mM NaCl, and 4 mM imidazole. The cell suspension was lysed with an ultrasonic liquid processor (Misonix, Farmingdale, NY). Supernatant was collected after centrifugation at 4°C, 35,000 x g for 30 minutes.

Recombinant proteins were then bound to Ni^{2+} NTA agarose beads (QIAGEN, Valencia, CA). The column was washed sequentially with two column volumes of Wash Buffer 1 (20 mM Tris-HCl pH 7.4 and 300 mM NaCl), one column volume of Wash Buffer 2 (20 mM Tris-HCl pH 7.4, 0.1% TritonX-100, and 1.5 M NaCl), and two column volumes of Wash Buffer 3 (20 mM Tris-HCl pH 7.4, 500 mM NaCl, and 10 mM imidazole), and eluted with a buffer containing 20 mM Tris-HCl pH 7.4, 200 mM NaCl, and 300 mM imidazole. Recombinant proteins were dialyzed overnight at 4°C in a buffer containing 20 mM Tris-HCl pH 7.4, 50 mM NaCl, and 1 mM DTT. Protein purity was assessed with SDS-PAGE and Coomassie G-250 staining (Bio-Rad, Hayward, CA), and concentrations were determined with the Bradford assay with known amounts of bovine serum albumin (Thermo-Fisher Scientific Inc., Rockford, IL) as standards. Concentrations of fluorescent-fusion proteins were determined by fluorescence intensities measured on a FlexStationII[384] (Molecular Devices, Sunnyvale, CA).

Cell Lines

HEK293 and MDCK cells were cultured in DMEM (Gibco, Carlsbad, CA), supplemented with 10% fetal bovine serum (Gibco), penicillin-streptomycin (Gibco), and 2 mM L-Glutamine (Gibco).

***In silico* SUMOylation Site identification**

Three SUMOylation site prediction tools were queried with the amino acid sequence of NS1 protein from the influenza PR8 virus. GPS 1.0 SUMOylation prediction tool was used with a medium SUMOylation threshold for identification of potential SUMOylated lysine residues[84]. The NS1 amino acid sequence was also submitted to

SUMOplot™ (https://abgent.com/sumoplot)[75,80]. The NS1 sequence was also submitted to PCI-SUMO tool and lysine residues that overlapped with two or more site prediction tools were picked[83].

Generation of HA-MDCK cells

MDCK cells were cultured to 90% confluence and transfected with pcDNA3.1-HA using Lipofectamine 2000 (Invitrogen, Carlsbad, CA). Twenty-four hours after transfection, the cells were split into 100 mm dishes at low density and selected by the addition of 100 μg/mL hygromycin B (Invitrogen). After visible cell colonies were formed on the dishes, 48 cell colonies were picked up and cultured for HA expression screening by western blot. The stable cell line with the highest level of HA expression (HA-MDCK) was maintained in DMEM containing 100 μg/mL hygromycin B.

***In vitro* qFRET Assay**

To identify the SUMO sites of PR8 NS1, all components of the SUMOylation assay (1 mM CyPet-SUMO1, 50 nM E1, 100 nM E2, and 2 mM YPet-Linker2-NS1 or its mutants) were combined in a buffered solution containing 50 mM Tris-HCl pH 7.4, 1 mM DTT, and 4 mM MgCl2 in a total volume of 60 μL. The sample mixtures were incubated in a Greiner 384-well plate (Sigma-Aldrich) at 37 °C. After adding 1 mM ATP to the sample well, the fluorescence emissions over time were measured by using FlexstationII[384]. Emission intensities were measured at three wavelengths: 475 and 530 nm after excitation at 414 nm, and 530 nm after excitation at 475 nm[2].

Em_{FRET} Analysis

As described in the plasmid constructs section, we fused the CyPet and YPet to the amino termini of the SUMO1 and NS1, respectively. The peak wavelengths of excitation and emission are 414 nm / 475 nm for CyPet and 515 nm / 530 nm for YPet. When the FRET pair are in close proximity (between 1 – 10 nm), the excitation of the donor at 414 nm will result in an energy transfer from the donor to the acceptor, resulting in the quenching of the donor and excitation of the acceptor. Upon the SUMOylation of YPet-NS1 with a CyPet-SUMO1 FRET can occur, resulting in a 530 nm emission given a 414 nm excitation. However, if the mutant NS1 does not contain the target lysine amino acid, the 530 nm emission will have a distinctively lower peak than a NS1 with the target lysine amino acid (Figure 2.1A). In addition, anything that prevents SUMOylation – absence of SUMO E1, E2, or ATP – would also result in no increase of the emission at 530 nm.

The real FRET emission (Em_{FRET}) was used to monitor the formation of the SUMO1-NS1 complex. We defined Em_{FRET} as shown in Equation 1[2]. To obtain the real FRET emission, which is correlated with the amount of bound CyPet-SUMO1 and YPet-NS1, the direct emissions at 530 nm from free CyPet-SUMO1 and YPet-NS1 need to be determined and subtracted from the total emission intensity at 530 nm. To account for the contributions to the total emission at 530 nm, we used a previously established spectrum analysis for determining the Em_{FRET}. In short, the total fluorescent emission at 530 nm given a 414 nm excitation (Em_{total}) is differentiated into three fractions: real FRET emission (Em_{FRET}), CyPet direct emission, and YPet direct emission. The direct fluorescence contribution of the CyPet at 530 nm is proportional to its peak emission at 475 nm (FL_{DD})

when excited at 414 nm with a ratio coefficient of $\alpha = 0.378$. The direct emission of YPet at 530 nm is proportional to its emission at 530 nm given a 475 nm excitation (FL_{AA}) with a ratio coefficient of $\beta = 0.026$.

$$Em_{FRET} = Em_{total} - \alpha\, FL_{DD} - \beta\, FL_{AA} \qquad \text{(Equation 1)}$$

The Em_{FRET} was compared across all time points for each sample for a duration of 25 minutes. The amount of SUMOylated YPet-NS1 and YPet-NS1 mutants were also determined by western blot using an anti-NS1 antibody (Santa Cruz Biotechnology, Santa Cruz, CA).

Generation of Influenza A Viruses

The plasmid-based reverse genetic techniques to generate recombinant influenza viruses have been described previously[87,89,90]. To generate the wildtype virus, eight ambisense plasmids encoding the individual segments of the influenza A virus were mixed together in 1 μg quantities in 50 μL of serum-free DMEM per transfection. NS1 and M1 mutant viruses were generated with plasmids containing the mutations instead of the wildtype plasmid. 10 μL of lipofectamine 2000 was used per transfection in 250 μL of serum-free DMEM. The plasmid cocktail and lipofectamine were mixed and incubated at room temperature for 15 minutes before added to a mixture of MDCK and HEK293 cells. One 10 cm plate of HEK293 and another 10 cm plate of MDCK at 90% confluence were aspirated and washed with 5 mL of 1x PBS followed by resuspension with 0.25% Trypsin-EDTA. The cells were placed in separate 15 mL conical tubes and centrifuged at 800 xg for 5 minutes, the MDCK cells were resuspended with 3 mL of DMEM with 10% FBS and 1x Pen-Strep Glutamine (Supplemented DMEM). The 3 mL MDCK suspension was then

used to resuspend the HEK293 cells and 250 μL of the cell suspension was distributed among two 6-well tissue culture-treated plates. 1 mL of supplemented DMEM was added to each well containing cells and the transfection mixture was added to each well, viruses were generated in triplicate. Twenty-four hours after transfection, the transfection medium was replaced with 1x MEM containing 0.3% BSA, 1% Pen-Strep Glutamine, 0.1% FBS, 1 μg/mL tosylsulfonyl phenylalanyl chloromethyl ketone (TPCK)-trypsin for 48 hours to produce mature viral particles. The supernatant was then passaged to infect fresh MDCK cells at 90% confluence in 6-well plates. To generate PR8 influenza viruses with the YPet-YPet reporter, the pDZ-HA packaging plasmid with the YPet-YPet insert was used in lieu of pDZ HA for the eight plasmid transfection of HA-MDCK cells[91].

Growth Kinetics of Wildtype and NS1 mutant PR8 viruses in MDCK cells

PR8 influenza A viruses were reconstituted by plasmid DNA as described above. Growth rates of the NS1 wildtype and mutants were evaluated by infecting MDCK cells in triplicate with a multiplicity of infection (MOI) of 0.001. MDCK cells were incubated with diluted virus for one hour at 37°C and washed three times with PBS and 2 mL of Opti-MEM with 1 μg/mL TPCK-treated trypsin was added per well. The cells were incubated at 37°C. Samples were collected at 12, 24, 36, 48, and 60 hours after infection. The titers were determined by plaque assay in MDCK cells.

Fluorescent Virus Assay

PR8 viruses with the HA-YPet-YPet reporter were generated as described above with the exception of using the HA-YPet-YPet pDZ plasmid instead of the HA pDZ plasmid and using MDCK-HA cells for transfection and propagation. The growth rates of the recombinant PR8 viruses with the YPet-YPet reporter and the NS1 mutants were evaluated by infecting MDCK-HA cells in triplicate at a MOI of 0.001. The cells were incubated with the diluted virus for one hour at 37°C and washed three times with PBS, and 2 mL of DMEM with 0.3% BSA and 1 μg/mL TPCK-treated trypsin was added per well. The infected cells were harvested after 48 hours and the fluorescence of the cells was measured with an excitation of 475 nm and an emission of 530 nm.

Plaque Assay

Confluent MDCK cells in 6-well plates were infected with 10-fold dilutions of virus and incubated for one hour at 37°C. After one hour, the inoculum was removed and the cells were overlaid with 0.65% agar (Oxoid Ltd.) in MEM supplemented with 0.4% BSA, penicillin-streptomycin, 0.01% DEAE dextran, and 1 mg/mL TPCK-treated trypsin. The plates were incubated for 48 hours at 37°C; then the agar overlay was removed, and the cells were fixed with 3.7% PFA and stained with 20% methanol and 1% crystal violet.

Results

A FRET-based approach to identify SUMOylation sites

The SUMOylation of the NS1 protein of the influenza A virus has been proven to be a *bona fide* target of SUMOylation *in vitro* and *in vivo*[77]. We have designed and implemented an *in vitro* FRET-based SUMOylation assay to screen the influenza A virus proteome for SUMOylation targets. To establish our system, we have chosen the NS1 protein from the PR8 IAV to verify the accuracy. The FRET-based SUMOylation assay was performed by mixing CyPet-SUMO1, YPet-NS1, UBA2, AOS1, and UBC9 in a tris-buffered system. After the addition of ATP, we monitored the FRET signal in real-time, showcasing the formation of the SUMOylated NS1 product as shown in Figure 2.1A. Our real FRET emission is specific to the formation of the SUMOylated product and is not contaminated by the presence of the FRET pair in the well as shown in Figure 2.1B. Western blotting with anti-NS1 verifies our real FRET emission is produced by the formation of the SUMOylated product (Figure 2.1C).

Previously reported SUMOylation sites of NS1 are not essential for SUMOylation

After confirming the influenza, A PR8 virus' NS1 protein is a SUMOylation substrate *in vitro*, we mutated previously published reported SUMOylated lysine residues to alanine, specifically K70 and K219[75]. The amino acid sequence of the IAV PR8 strain was submitted to three different SUMO site prediction tools (GPS-SUMO 1.0, SUMOplot™, and PCI-SUMO) and we chose the overlapping lysine residues- 70, 175, 219 (Figure 2.2A)[82–84]. However, our FRET results using the three different mutations

K70A, K175A, K219A did not result in reduced SUMOylation compared with the wildtype NS1 (Figure 2.2B). Creating a NS1 mutant with all three lysine-to-alanine mutations still did not abolish the SUMOylation of NS1 (Figure 2.2C).

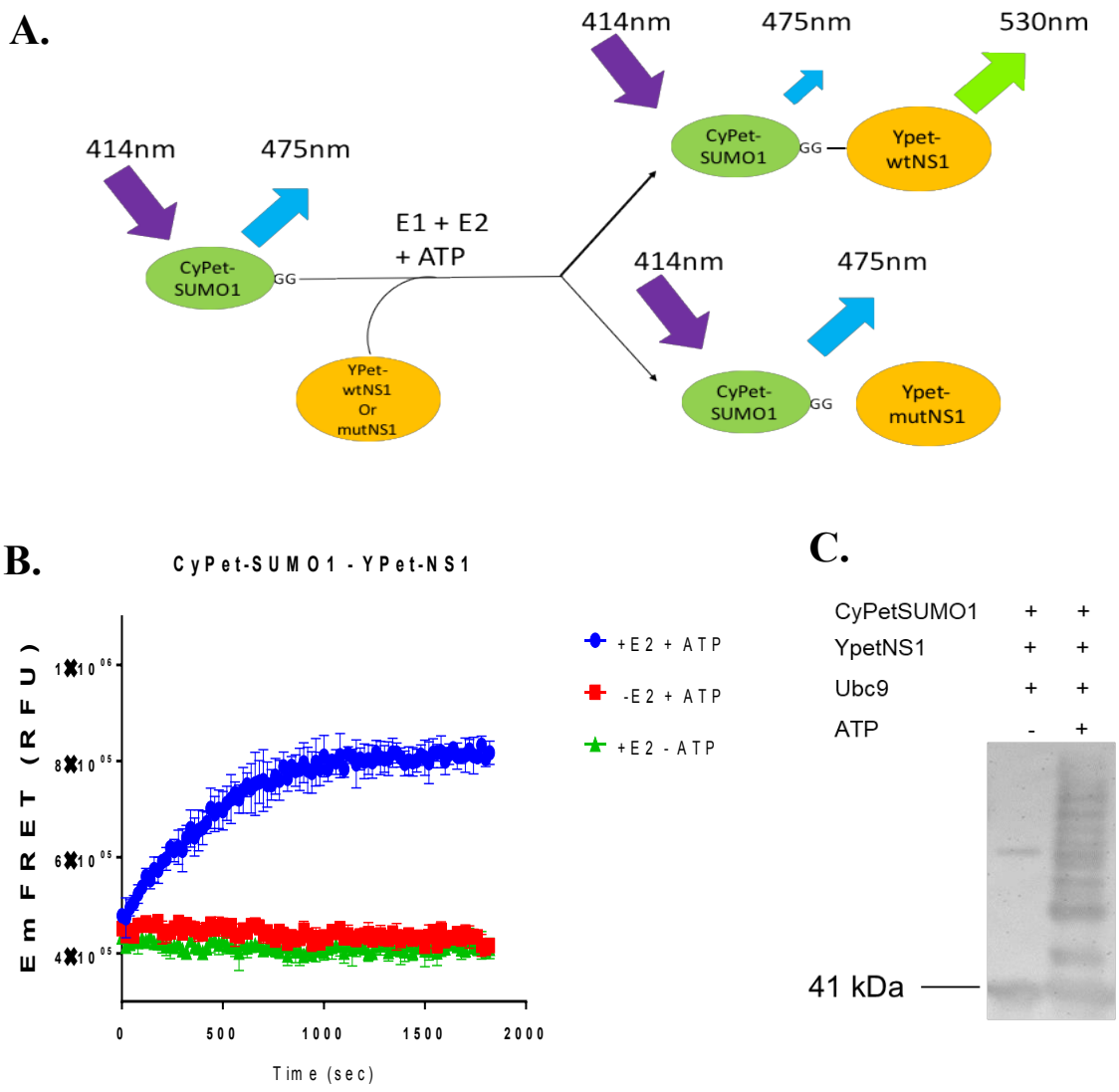
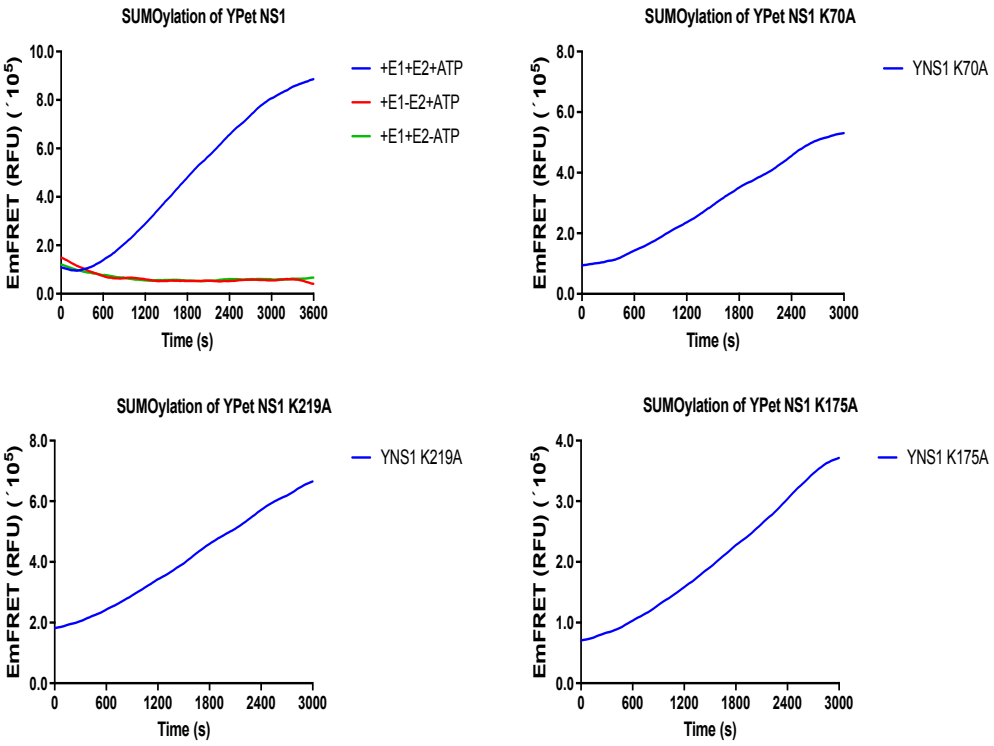


Figure 2.1. Development of an *in vitro* FRET-based SUMOylation assay of IAV PR8 NS1. A) The FRET-based SUMOylation assay of NS1 is specific to the SUMOylation of NS1. A mutant that cannot be SUMOylated will not have an increase in the Em_{FRET} B) The Em_{FRET} signal is specific to the SUMOylation of NS1, with the presence of SAE1, E2, and ATP. C) Western blot of SUMOylation of NS1 in the presence or absence of ATP.

A.

GPS-SUMO1.0	SUMOplot™	PCI-SUMO
K 70	K 70	K 78
K 175	K 219	K 108
		K 131
		K 175
		K 219

B.



C.

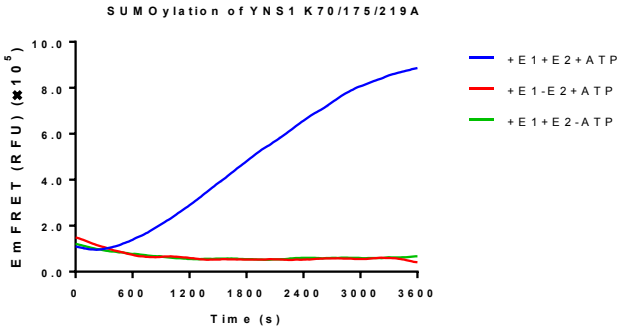


Figure 2.2. Validation of SUMO site prediction tools. A) Predicted SUMOylated lysine residues from GPS-SUMO 1.0, SUMOplot™, and PCI-SUMO. B) SUMOylation of NS1 is not dependent on lysine residues K70, 175, or 219. C) SUMOylation can still occur in a mutant containing all three mutations.

Systematic screening of lysine residues of NS1 for SUMOylation with FRET assay

We conducted a systematic mutagenesis on all possible SUMOylation sites by mutating lysine residues in a stepwise fashion until we had a lysine-deficient mutant (Figure 2.3A). Each mutant was then expressed, purified, and run in an *in vitro* SUMOylation assay (Figure 2.3B). NS1 mutant #7 was able to be SUMOylated, however, NS1 mutant #8 – the first mutant with K131A – did not lead to an increase in the Em_{FRET} over time. Similar results were shown with the other two mutants containing K131A.

Confirming K131 as the essential SUMOylation site in NS1

To confirm K131 is the SUMOylated lysine residue, I created an NS1 mutant with only the K131A mutation. Compared to the wild type, it had a significant loss of SUMOylation as shown in Figure 2.4A and 2.4B. Parallel reactions were set up and were used for a western blot. The western blots were probed with anti-NS1 and confirmed that K131A is not SUMOylated while the wildtype NS1 and NS1 K70/219A are SUMOylated shown in Figure 2.4C.

A.

NS1 Mutant	Lysine Residues mutated to Alanine
1	70, 175, 219
2	70, 108, 175, 217
3	41, 70, 108, 175, 217
4	41, 62, 70, 126, 108, 175, 217
5	41, 62, 78, 126, 108, 175, 217
6	20, 41, 62, 78, 126, 108, 175, 217
7	20, 41, 62, 78, 108, 110, 126, 175, 217
8	20, 41, 62, 78, 108, 110, 126, 131, 175, 217
9	20, 41, 62, 78, 108, 110, 126, 131, 175, 217, 219
10	20, 41, 62, 70, 78, 108, 110, 126, 131, 175, 217, 219

B.

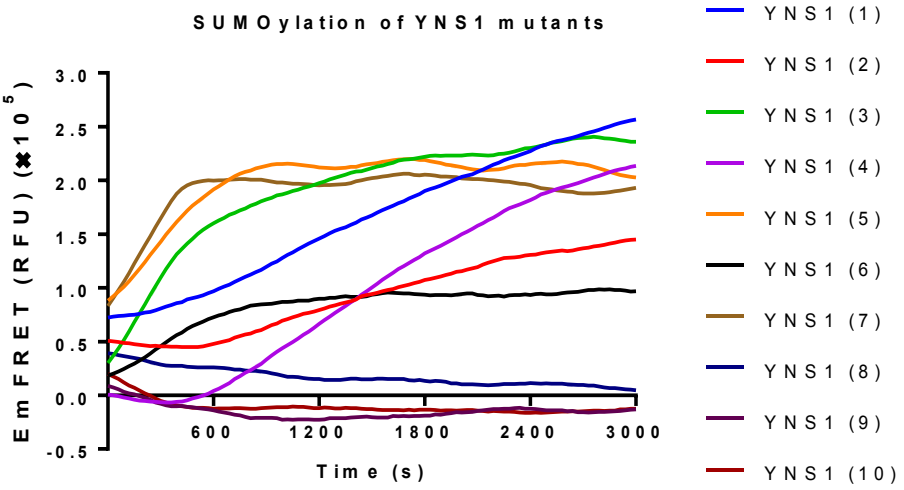


Figure 2.3. Systematic mutation of lysine residues. A) Mutations added to the NS1 protein. B) Em_{FRET} of NS1 mutants over time, Em_{FRET} does not increase over time with the addition of the K131A mutation.

Effects of the NS1 K131A mutant on virus replication

The NS1 protein of influenza A has many diverse functions in the scope of the virus lifecycle, including inhibition of the host-IFN stimulated immune response[75]. Knocking down Ubc9 expression in host cells reduces the replication of efficiency of the influenza A virus which is indicative of the importance of SUMOylation for influenza A virus production[74]. To evaluate the importance of the SUMOylation of NS1 in the context of the PR8 influenza virus developed a fluorescent reporter strategy to quantify the extent of viral replication. For our approach, we incorporated a tandem YPet-YPet gene in the middle of the HA gene but kept the packaging regions both ends to ensure the incorporation of the recombinant segment into progeny virions (Figure 2.5A and 2.5B)[88]. The expression of the YPet-YPet reporter correlates with viral replication in MDCK-HA expressing cells. Comparing the fluorescent emissions from the wildtype and the two mutants with a two-tailed student's t-test ($\alpha = 0.05$), there was no significant difference in the fluorescent emission between the wildtype NS1 and NS1 K70/219A ($p = 0.136$) but there was a significant difference between wildtype NS1 and NS1 K131A ($p = 0.033$) (Figure 2.5C). These data highlight the importance of K131A in the influenza A virus lifecycle.

To verify that NS1 K131A mutation is important for influenza virus replication, we generated wildtype PR8 and two NS1 mutant viruses (PR8NS1-70A and PR8NS1-131A) and compared the multi-growth properties in MDCK cells when infected with a low MOI (0.001). All three viruses were able to replicate; however, the viral titers in PR8NS1-131A infected cells were significantly lower than the wildtype and the PR8NS1-70A infected

cells at 36, 48, and 60 hours after infection (Figure 2.5D). These results validated that K131 in NS1 is an important factor in PR8 replication. A viral plaque assay also confirmed that NS1 K131A mutation attenuated the viruses' ability to form plaques (Figure 2.5E).

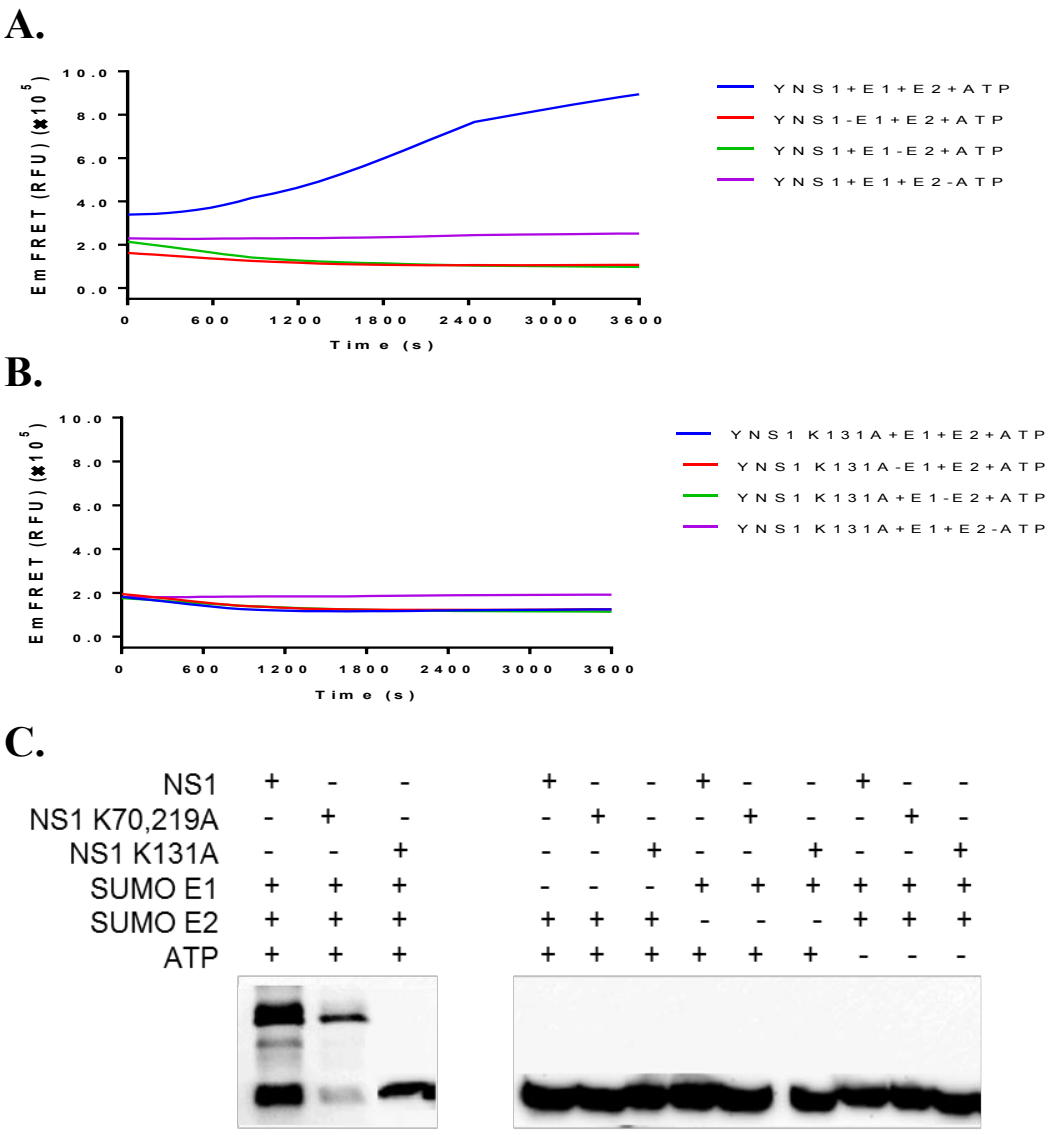


Figure 2.4. Confirmation of SUMOylation site. A) FRET assay of wildtype NS1 is specific to the SUMOylation of NS1. B) SUMOylation of the NS1 K131A mutant is not detected on the FRET assay. C) SUMOylation of NS1 and NS1 mutants in the presence or absence of E1/E2/ATP. Western blots were probed with anti-NS1.

SUMOylation of M1 is critical for virus generation and replication

The matrix protein 1 (M1) has an essential role in the assembly of the influenza virus. It has been shown to interact with vRNPs to create progeny virions[92,93]. The virus coordinates many protein-protein interactions to stitch the different parts of the virus together to create new viruses. SUMOylation has been shown to be an important factor for the translocation of M1 which directly affects the morphology of the virions formed, attenuating the virus. We have tested a previously identified SUMOylation site – K242 – and have found it does not prevent the virus from being able to infect and replicate at similar titers to wildtype[74] (Figure2.6.A and B). Vipul Madahar ran tandem liquid chromatography and mass spectrometry (LC/MS) on *in vitro* SUMOylation reactions of M1 to identify SUMOylated lysine residues. From our initial run, we have identified six possible SUMOylated lysine residues: K21, K35, K187, K230, K242, and K252. We then conducted a mutagenesis to demonstrate the role of SUMOylating M1 in the context of the IAV lifecycle. We tried to generate an IAV M1 mutant with the K21/35/187/230/242/252R (6x K>R) mutations and could not generate any viruses after three independent triplicate transfections (Figure 2.6.C).

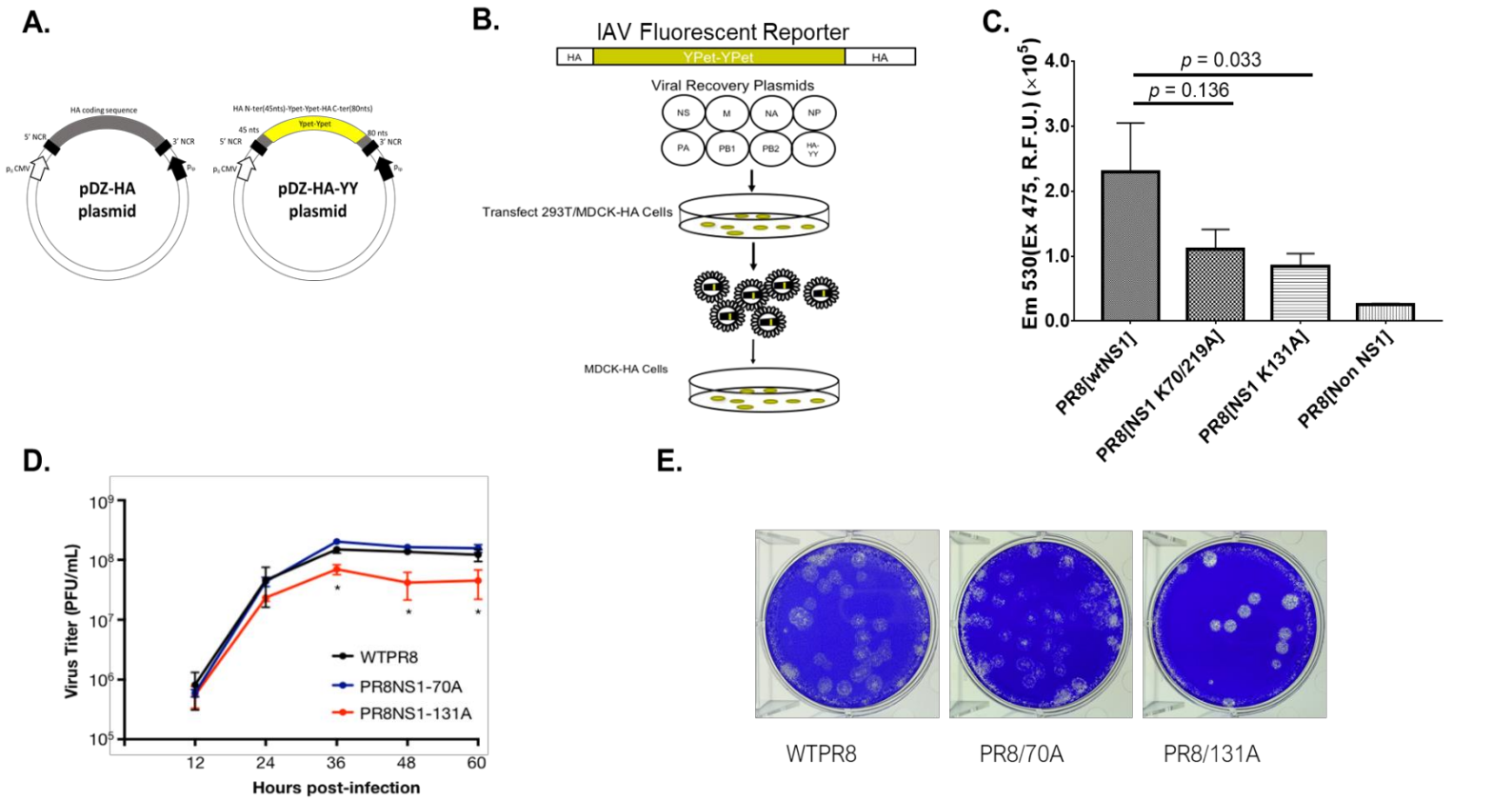


Figure 2.5. Evaluating the impact of a mutant NS1 on the influenza A virus. A) Plasmid design with the 2x YPet gene incorporated in the middle of the HA gene in the pDZ plasmid. B) Generating influenza viruses with the 2xYPet reporter gene. C) Fluorescence assay of HA-Ypet-Ypet with NS1 mutants. Wildtype PR8 and NS1 K70/219A do not have a significant difference in fluorescent emission (p = 0.136), while wildtype PR8 and NS1 K131A have a significant difference in the fluorescent emission using a two-tailed student's t-test (p = 0.033). D) comparison of multicycle growth of wildtype PR8, PR8 NS1 70A, and PR8 NS1 131A. D) Plaques were determined by infecting cells with low MOI (0.001) and were determined at 36, 48, and 60h post-infection. Error bars indicate the standard deviation of the triplicates.

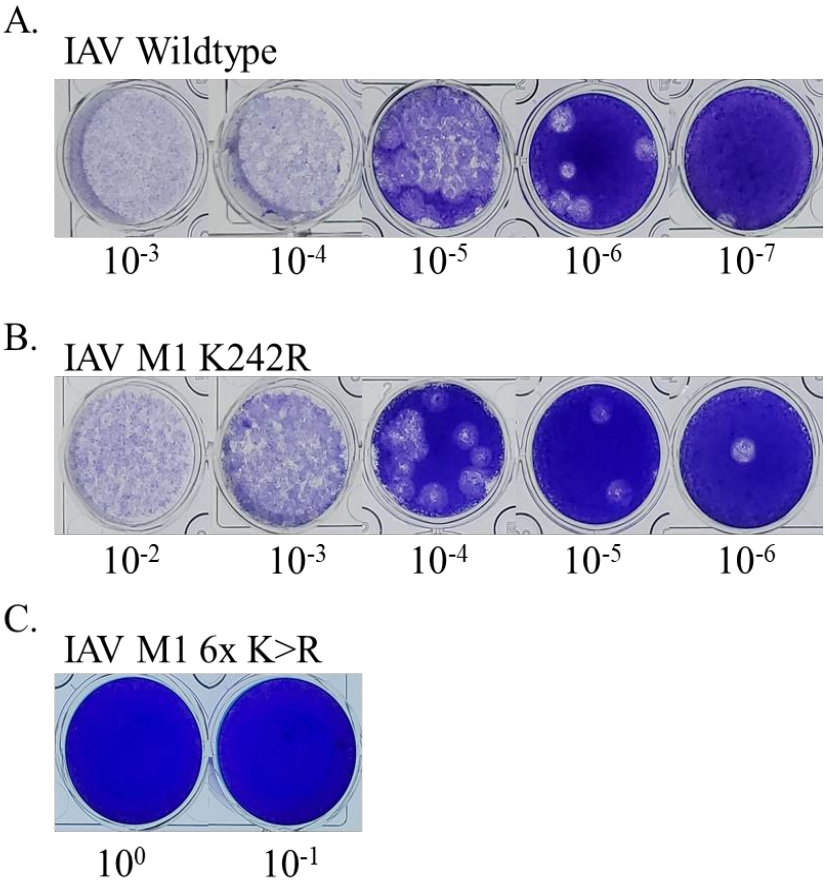


Figure 2.6. Plaquing results from influenza A viruses. A) Titration of IAV Wildtype starting from 10^{-3} to 10^{-7}. B) Titration of IAV M1 K242R starting from 10^{-2} to 10^{-6}. C) Titration of IAV M1 6x K>R at undiluted stock and ten-fold diluted stock.

Discussion

Many contributions have been made in the last century towards the understanding of the influenza A virus. Despite the accumulation of knowledge to understand the role each protein plays in the influenza lifecycle, understanding the delicate dance the virus plays with the host requires further study. SUMOylation has been reported as an important host process responsible for promoting viral replication[13,94]. The SUMOylation of several influenza virus proteins has been reported, including NS1- a critical virulence factor that

promotes influenza virus replication through multiple mechanisms, including the IFN-stimulate response[95]. Pal et al. was the first group to identify the influenza A virus NS1 as a target for SUMOylation which supported the idea that SUMOylation is likely to be a key regulatory factor for virus replication[77]. Xu et al. confirmed the SUMOylation of NS1 and further mapped the SUMOylated sites to the C-terminal portion of a H5N1 influenza A virus isolate. They further demonstrated the stability SUMOylation provided for NS1 which also led to increases in viral titers[73]. Santos et al. identified the main SUMOylation sites in the NS1 protein of the PR8 virus to be K70 and K219 and demonstrated that a SUMOylation deficient NS1 had a lower ability to antagonize the cellular IFN response[75]. Interestingly, the NS1 protein of the H1N1 pandemic strain contained a relatively rare SUMOylation site at K227 which promotes the ability of NS1 to partition in RNA granules, facilitating NS1-mediated inhibition of host gene expression[80]. In this study, we have revisited the identification of PR8 NS1 SUMOylation sites by using a novel FRET-based SUMOylation assay and found, in contrast to previous results, K70 and K219 did not have a significant role in the SUMOylation of NS1. Therefore, we systematically mutated each lysine residue in PR8 NS1 and mapped the SUMO acceptor site to K131. The discrepancy in these data may be attributed to a combination of techniques used to determine the SUMOylation site.

Previously, a fluorescent reporter strategy was developed for use in an influenza virus assay and we adapted this strategy to create a more sensitive fluorescent reporter to measure the effect of site-specific mutations on virus replication[88]. We utilized this methodology to highlight the significance of the K131 in NS1 on virus growth. This was

confirmed through the comparison of the multi-growth cycles of the wildtype, NS1 K70A, and NS1 K131A in MDCK cells. K131A was the only mutation that significantly reduced the multicycle replication of PR8 virus in MDCK cells.

FRET is used in many biological applications due to its high sensitivity and specificity. In this study, we have developed a FRET-based approach to determine SUMOylated lysine residues in an influenza A virus protein. There are several clear-cut advantages to using this approach over traditional methods. First, the FRET signal is proportionally to the number of molecular interaction events which can allow for the quantification of interacting partners. Previous methods have used co-immunoprecipitation techniques to determine the SUMOylation site which has a limit of detection based on the abundance of the protein as well as the affinity of the antibodies being used. Depending on how the antibodies were made, the epitope may become unavailable after PTMs are bound to the protein and the antibodies can have different affinities for the target epitope which can cause false-positive results. Also, the real FRET emission can be monitored in real-time, which can provide information on kinetics which can be compared between mutants. This method can also be used with crude protein extracts and fluorescent proteins can be added through simple molecular subcloning. Mass spectrometry has also been used to determine SUMOylation sites, but this approach requires the use of mutant SUMO proteins for simplified analysis and analysis requires more time while not being able to collect real-time monitoring of SUMOylation[96].

Viral proteins can engage with the host's SUMOylation cascade to enhance viral replication or evade host antiviral responses. SUMOylation has been documented to

positively regulate wildtype IAV infection as shown with a global increase in SUMOylation events during IAV infection and a decrease of the influenza virus' rate of replication after host SUMO E2 knockdown[71,74,75]. Our results demonstrate that a SUMOylation deficient NS1 PR8 virus has a lower growth rate than a PR8 wildtype virus.

There are clear benefits of SUMOylating the NS1 protein for the virus. The SUMOylation of NS1 enhances its stability and its ability to partition RNA granules helping the virus subvert the host gene expression[73,80]. Further studies can be done to determine whether SUMOylation of NS1 at the K131 residue is involved in the stability and/or function of the NS1 protein. We also cannot exclude that the K131 residue is functionally involved in NS1 function independent of SUMOylation. Moreover, since K131 is not completely conserved among all influenza A virus strains, it is conceivable that NS1 polymorphisms may compensate for a lack of the K131 residue or other lysine residues are SUMOylated in strains lacking a K131 residue. For example, binding of NS1 to CPSF30 inhibits antiviral gene expression, in strains of influenza lacking this function, the loss is compensated by mutations in other viral proteins[97].

SUMOylation is believed to be a critical factor for efficient influenza virus growth and infection, the direct importance of SUMOylation for influenza infection still needs to be studied more in depth. We have reported that the K131 residue of the PR8 NS1 protein is important for influenza virus replication, but it is not essential. We expect to find SUMOylation sites in other IAV proteins and expand upon the understand the role SUMOylation plays in the function of those proteins. In addition, our identification novel

SUMOylated lysine residues in the M1 protein are essential for the influenza virus lifecycle. Our FRET-based assay can be easily converted into a high-throughput format to identify other SUMOylation sites in virus and host proteins during viral infection.

Given the role SUMOylation plays in efficient influenza infection, the development of SUMOylation inhibitors can be invaluable tools to aid in the development of novel antiviral drugs to effectively treat a broad spectrum of strains. Efforts in this direction can also contribute to our knowledge of the functional interactions between SUMOylation and viral proteins.

Chapter 3: Evaluate SUMOylation as a potential target for influenza virus drug development

Abstract

The influenza A virus has afflicted the world for the last century with recurring epidemics and the occasional pandemic. The world's efforts to combat this virus has resulted in the development of vaccines as well as drugs. Vaccines rely on strict surveillance data to accurately predict the viruses that will be circulating during the next influenza season. Current antiviral drugs target viral proteins, but the influenza is able to develop drug resistant mutations due to its error prone replication. There is a need for a new class of treatments to alleviate the burden of the influenza on society, however there has not been a strong push to change the current rationale for drug design. Here, we propose to target host factors that the influenza A virus hijacks throughout its lifecycle. SUMOylation is a ubiquitin-like modification which is responsible for a wide-variety of processes inside the cell and a host factor exploited by the influenza virus. Our findings determined an essential lysine residue for the influenza A virus in the M1 protein. These findings validate our strategy to target host factors the influenza virus hijacks as an antiviral therapy.

Introduction

The influenza A virus circulates around the world in seasonal epidemics with the occasional pandemic. Over the last century, there have been four major pandemics caused by influenza. Aquatic fowl are the primary reservoirs for the influenza A virus throughout

the year. The influenza A virus is bound to different species by the recognition of the sialic acid α-linkage on the surface receptors. Avian IAV recognize the α2-3 linkage and human IAV recognize α2-6 linkages[98]. Although these differences prevent most viruses from crossing the species barrier, some animals contain both α2-3 and α2-6 linkages which can give rise to novel viruses that can be more virulent than both parental viruses that reassorted within the host like the H1N1 swine flu pandemic in 2009[99].

Public health organizations and governments have supported many different efforts to prevent the spread of the IAV, IBV, and many other viruses. Vaccines have been utilized as a primary control method to prevent the spread of seasonal epidemics from growing into global pandemics. Vaccination is a proven method to cull the spread of disease but vaccines for the influenza require extensive collaborative efforts among surveillance sites to accurately predict which strains will be circulating during the upcoming flu season. Anti-influenza drugs have also been developed to help stop the spread of the virus and save lives. However, the virus can develop resistances to the anti-influenza drugs currently approved by health organizations. This arms race between drug research and nature needs to be reevaluated.

Novel therapies to treat influenza virus need to be developed. Several studies have screened for host factors which are important for the influenza A virus to infect and replicate[13,94,100,101]. This evidence illustrates the possibility of developing novel classes of anti-influenza drugs. SUMOylation was a common factor identified among the genome-wide screens. Inhibiting SUMOylation can provide valuable insight towards the

development of a new class of inhibitors to treat influenza virus infections as well as other viruses that utilize SUMOylation[102]. In this study, we evaluate the importance of SUMOylation for the influenza virus using our SUMOylation inhibitor, STE.

Materials and Methods

Cloning and Constructs

Influenza A virus segments and influenza B virus segments were cloned in ambisense plasmids – pDZ – which were generously gifted by Dr. García Adolfo Sastre. Mutant genes were created by site-directed mutagenesis and ligated into pDZ plasmids using EcoRI (N-terminus) and NheI restriction sites (C-terminus).

Generation of Influenza viruses

Wildtype PR8 and IAV 3C was rescued from pDZ plasmids based on methods from Martinez-Sobrido and García-Sastre[103]. In short, a cocktail composed of 1 μg of each pDZ plasmid (PB2, PB1, PA, HA, NP, NA, M, and NS) was mixed with 36 μL of P3000 reagent 50 μL of serum-free DMEM. In another tube, 250 μL of serum free DMEM was mixed with 10 μL of Lipofectamine 3000 (ThermoFisher Scientific Waltham, MA). Both tubes were then mixed together and incubated for 15 minutes. One 10 cm plate of HEK293 and MDCK cells were resuspended and mixed in 3 mL of DMEM 10% FBS and 1% Pen-Strep L-Glutamine (Gibco). 250 μL of the cell suspension was added to each well in a tissue culture treated 6-well plate. 1 mL of DMEM supplemented with 10% FBS, 1% Pen-Strep L-Glutamine was added to each well. The transfection mixture was added to each well for

each virus and incubated in a humidified 5% CO_2 incubator overnight. After overnight incubation, the transfection mixture was collected and centrifuged at 800 xg for 5 minutes. The supernatant was collected and stored at -80° C. 2 mL of post-transfection medium (DMEM, 0.01% FBS, 0.42% BSA, 1x Pen-Strep Glutamine, and 1 μg/mL TPCK-trypsin) was added to each well and the plates were incubated for 48 hours in a humidified 5% CO_2 incubator. 1.5 x 10^6 MDCK cells were seeded into tissue-culture treated 6-well plates at 24 hours after the post-transfection medium was added. The supernatant from each virus was collected in 2 mL conical tubes and centrifuged at 5,000 xg for 5 minutes and 300 μL of supernatant was added to each well and incubated at 37° C for one hour with gentle rocking every 10 minutes. The supernatant was then aspirated from each well, 2 mL of post-infection medium (1x MEM (Gibco Ref: 11430-030), 20 mM HEPES pH 7.4, 0.15% sodium bicarbonate, 0.42% BSA, 1x Pen-Strep Glutamine, and 1 μg/mL TPCK-trypsin) was added, and the cells were incubated at 37° C for 72 hours. After 72 hours, the supernatant was collected in 2 mL conical tubes and centrifuged at 5,000 xg for 5 minutes, cleared lysate was transferred into new 2 mL conical tubes and stored at -80° C.

IAV 3C was generated using the same protocol above but swapping the NA pDZ with an NA 3C pDZ plasmid which contains the NA segment from the A/New York/ 08-136/2008 (H1N1) virus, harboring the oseltamivir-resistant S247N and H275Y mutations on the NA gene. IBV Yamagata/88 was generously gifted from Dr. Rong Hai and the virus was propagated in MDCK cells.

Plaque Assay

1.5 x 10^6 MDCK cells were seeded into each well of a 12-well plate tissue culture treated plate. Viruses were serially diluted in a total volume of 300 μL in a medium consisting of 1x PBS, 6.11 mg $CaCl_2$ dihydrate, 10.7 mg $MgCl_2$ hexahydrate, 0.42% BSA, and 1x Pen-Strep Glutamine (Gibco Ref: 10378-016). 200 μL of diluted virus samples were added to each plate and incubated at 37˚ C for IAV and 33˚ C for IBV in a humidified 5% CO_2 incubator for one hour with rocking every ten minutes. The supernatant was aspirated and 1 mL of plaquing medium (1x MEM (Gibco Ref: 11430-030), 0.42 % BSA, 1x Pen-Strep Glutamine (Gibco Ref: 10378-016), 10 mM HEPES pH 7.4, 0.1 % sodium bicarbonate, 0.1% dextrose, 1 μg/mL TPCK-trypsin and 0.76% Avicel RC-591) was added to each well. The plates were incubated in a humidified 5% CO_2 incubator at 37˚ C for 48 hours (IAV) and 33˚C for 72 hours (IBV). The cells were then fixed with 500 μL of 4% (w/v) paraformaldehyde in PBS for one hour and stained with 500 μL of a solution containing 1% crystal violet (w/v) in 10% methanol (v/v).

Plaque Reduction Assay

3.0 x 10^6 MDCK cells were seeded in each well of a 6-well tissue culture treated plate. The cells were treated with specified concentrations of inhibitors overnight. 300 μL of virus dilution buffer with 10 – 30 plaque forming units were added to each well and incubated at 37˚C for one hour with rocking every ten minutes to prevent the cells from drying. Each well was aspirated and 2 mL of plaquing (1x MEM (Gibco Ref: 11430-030), 0.42 % BSA, 1x Pen-Strep Glutamine (Gibco Ref: 10378-016), 10 mM HEPES pH 7.4,

0.1 % sodium bicarbonate, 0.1% dextrose, 1 μg/mL TPCK-trypsin and 0.76% Avicel RC-591) and inhibitor or DMSO (1% final concentration v/v) was added to each well. The plates were then incubated in a humidified 5% CO_2 incubator for 48 hours, IBV-infected cells were incubated for 72 hours at 33˚C in a humidified 5% CO_2 incubator. The cells were fixed with 1 mL PBS with 4% paraformaldehyde for one hour and stained with 500 μL of a solution containing 1% crystal violet (w/v) in 10% methanol (v/v).

Results

Evaluating the importance of SUMOylation for the influenza A virus and influenza B virus

We have found our SUMOylation inhibitor, STE, displays inhibitory effects on both influenza A and influenza B virus in virus plaque assays. When increasing the concentration of STE from 0.5 μM to 3.5 μM, there the influenza A (A/Puerto Rico/8/1934/H1N1) and influenza B (Yamagata/1988) are completely inhibited at 3.0 μM as shown in Figure 3.1A and 3.1B. A drug resistant NA segment 3C from A/New York/287/2009/H1N1 which contains two amino acid substitutions- S247N and H275Y- which confer a heightened resistance to the neuraminidase inhibitor oseltamivir, is also susceptible to the SUMOylation inhibitor STE (Figure 3.1C). These data demonstrate that inhibiting SUMOylation can inhibit two different influenza viruses as well as known oseltamivir-resistant influenza A virus. To understand the mechanism of how the SUMOylation inhibitor prevented viral growth, we studied the SUMOylation of an essential protein for the IAV as well as a previously described target of SUMOylation[74].

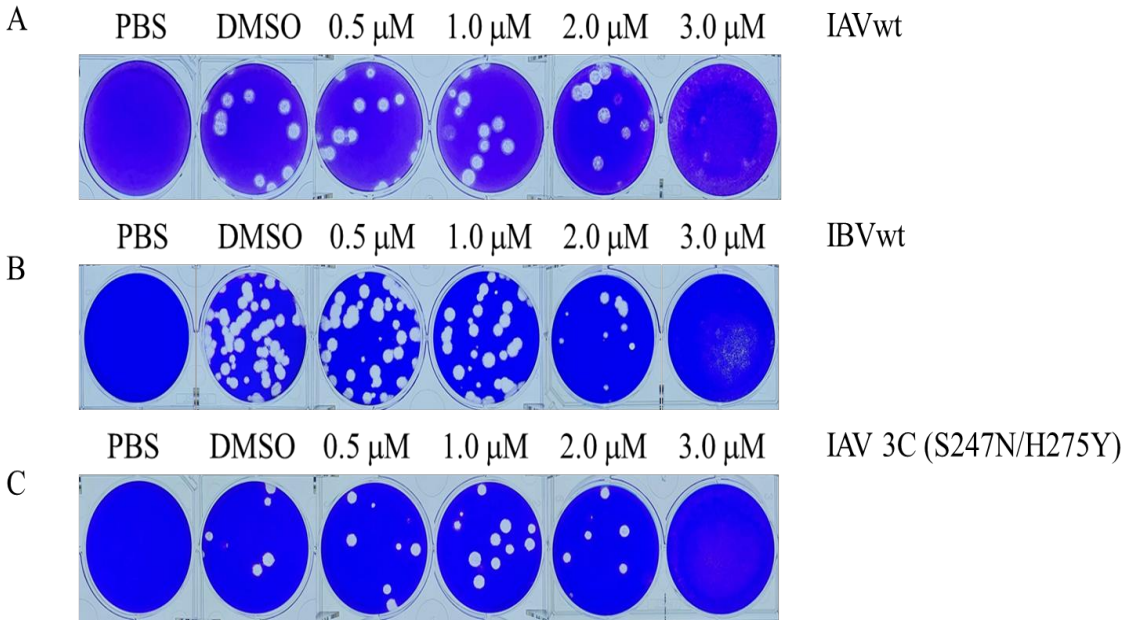


Figure 3.1. Plaque reduction assay on influenza virus. A) Wildtype IAV PR8 infecting MDCK cells with STE at 0.5, 1.0, 2.0, and 3.0 µM. B) Wildtype IBV Yamagata/1988 infecting MDCK cells with STE at 0.5, 1.0, 2.0 and 3.0 µM. C) Oseltamivir-resistant IAV PR8 (NA S247N and H275Y) infecting MDCK cells with STE at 0.5, 1.0, 2.0 and 3.0 µM.

SUMOylation is essential for the virus lifecycle

Previous studies have shown that the influenza A virus interacts extensively with the SUMOylation but these studies have not observed the effects of inhibiting SUMOylation in the host[104]. The SUMOylation of NS1, M1, and NP have been previously studied and have been found to be important for the influenza A virus[72,74,105,106]. Lysine residues involved in SUMOylation have been identified via western blots but that approach lacks the sensitivity needed to properly identify SUMOylated lysine residues[105]. Here, we utilized tandem liquid chromatography and mass spectrometry to identify SUMOylation sites within the M1 protein to propose a mechanism of action by which STE works to inhibit the influenza viruses' ability to continue through the lifecycle. From the mass

tandem LC/MS experiments conducted by Vipul Madahar and the UCR IIGB Facilities, we were able to identify six possible lysine residues – K21, K35, K187, K230, K242, K252 – target by SUMOylation. After mutating each of the lysine residues to arginine in the pDZ M1 plasmid, we created an M1 segment lacking the prespecified lysine residues (6x K>R). Using the eight-plasmid transfection system to generate recombinant IAV PR8, we tested the effects of novel SUMOylation sites identified by LC/MS compared to wildtype after one propagation following transfection (Figure 3.2). Compared to the wildtype PR8, the

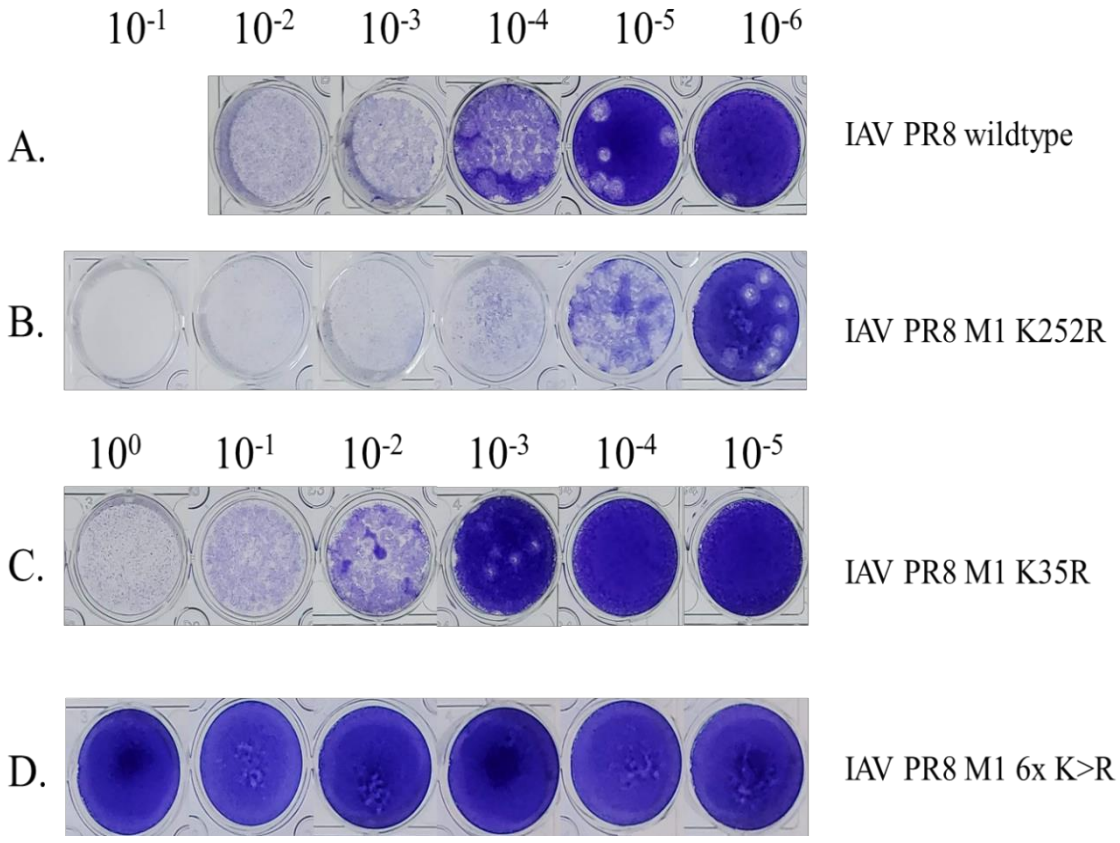


Figure 3.2. Plaque assay of the second propagation of influenza A viruses with mutant M1 proteins. A) Wildtype IAV PR8 titrated from 10^{-2} to 10^{-6}. B) IAV PR8 with an M1 K252R mutant titrated from 10^{-1} to 10^{-6}. C) IAV PR8 M1 with a K35R mutation titrated from 10^{0} to 10^{-5}. D) IAV PR8 M1 with K21/35/187/230/242/252R (6x K>R) mutations titrated from 10^{0} to 10^{-5}.

M1 K252R's growth was not attenuated but the M1 K35R mutant's growth was attenuated (Figure 3.2 A.B.C). Not surprisingly, the PR8 6x K>R was not able to form virus after three independent transfections and propagations (Figure 3.2.D).

Identification of an essential lysine residue for influenza A virus

After creating the 6x K>R mutant and the single mutants, we still were not able to fully explain if the SUMOylation of one lysine residue is more important than the others. To distinguish which SUMOylated lysine residue(s) is essential for the IAV, we mutated

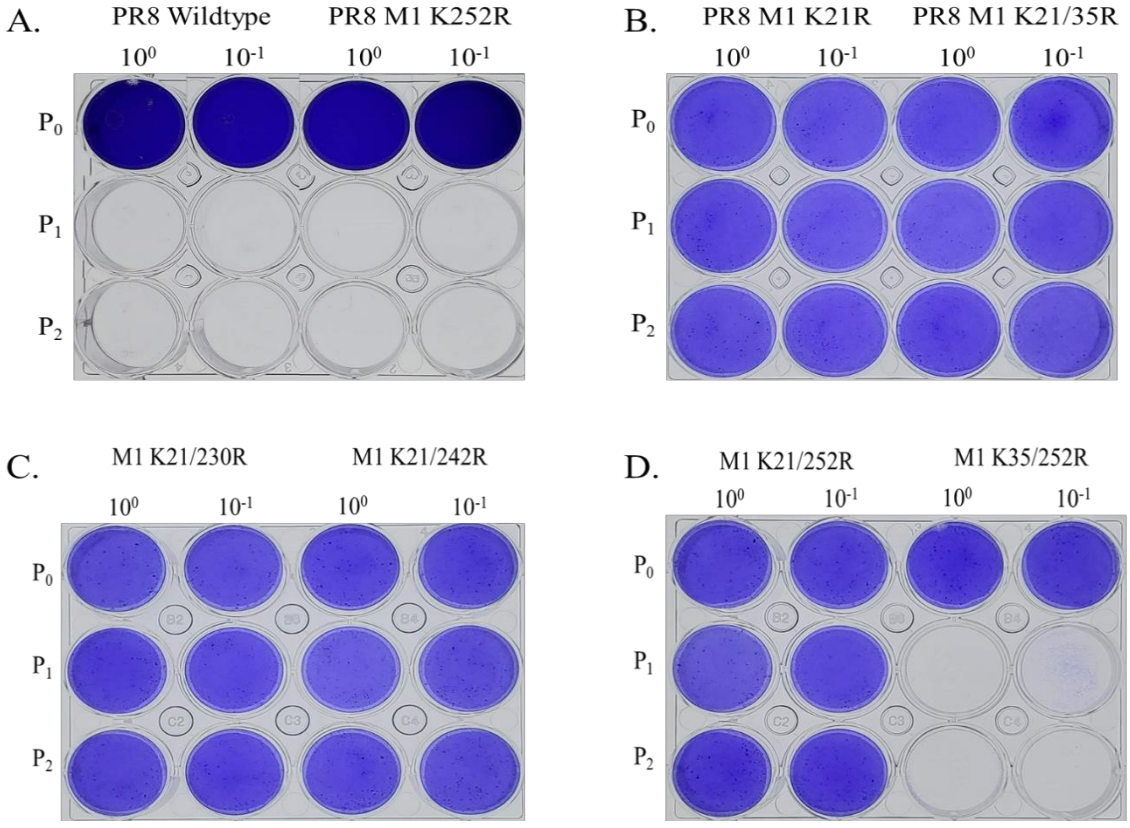


Figure 3.3. Plaque assay of IAV M1 mutants. A) Plaques from PR8 wildtype and PR* M1 K252R viruses formed after transfection and one or two propagations in MDCK cells. B) Plaques formed by PR8 K21R and PR8 K21/35R viruses after transfection and one or two propagations. C) Plaques from PR8 M1 21/230R and PR8 M1 21/242R viruses after transfection and one or two propagations. D) Plaques from PR8 M1 K21/252R and PR8 M1 K35/252R viruses after transfection and one or two propagations.

lysine residue 252 to arginine singularly as well as in combination with lysine residues 21 or 35 to arginine. We also mutated lysine residue 21 and 230 or 242 to arginine. Our initial transfections with the wildtype compared to PR8 M1 K252R and PR8 M1 K21R had a stark contrast in their results (Figure 3.3 A and B). There was no observable difference between the plaques formed by the wildtype and the mutant PR8 K252R, however, there were no observable plaques with the PR8 K21R mutant. When I generated recombinant viruses, PR8 K21/230R and PR8 K21/242R, there were also no observable plaques formed from the virus (Figure 3.3C). To clarify whether the PR8 K21R mutation is responsible, I generated another set of recombinant viruses with the M1 K21/252R mutations or M1 K35/252R mutations and found that M1 K35/252R is able to form plaques and that the K21R mutation by itself is detrimental for the influenza A virus (Figure 3.3D).

To verify the significance of SUMOylation of K21 on the M1 protein for the IAV, we created HEK293 and MDCK cell lines that stably express the wildtype M1 protein of the IAV. These cell lines were used for the transfection and propagation of the recombinant influenza A viruses to provide a wildtype copy of the M1 protein for the recombinant viruses. After three independent transfections with successful wildtype controls, we generated recombinant influenza A viruses with K242R, K21R, or K21/242R mutations on the M1 proteins and performed plaque assays on wildtype MDCK cells and the M1 MDCK stable cell line. The wildtype PR8 and PR8 M1 K242R were able to form plaques in both the M1 MDCK stable cell line and wildtype MDCK cell line (Figure 3.4A and C). To our surprise, we were unable to generate a virus that can make plaques at P_0, P_1, or P_2 with the K21R or K21/242R mutations using the M1 MDCK stable cell line and in the wildtype

MDCK cell line(Figure 3.4B and D). From our data, K21 is an essential SUMOylated lysine residue and lacking this lysine residue will have detrimental effects on the influenza A virus' ability to form plaques. The data suggest that the mutation of the SUMOylated lysine K21R may have a dominate negative effect on the growth of the influenza A virus.

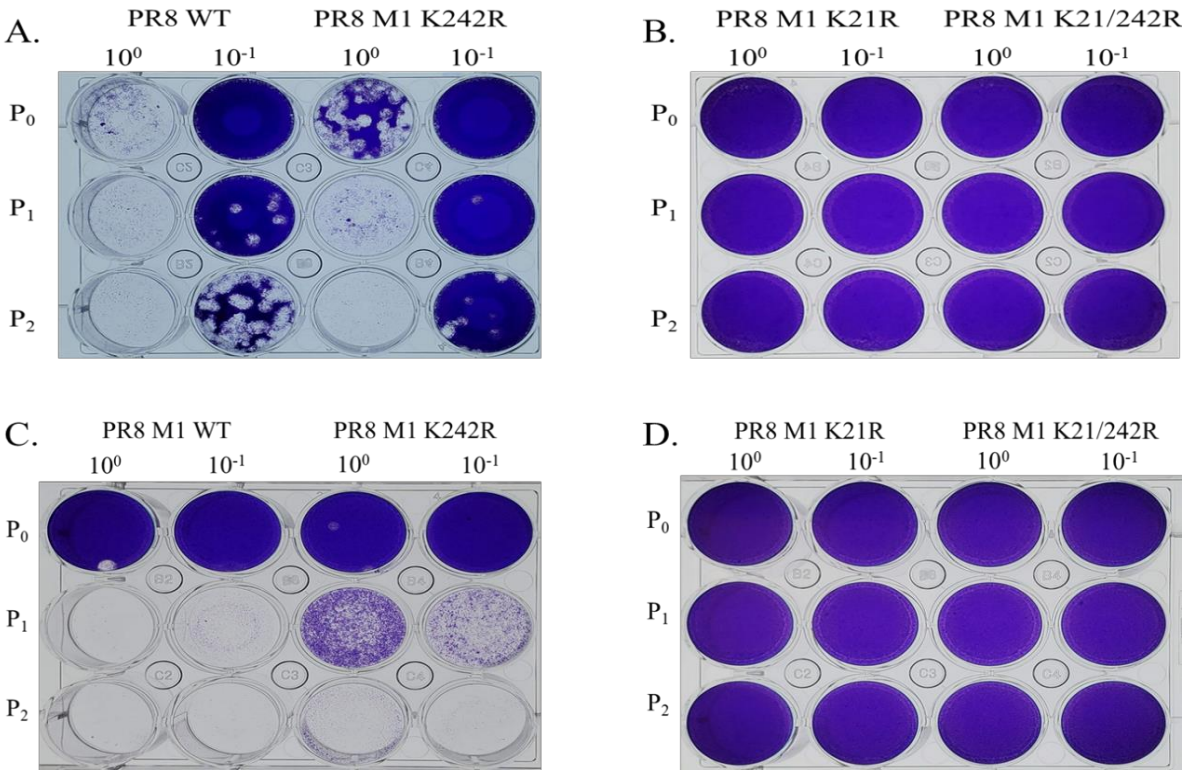


Figure 3.4. Plaque assay results from generating IAV in HEK293 and MDCK cell lines stably expressing wildtype M1. A) Plaques formed in MDCK cells stably expressing M1 with wildtype PR8 IAV and PR8 with the K242R mutation on the M1 protein after transfection, one propagation, and two propagations. B) Plaques formed by PR8 M1 K21R and PR8 M1 K21/242R in MDCK cells stably expressing M1 after transfection, one propagation, and two propagations. C) Plaquing results from wildtype PR8 and PR8 M1 K242R after transfection, one propagation, and two propagations in wildtype MDCK cells. D) Plaques formed by PR8 K21R and PR8 K21/242R in wildtype MDCK cells after transfection, one propagation, and two propagations.

Evaluation of other SUMOylation inhibitors and a NEDDylation inhibitor

In an effort to demonstrate the efficacy of SUMOylation inhibitors effects on IAV replication, we also tested two SUMO inhibitors found by Takeda Pharmaceuticals – ML-792, and TAK-981[107]. From three independent tests, MDCK cells were pretreated with 1.0, 2.5, 5.0, and 10.0 μM of ML-792 or TAK981 and later infected wildtype IAV. My findings show there was no significant difference in the amount of plaques between the DMSO control and the highest concentration of inhibitor treatment as shown in Figure 3.5 A and B. However, when MDCK cells were pretreated with 1.0, 2.5, 5.0, and 7.5 μM of MLN

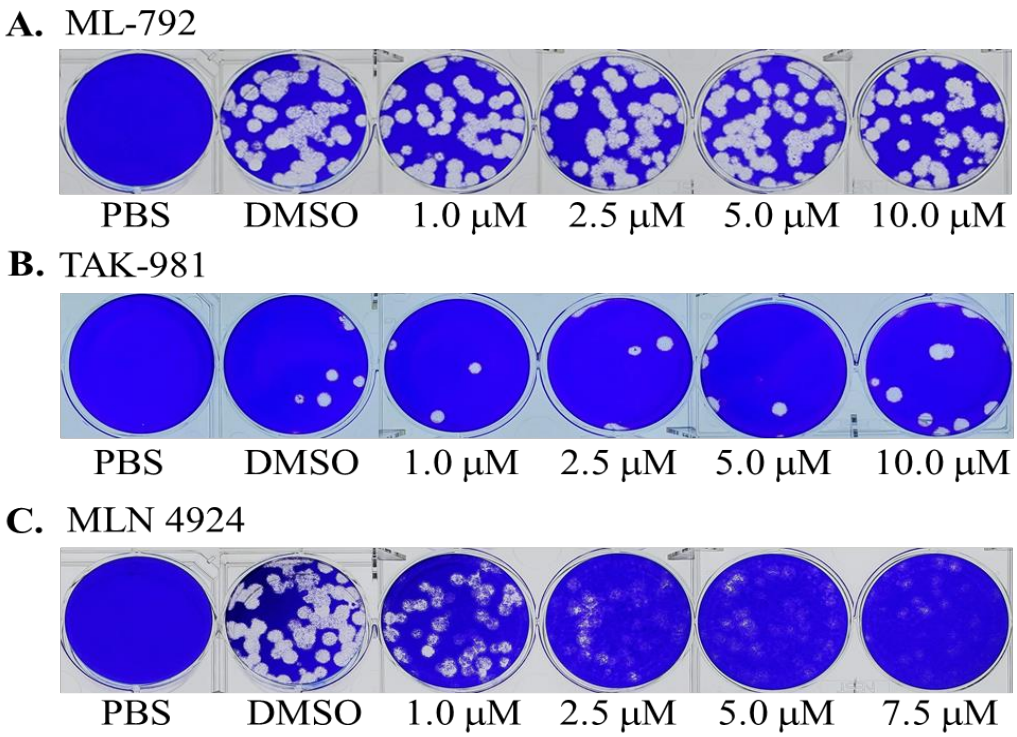


Figure 3.5. SUMOylation and NEDDylation inhibitors effects on IAV plaques. A) Inhibition of IAV plaques with SUMOylation inhibitor ML-792 ranging from 1.0 – 10.0 μM. B) Inhibition of IAV plaques with SUMOylation inhibitor TAK-981 ranging from 1.0 – 10.0 μM. C) Inhibition of IAV plaques with NEDDylation inhibitor MLN 4924 ranging from 1.0 μM to 7.5 μM.

4924 the amount of plaques the IAV can form decreased noticeably starting from 1.0 μM as shown in Figure 3.5C. These results show the importance of NEDDylation for IAV and are in agreement with Sun, H. et al.[108].

I further studied this phenomenon to test if it can inhibit an oseltamivir-resistant strain of influenza A virus (Osel P20) as well as the influenza B virus. When MDCK cells were pretreated with MLN 4924 and later infected with wildtype IAV, Osel P20, or wildtype IBV in the presence of MLN 4924. Surprisingly, the NEDDylation inhibitor was able to inhibit 100% of the IAV wt and Osel P20 plaques at 5.0 μM but a couple plaques were able to form at 7.5 μM, suggesting there may be an optimal concentration (Figure 3.6 A and B). Also, MLN 4924 showed signs of plaque reduction on IBV, although it does not completely inhibit the virus' ability to form plaques, it still demonstrates its ability to decrease the number of plaques as shown in Figure 3.6 C.

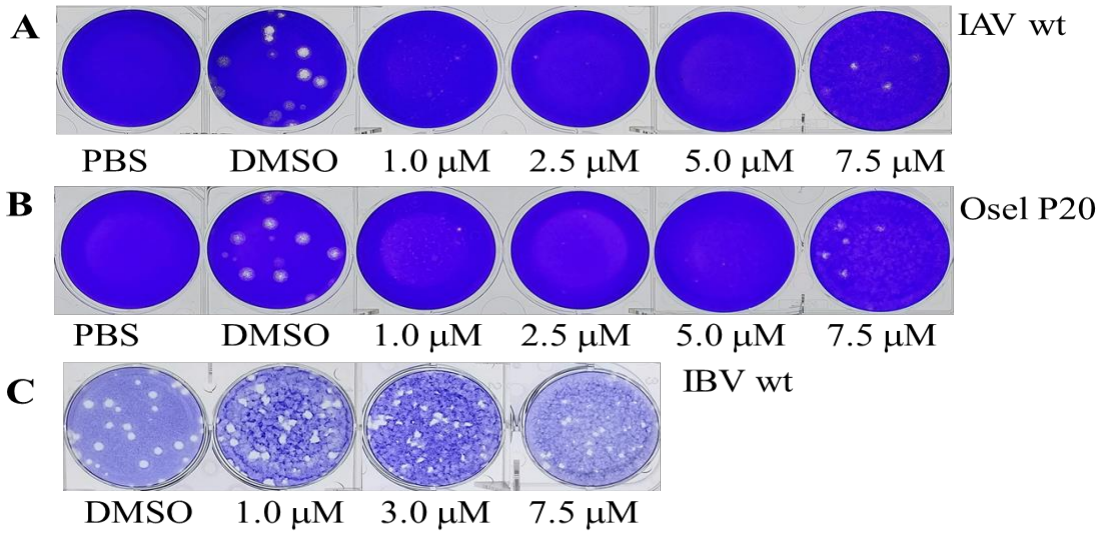


Figure 3.6. Plaque reduction assay of influenza viruses with MLN 4924. A) Inhibition of wildtype IAV with MLN 4924 ranging from 0 – 7.5 μM. B) Inhibition of Osel P20 with MLN 4924 ranging from 0 – 7.5 μM. C) Inhibition of IBV with MLN 4924 from 0 – 7.5 μM.

Discussion

Targeting host factors as an antiviral therapy is a relatively new idea to the field of virology. Identifying key host factors that are hijacked by viruses is the first challenge and finding an inhibitor that works *in cellulo* is the next big challenge of this approach. Several host factors have been identified to be utilized by the influenza A virus in genome-wide screens with SUMOylation showing up in each one[14,94,101]. The role SUMOylation plays in the IAV lifecycle has been probed in past studies at the protein level but the global inhibition of SUMOylation has not been studied[71]. The SUMOylation of NS1 has been shown to lead to a decrease in the IAV growth kinetics, aids in the ability of NS1 to antagonize cellular IFN response and inhibit host gene expression[75,77,105]. SUMOylation of NP was shown to be indispensable for intracellular trafficking of NP[72]. The SUMOylation of M1 has also been studied and found to play a critical role in virus assembly and morphogenesis[74]. These studies have dissected the importance of the individual SUMOylated IAV proteins but the SUMOylation of each individual protein may not be sufficient to explain why our SUMOylation inhibitor, STE, inhibits the IAV and IBV.

The inhibition of SUMOylation has detrimental effects to IAV and IBV. I have demonstrated that inhibiting SUMOylation in MDCK cells diminishes the ability of IAV and IBV to form plaques. Interestingly, an IAV harboring a highly oseltamivir-resistant NA segment (3C) was also inhibited with STE demonstrating an advantage of this approach. This evidence reinforces the idea of developing inhibitors which target host factors

exploited by viruses as potential broad-spectrum anti-influenza virus therapies. The understanding of the role host factors play in the influenza viruses' lifecycles will be critical in the development of novel antiviral therapies.

To understand the mechanism by which our SUMOylation inhibitor works, we investigated the SUMOylation of M1 using tandem liquid chromatography and mass spectrometry. Our results from tandem liquid chromatography and mass spectrometry identified novel SUMOylated lysine residues in the M1 protein of the IAV. One of the SUMOylated lysine residues – K242 – was previously identified by CY Wu et al but did not completely inhibit the virus' ability to infect and replicate[74]. Our approach is more sensitive than the traditional biochemical approach to identify SUMOylation sites because can identify SUMOylated lysine residues with high accuracy. This method resulted in the identification of a lysine residue – K21 – that is essential for the IAV. We created HEK293 and MDCK cell lines stably expressing IAV M1 to create an IAV which can only grow in the stable cell lines. Unfortunately, I was not able to generate any viruses with the K21R or K21/242R mutation on the M1 protein. This may be due to a dominate negative effect which will decrease the efficiency of generating recombinant viruses from a transfection with low efficiency.

Although our SUMOylation inhibitor, STE, has antiviral effects, other inhibitors had different results. The Takeda compounds ML-792 and TAK-981 have been previously shown to inhibit SUMOylation *in vitro* and can inhibit the proliferation of cancer cell lines and cells overexpressing the Myc oncogene, I was unable to inhibit IAV plaque formation[107,109]. These results really highlight the advantage of our inhibitor which can

inhibit SUMOylation *in vitro* and *in cellulo* at concentrations less than 10 μM. Interestingly, the NEDDylation inhibitor MLN 4924 can inhibit wildtype IAV, an oseltamivir-resistant IAV, and wildtype IBV. Previous studies have shown the inhibition of NEDDylation in A549 cells has antiviral activity against different IAV subtypes by inhibiting the CRL/NF-κB pathway and suppressing the expression of pro-inflammatory cytokines induced by IAV[108]. Our results expand upon the previous MLN 4924 findings by demonstrating the disruptive effects it has on IBV plaque formation and an oseltamivir-resistant IAV.

These results demonstrate our approach to target host factors has the potential for the development of a broad-spectrum antiviral therapy. Current antiviral treatments focus on easy-to-target viral proteins. The beauty of this method is the straightforward approach to target conserved regions of the proteins across different strains or inhibit catalytic domains, preventing the protein from performing its function. However, this approach inevitably leads to an arms race between the ever-evolving viruses and pharmaceutical research as seen with the influenza virus. Our approach to target host factors has the direct advantage of circumventing the biological arms race by preventing viruses from hijacking host machinery but at the cost of possible off-target effects. Overall, these data provide fundamental evidence of inhibiting SUMOylation and NEDDylation as potential broad-spectrum anti-influenza virus therapeutics.

Chapter 4: Determining the acquisition of drug-resistant mutations by targeting host factors or viral proteins of the influenza A virus

Abstract

The influenza A virus is a seasonal virus that causes localized epidemics and widespread pandemics on occasion. Vaccines are the primary line of defense to protect the public from pandemics but the virus' ability to mutate and reassort make vaccine predictions a delicate dance between pandemics and controlled flu seasons. Anti-influenza drugs have also been developed to curb the spread and burden of the annual influenza season. The FDA-approved anti-influenza drugs target viral proteins and are prone to become ineffective after many years of use in the general population. Many efforts have been made to determine how fast drug-resistant mutants arise in population sampling studies and *in vitro* studies. However, there is still yet to be a study to bridge the gap between mutations that arise due to selection pressure from direct and indirect inhibitors. This investigation seeks to understand the relationship between the influenza A virus and the acquisition of drug-resistant mutations. Our findings provide advantages for targeting host factors compared to viral proteins for future antiviral therapies.

Introduction

The influenza A virus has had immeasurable impacts on society. The typical flu season can expect more than 3 million infected individuals and the number of deaths upwards of 250,000 globally according to the WHO[110]. In the 2019 – 2020 United States flu season, there were more than 39 million flu illnesses and more than 24,000 flu deaths according the CDC estimates[111]. Various methods have been implemented to help minimize the impact of the influenza virus on the world.

Vaccinations have proven useful to prevent the spread of the influenza virus, but their efficacy primarily relies on predictions. The accuracy of the predictions may vary from year to year and when the predictions fail to protect the population from circulating viruses, widespread pandemics can bring the world economy to a standstill. In an ongoing effort to increase the efficacy of vaccines, new technology and methods have been developed. Currently there are three types of licensed vaccines: inactivated, live attenuated, and recombinant HA vaccines[112,113]. Recently, DNA/RNA vaccines have taken the spotlight showing promising results for vaccine development to prevent influenza virus outbreaks as well as SARS-CoV2.

On the other hand, drugs have also been developed to help protect people against the influenza once they have been infected. Currently there are three classes of FDA-approved drugs to treat the influenza: M2 ion channel blockers, NA inhibitors, and PA inhibitors. Each class of drug targets a specific protein of the virus to inhibit their functions. Although these drugs are effective at combating the IAV, the virus has developed drug-resistant mutations[114]. Influenza viruses are more prone to mutations than their host

counterparts due to their RdRp complex lacking a proofreading function. The error rate of the influenza A virus' can produce 2.3 x 10^{-5} substitutions per nucleotide per cell infection within its 13.6 kb genome, this substitution rate is much higher than the base pair substitution rate in humans which has been determined to be 2.66 x 10^{-9} per mitotic division[115,116]. The current disadvantages of these drugs provide reason to rethink the traditional approach to the development of antiviral drugs.

In this study, we compare a strategy targeting host factors and a virus protein. Using a novel SUMOylation inhibitor (STE) and a neuraminidase inhibitor (oseltamivir), IAV PR8 has been grown in the presence of the individual inhibitors to better understand the virus' ability to accumulate mutations to decrease the efficacy these inhibitors. After 20 propagations, the influenza A virus has developed several drug-resistant mutations for oseltamivir and drug-resistant mutations in response to STE.

Materials and Methods

Cloning and Constructs

The ambisense pDZ constructs (PB2, PB1, PA, HA, NP, NA, M, NS) used for generating the influenza A virus was generously gifted from Dr. Adolfo García-Sastre. Each plasmid contained the genetic information of a corresponding influenza A virus segment.

Generation of virus from eight plasmid transfection

Recombinant IAV was created based on Hoffman et al.[89]. 1 μg of each pDZ plasmid (PB2, PB1, PA, HA, NP, NA, M, NS) was added to 50 μL of serum-free DMEM and 12 μL of P3000 was added to a 1.5 mL conical tube. In another conical tube, 150 μL of serum-

free DMEM was mixed with 10 mL of Lipofectamine 3000. The plasmid cocktail and Lipofectamine 3000 DMEM were mixed together and incubated at room temperature for 15 minutes. One 10 cm dish of 80-90% MDCK and one 10 cm dish of 80-90% HEK293 were stripped with 0.25% Trypsin-EDTA (Gibco, Carlsbad, CA) and resuspended together in 3 mL of DMEM supplemented with 10% FBS and 1% Pen-Strep Glutamine (Gibco, Carlsbad, CA). 250 μL of the cell suspension was added to each well in a 6-well culture plate and 1 mL of supplemented DMEM was added to each well. After the 15-minute incubation, the transfection mixture was added to its corresponding well, each recombinant virus was transfected in triplicate.

Generating Oseltamivir-carboxylate and STE-resistant influenza A viruses *in vitro*

MDCK cells were seeded at 1 x 10^6 cells per well in a 12-well cell culture plate. Both STE and oseltamivir carboxylate were solubilized in DMSO. The cells were pretreated with a specified inhibitor concentration – up to 1% DMSO – the next day and cultured overnight in DMEM supplemented with 10% FBS and 1% Pen-Strep Glutamine. MDCK cells were then infected with IAV PR8 at a MOI 0.001 PFU/cell and cultured in post infection medium (1x MEM containing 0.3% BSA, 1% Pen-Strep Glutamine, 1 μg/mL TPCK-trypsin) with the specified inhibitor for 72 hours. The viruses were serially passaged by infecting pre-treated MDCK cells at a MOI 0.001 PFU/cell with the previous viral stock. At 72 hours, the virus medium was collected, centrifuged at 2,000 xg for 5 minutes and cleared supernatant was aliquoted and stored at -80°C for later analysis. The concentrations of the inhibitor compounds increased every five generations, allowing for the virus to

adjust to each challenge and creating inhibitor-resistant compounds overtime. This selection was carried out for 20 generations, starting with 0.5 – 3.5 μM for STE or 0.5 – 15 μM for oseltamivir carboxylate as shown in Table 4.1. A DMSO-treated control was serially passaged alongside the other two viruses as a control.

Table 4.1 *In vitro* selection of serially passaged virus in STE and Oseltamivir

		Drug concentration used in selection (μM)		
Virus	DMSO	Oseltamivir	STE	Number of days in culture
Wildtype PR8	0	0	0	3
P1	1%	0.5	0.5	6
P2	1%	0.5	0.5	9
P3	1%	0.5	0.5	12
P4	1%	0.5	0.5	15
P5	1%	1.0	1.0	18
P6	1%	1.0	1.0	21
P7	1%	1.0	1.0	24
P8	1%	1.0	1.0	27
P9	1%	1.0	1.0	30
P10	1%	5.0	2.5	33
P11	1%	5.0	2.5	36
P12	1%	5.0	2.5	39
P13	1%	5.0	2.5	42
P14	1%	5.0	2.5	45
P15	1%	10.0	3.0	48
P16	1%	10.0	3.0	51
P17	1%	10.0	3.0	54
P18	1%	10.0	3.0	57
P19	1%	10.0	3.0	60
P20	1%	15.0	3.5	63

Plaque Reduction Assay

3.0 x 10^6 MDCK cells were seeded in each well of a 6-well tissue culture treated plate. The cells were treated with specified concentrations of inhibitors overnight. 300 μL of virus dilution buffer with 10 – 30 plaque forming units were added to each well and

incubated at 37˚C for one hour with rocking every ten minutes to prevent the cells from drying. Each well was aspirated and 2 mL of plaquing (1x MEM (Gibco Ref: 11430-030), 0.42 % BSA, 1x Pen-Strep Glutamine (Gibco Ref: 10378-016), 10 mM HEPES pH 7.4, 0.1 % sodium bicarbonate, 0.1% dextrose, 1 μg/mL TPCK-trypsin and 0.76% Avicel RC-591) and inhibitor or DMSO (1% final concentration v/v) was added to each well. The plates were then incubated in a humidified 5% CO_2 incubator for 48 hours, IBV-infected cells were incubated for 72 hours at 33˚C in a humidified 5% CO_2 incubator. The cells were fixed with 1 mL PBS with 4% paraformaldehyde for one hour and stained with 500 μL of a solution containing 1% crystal violet (w/v) in 10% methanol (v/v).

RNA Extraction and cDNA synthesis

MDCK cells were infected at a MOI of 0.001 with the specified viruses and grown for 48 hours in 1x MEM supplemented with 20 mM HEPES pH 7.4, 0.15% sodium bicarbonate, 0.42% BSA, 1x Pen-Strep Glutamine, and 1 μg/mL TPCK-trypsin. Total RNA was extracted from the infected MDCK cells using the tissue RNA kit plus and the manufacturer's protocol, eluting with 50 μL of DEPC-treated water (San Diego, CA, Biomiga).

2 μM of each forward primer (Table 4.2) was added with 5 μg of purified RNA and DEPC-treated water to a total volume of 13 μL. The primers annealed at 65˚C for 5 minutes and incubated on ice for 1 minute. 7 μL of reverse transcription mix (1x SSIV Buffer, 5 mM DTT, RNaseOUT™ Recombinant RNase Inhibitor, and 200 units of SuperScript®

Reverse Transcriptase) was added to the annealed RNA and mixed via pipetting and centrifuged. The mixtures were incubated at 55˚C for 10 minutes followed by 80˚C for 10 minutes.

PacBio Sequencing

1 μL of cDNA was added to eight different PCR reactions (2 units Express High Fidelity Polymerase (San Diego, CA, Biomiga), 0.2 mM dNTPs, 1x HF Buffer, 0.2 μM forward primer with a 5AmMC6 modification, 0.2 μM reverse primer with a 5AmMC6 modification, and up to 20 μL DEPC-treated ddH_2O) each reaction containing a separate pair of primers for amplifying the different viral segments for each virus. The reactions were incubated at 98°C for 30 seconds and cycled at 98°C for 15 seconds followed by a 55°C annealing for 15 seconds and a 72°C extension for 1 minute and 15 seconds 25 times followed by a 7-minute 72°C final extension. 5 μL of the PCR reactions were checked on a TAE agarose gel.

1 μL of PCR product was used for a secondary PCR reaction to add barcoded sequences (Menlo Park, CA, Pacific Biosciences Part No. 101-629-100) for later identification. The barcoded primers were The reactions were incubated at 98°C for 30 seconds and cycled at 98°C for 15 seconds followed by a 64°C annealing for 15 seconds and a 72°C extension for 2 minutes 25 times followed by a 7-minute 72°C final extension. 5 μL of the PCR reactions were loaded onto a 1% TAE agarose gel and imaged on a UVP Biospectrum AC Imaging System. The polymerized products were then pooled according to their genes and purified with the GE illustra GFX DNA purification kit (Cytiva,

Marlborough, MA). The DNA samples were quantified with a nanodrop 2000c and submitted to UCI GHTF for adapter ligation and PacBio sequencing on the Sequel II system.

Table 4.2. Primer design for PacBio cDNA synthesis and PacBio sequencing.

Primer	Sequence
PB2 Fwd Pbio Seq	/5AmMC6/ gcagtcgaacatgtagctgactcaggtcacGCAGGTCAATTATATTCAAT
PB2 Rev Pbio Seq	/5AmMC6/ tggatcacttgtgcaagcatcacatcgtagCGTTTTTAAACTATTCGACA
PB1 Fwd Pbio Seq	/5AmMC6/ gcagtcgaacatgtagctgactcaggtcacAAAGCAGGCAAACCATTTGA
PB1 Rev Pbio Seq	/5AmMC6/ tggatcacttgtgcaagcatcacatcgtagTGAAGGACAAGCTAAATTCA
PA Fwd Pbio Seq	/5AmMC6/ gcagtcgaacatgtagctgactcaggtcacAAAGCAGGTACTGATCCAAA
PA Rev Pbio Seq	/5AmMC6/ tggatcacttgtgcaagcatcacatcgtagAAATAGTAGCACTGCCACAA
HA Fwd Pbio Seq	/5AmMC6/ gcagtcgaacatgtagctgactcaggtcacGGGGAAAATAAAAACAACCA
HA Rev Pbio Seq	/5AmMC6/ tggatcacttgtgcaagcatcacatcgtagCCTCATATTTCTGAAATTCTAA
NP Fwd Pbio Seq	/5AmMC6/ gcagtcgaacatgtagctgactcaggtcacGTAGATAATCACTCACTGAG
NP Rev Pbio Seq	/5AmMC6/ tggatcacttgtgcaagcatcacatcgtagAAACAAGGGTATTTTTCTTT
NA Fwd Pbio Seq	/5AmMC6/ gcagtcgaacatgtagctgactcaggtcacAGCGAAAGCAGGGGTTTAAA
NA Rev Pbio Seq	/5AmMC6/ tggatcacttgtgcaagcatcacatcgtagGAAACAAGGAGTTTTTTGAACAG
M Fwd Pbio Seq	/5AmMC6/ gcagtcgaacatgtagctgactcaggtcacGAAAGCAGGTAGATATTGAA
M Rev Pbio Seq	/5AmMC6/ tggatcacttgtgcaagcatcacatcgtagACAAGGTAGTTTTTTACTCC
NS Fwd Pbio Seq	/5AmMC6/ gcagtcgaacatgtagctgactcaggtcacCAAAAGCAGGGTGACAAAGA
NS Rev Pbio Seq	/5AmMC6/ tggatcacttgtgcaagcatcacatcgtagAGTAGAAACAAGGGTGTTTT

Sanger Sequencing

Eight different PCR reactions were set up for each virus, amplifying the eight different segments of the virus. The PCR reactions consisted of 1 μL of synthesized cDNA 0.2 μM of a forward and reverse primer set, 0.2 mM dNTPs, 1x HF Buffer, 2 units of EHF polymerase (Biomiga), and DEPC-treated water to a total volume of 25 μL. The reactions were incubated at 98°C for 30 seconds and cycled at 98°C for 15 seconds followed by a 55°C annealing for 15 seconds and a 72°C extension for 1 minute and 15 seconds 25 times followed by a 7-minute 72°C final extension. 5 μL of the PCR reactions were checked on a TAE agarose gel. PCR products were purified using illustra GFX DNA purification kit (Cytiva, Marlborough, MA) and eluted in 20 μL of molecular biology-grade water. 200 ng of PCR product was submitted for Sanger sequencing at the UCR Genomics core facility.

Data Analysis

PacBio sequencing data was analyzed via PacBio SMRT Analysis at UCI GHTF. Chromatograms from Sanger sequencing data was analyzed with chromas lite and alignments were made with CLC Sequence Viewer 7.

Results

Selecting inhibitor-resistant mutants

H1N1 A/Puerto Rico/8/1934 (PR8) was grown in the presence of inhibitors with stepwise increases in concentration every five generations as described in Table 4.1. After 20 generations, the IAV grown in the presence of inhibitors have developed a resistance, depicted in Figure 4.1. To highlight the extent to which the viruses have developed a

resistance to their respective inhibitors, the viruses were challenged with different concentrations of the inhibitor they were grown with and the other inhibitor the viruses have not been exposed to. After 15 propagations with the inhibitors, both viruses were sensitive to STE treatments and surprisingly, they were both sensitive to oseltamivir treatments as shown in Figure 4.2. After 20 propagations, the STE and Oseltamivir grown

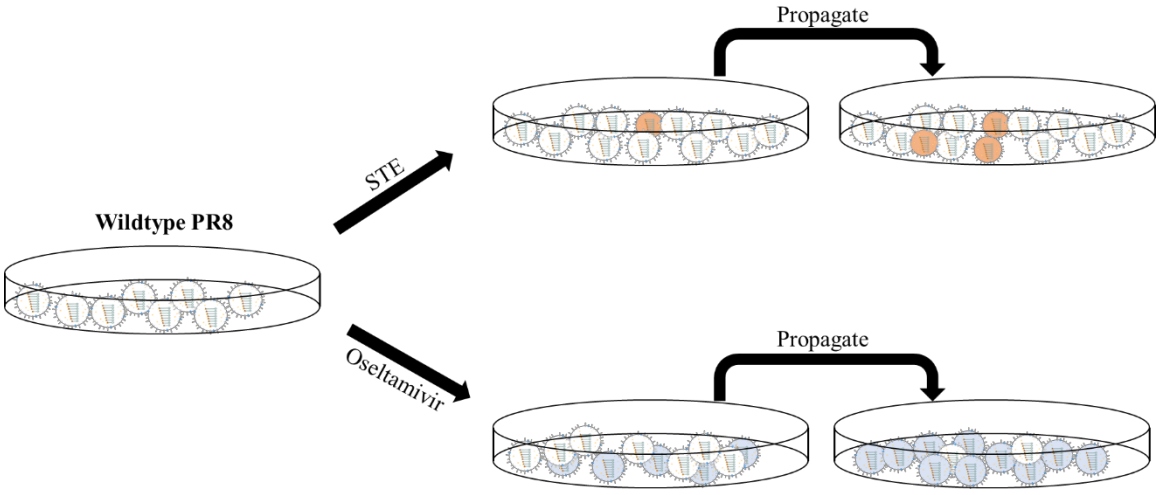


Figure 4.1. Schematic of selecting for inhibitor-resistant viruses. Wildtype virus is propagated in the presence of inhibitors, increasing every five propagations. Overtime, only inhibitor-resistant viruses will be cultured.

viruses were still sensitive to STE treatments. The viruses grown with oseltamivir were able to grow in 15 μM oseltamivir when the viruses grown with STE were not able to grow well in 1 μM oseltamivir. Interestingly, both lineages of virus were unable to grow in 3.5 μM of STE (Figure 4.2) which demonstrates the virus' inability to acclimate to a SUMOylation-deficient host. These data show the stark contrast between the virus' ability to mutate and develop resistances against virus-specific inhibitors and host factor inhibitors.

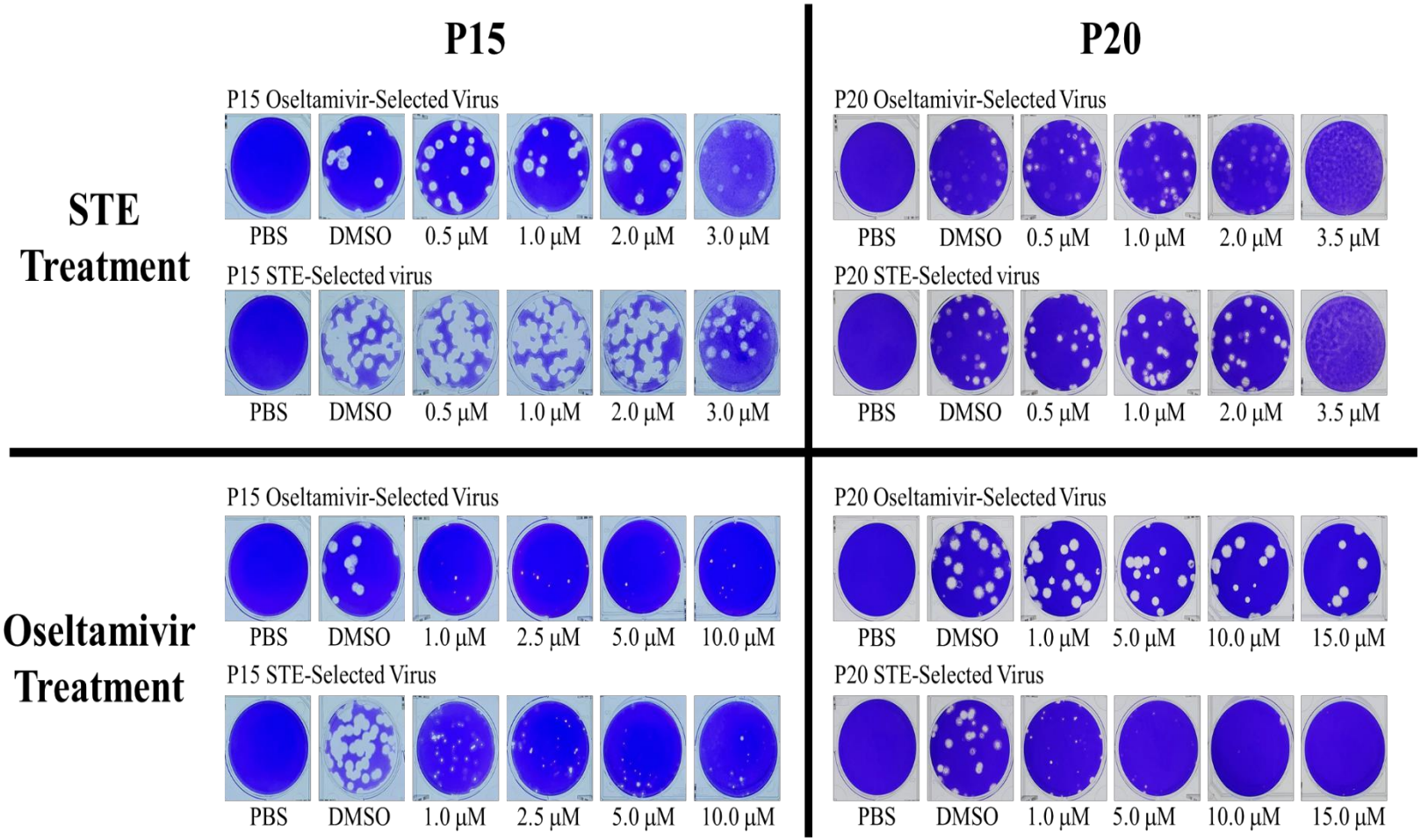


Figure 4.2. Plaque reduction assay results from virus propagations in the presence of inhibitors. The viruses grown in the presence of STE are inhibited by STE at propagation 15 and 20 but are sensitive to oseltamivir. Viruses grown in the presence of oseltamivir have developed a resistance to the inhibitor but are sensitive to STE.

NA segment mutations

To detect possible emergence of NA segments that developed drug-resistant mutations, the NA segment of P20 from each virus lineage (DMSO, Oseltamivir, and STE) was extracted from infected MDCK cells and PCR amplified. The PCR amplicons were then submitted from Sanger sequencing. The Sanger sequencing results of the NA segment in the 20th propagation of the DMSO, Oseltamivir, and STE lineages covered 100% of the NA segment with some redundancy. Each of the segments were sequenced from the 5' end, 3' end, 5' to 3' from bp position 401, and 3' to 5' from bp position 965 to cover as much of the segment as possible. From sequence alignment and chromatogram analysis, the different amino acid substitutions were detected as shown in Table 4.3. From the chromatogram data, the nucleotide substitutions are minorities at the corresponding

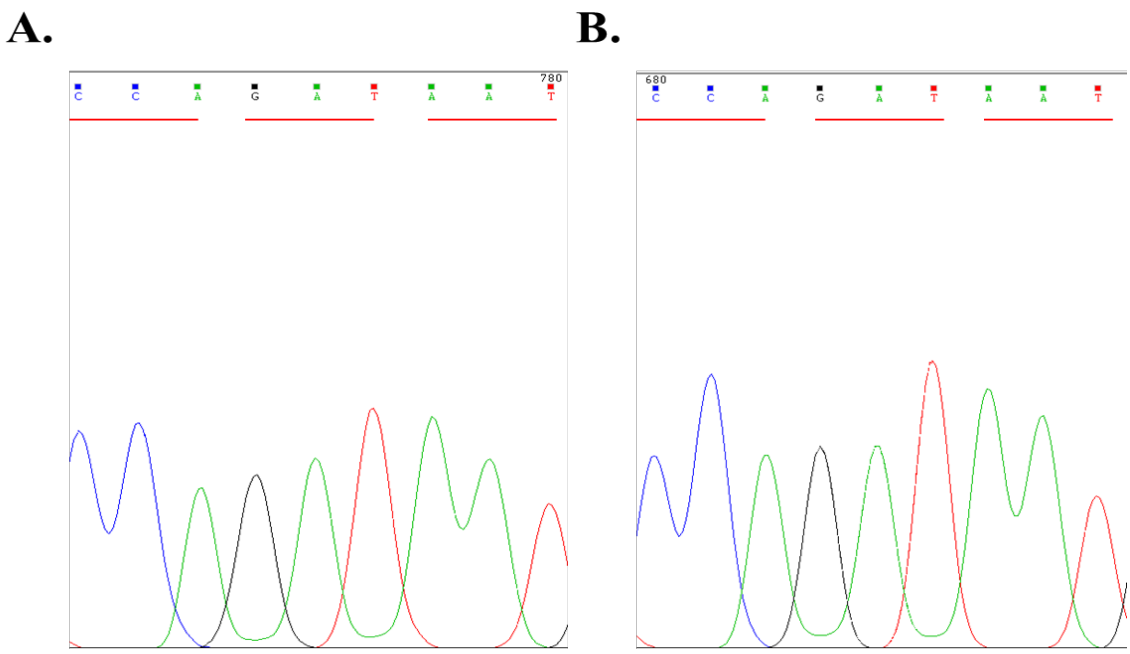


Figure 4.3. Chromatograms from sequencing the PCR products of the NA segment of the DMSO P20 virus (A) and STE P20 virus (B). The red bars underlining the consensus reads represent the codon positions.

positions. These single nucleotide polymorphisms result in amino acid substitutions shown in Table 4.3. The chromatogram files for DMSO P20 and STE P20 had remarkable resemblance with one another as shown in Figure 4.3A and B, respectively. There were remarkably few mutations at key residues for oseltamivir resistance, however, there was a point mutation which resulted in the amino acid mutation of glutamate at position 119 to lysine as shown in Table 4.3. Interestingly, the Osel P20 chromatogram files revealed several known drug-resistant mutations to oseltamivir. A transition mutation from G to A at nucleotide position 410 results in an arginine to lysine mutation at amino acid position 152 but is only a minor variant in the overall population from the Sanger sequencing data (Figure 4.4.A). Another transition mutation from G to A at nucleotide position 552 which resulted in an aspartate to asparagine mutation at amino acid position 199 as shown in Figure 4.4.B.

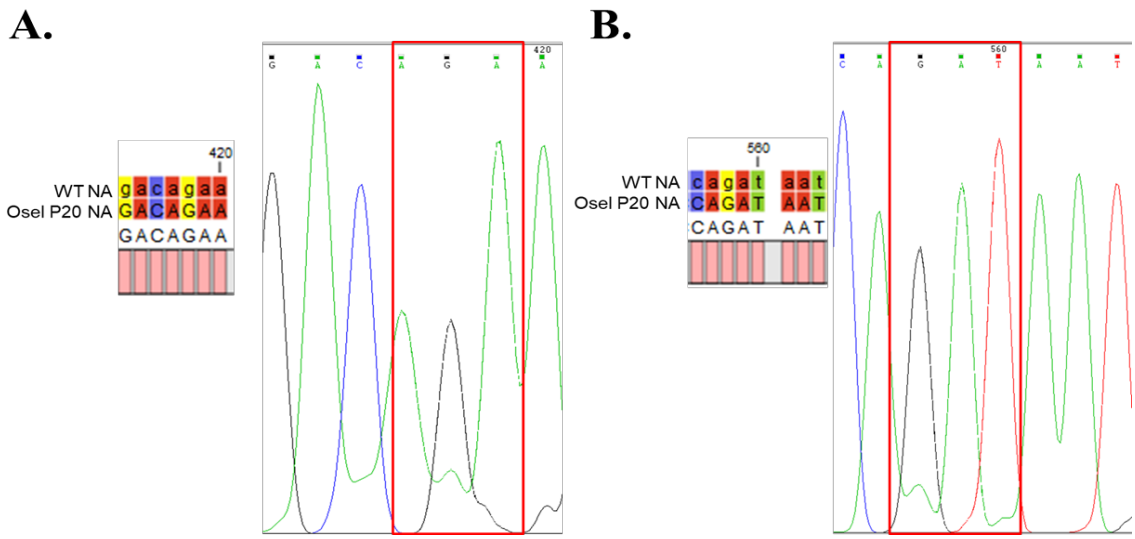


Figure 4.4. Chromatograms and corresponding alignments of the wildtype NA segment and the Osel P20 NA segment. A) Identification of a minor SNP at nucleotide position 411 resulting in an arginine to lysine at amino acid position 152. B) Identification of a minor SNP at nucleotide position 552, resulting in an aspartate to asparagine at amino acid position 199.

The different amino acid mutations were compiled in Table 4.3. Both the DMSO P20 and STE P20 NA segments harbored a E119K mutation as the only detectable mutation from the Sanger sequencing data. In contrast, the Oseltamivir P20 NA segment contained several known mutations which confer a heightened drug-resistance to oseltamivir (R152K, D199N, I223T, H275Y, and R293K) which are consistent with what is expected when selecting for oseltamivir-resistant viruses.

Table 4.3 Amino Acid variants identified from Sanger sequencing in different viruses

Virus	Position	Reference Amino Acid	Amino Acid change
DMSO P20	119	E	K
STE P20	119	E	K
Oseltamivir P20	152	R	K
	199	D	N
	223	I	T
	275	H	Y
	293	R	K

Discussion

Targeting essential host factors for viruses can provide essential information to develop novel antiviral therapies. Our data demonstrate the distinct advantage of targeting the host SUMOylation cascade to inhibit viral growth. Inhibiting SUMOylation did not increase the virus' ability to grow in the presence of the inhibitor as it did with oseltamivir as shown in Figure 4.2. This is a distinct advantage over the established strategy to target viral proteins. While the influenza A virus' substitution rate per base pair per cell infection (2.3×10^{-5}) is much higher than human's substitution rate per base pair per mitotic division (3.3×10^{-11}), it is not able to accumulate drug-resistant mutations to our SUMOylation inhibitor as fast as it is able to accumulate drug-resistant mutations to oseltamivir[115,116].

From the Sanger sequencing results, we can see a difference in the amount of mutations between the three different viruses. These data depict the complexity of variants in the population but not in a quantitative capacity. Also, the Sanger sequencing methodology will not allow us to resolve where the observed mutations originated. This information will help separate the difference between a segment that would be found in an infective viral particle versus a segment that may contain nonsense mutations which would not produce an infectious viral particle. Using the PacBio sequencing methodology, we would be able to discern the sequencing coverage, barcode each segment individually, and have a better understanding of where these mutations arose. Nevertheless, these sequencing results still provide insight into the amount of mutations the virus can develop in response to a drug that targets a viral protein and a potential drug which targets a host factor necessary for the virus lifecycle.

Our results demonstrate the accumulation of drug-resistant mutations for oseltamivir are much higher than STE and DMSO. When evaluating the amino acid positions known to mutate and confer drug-resistance to oseltamivir, there were more drug-resistant mutations in the NA segment of Osel P20 than STE and DMSO. This phenomenon can be attributed to a couple differences between the two inhibitors. First, oseltamivir targets the viral neuraminidase protein which prevents the progeny virions from budding off the host cell[66]. Second, the SUMOylation inhibitor, STE, inhibits an essential host factor for the influenza A virus. Though the virus is prone to mutate, many mutations will be deleterious and only a handful will be advantageous for the next generation of viruses. Expecting the virus to develop a dependency on a new host factor that will fulfill the role of SUMOylation will take more time than 60 days in culture with the inhibitor. Understanding the whole system can provide the proper insight to understand the cause and effect of using different approaches towards the development of antiviral drugs. Targeting specific elements can elicit different responses, virus proteins are more susceptible to change and developing drug-resistance than host factors. Although the sequencing data are not as comprehensive as data from next-generation sequencing, it still provides sufficient evidence that SNPs which result in amino acid mutations do exist within the segment populations.

From the data collected, we have found distinct advantages of developing antiviral drugs that target the host factors instead of the viral proteins. Due to a culmination of factors such as; the differences between the host cells' rate of base pair substitution and the virus' rate of base pair substitution, and the virus' dependence on host factors, the development of antiviral therapies targeting host factors are more attractive.

References

1. Ouyang, M. *et al.* Visualization of polarized membrane type 1 matrix metalloproteinase activity in live cells by fluorescence resonance energy transfer imaging. *J. Biol. Chem.* **283**, 17740–17748 (2008).
2. Song, Y., Madahar, V. & Liao, J. Development of FRET assay into quantitative and high-throughput screening technology platforms for protein-protein interactions. *Ann. Biomed. Eng.* **39**, 1224–1234 (2011).
3. Hochreiter, B., Garcia, A. P. & Schmid, J. A. Fluorescent proteins as genetically encoded FRET biosensors in life sciences. *Sensors (Switzerland)* **15**, 26281–26314 (2015).
4. Liu, Y., Song, Y., Madahar, V. & Liao, J. Quantitative Förster resonance energy transfer analysis for kinetic determinations of SUMO-specific protease. *Anal. Biochem.* **422**, 14–21 (2012).
5. Song, Y. & Liao, J. Systematic determinations of SUMOylation activation intermediates and dynamics by a sensitive and quantitative FRET assay. *Mol. Biosyst.* **8**, 1723 (2012).
6. Nguyen, A. W. & Daugherty, P. S. Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nat. Biotechnol.* **23**, 355–360 (2005).
7. Chen, H., Puhl, H. L., Koushik, S. V., Vogel, S. S. & Ikeda, S. R. Measurement of FRET Efficiency and Ratio of Donor to Acceptor Concentration in Living Cells. *Biophys. J.* **91**, 39–41 (2006).
8. Chen, H., Puhl, H. L. & Ikeda, S. R. Estimating protein-protein interaction affinity in living cells using quantitative Förster resonance energy transfer measurements. *J. Biomed. Opt.* **12**, 054011 (2007).
9. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016).
10. Isogai, S. & Shirakawa, M. Protein modification by SUMO. *Seikagaku* **79**, 1120–1130 (2007).
11. Kessler, J. D. *et al.* A SUMOylation-Dependent Transcriptional Subprogram Is Required for Myc-Driven Tumorigenesis. *Science (80-.).* **335**, 348–353 (2012).
12. Flotho, A. & Melchior, F. Sumoylation: A Regulatory Protein Modification in Health and Disease. *Annu. Rev. Biochem.* **82**, 357–385 (2013).
13. Karlas, A. *et al.* Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature* **463**, 818–822 (2010).
14. Shapira, S. D. *et al.* A Physical and Regulatory Map of Host-Influenza Interactions Reveals Pathways in H1N1 Infection. *Cell* **139**, 1255–1267 (2009).
15. Domingues, P. *et al.* Global Reprogramming of Host SUMOylation during Resource Global Reprogramming of Host SUMOylation during Influenza Virus Infection. *CELREP* **13**, 1467–1480 (2015).
16. Su, H. L. & Li, S. S. L. Molecular features of human ubiquitin-like SUMO genes and their encoded proteins. *Gene* **296**, 65–73 (2002).
17. Knipscheer, P. *et al.* Ubc9 Sumoylation Regulates SUMO Target Discrimination. *Mol. Cell* **31**, 371–382 (2008).

18. Takahashi, Y. & Strunnikov, A. In vivo modeling of polysumoylation uncovers targeting of Topoisomerase II to the nucleolus via optimal level of SUMO modification. *Chromosoma* **117**, 189–198 (2008).
19. Sharrocks, A. D. PIAS proteins and transcriptional regulation - More than just SUMO E3 ligases? *Genes Dev.* **20**, 754–758 (2006).
20. Renzini, A., Marroncelli, N., Noviello, C., Moresi, V. & Adamo, S. HDAC4 Regulates Skeletal Muscle Regeneration via Soluble Factors. *Front. Physiol.* **9**, 1–11 (2018).
21. Yamashita, D., Moriuchi, T., Osumi, T. & Hirose, F. Transcription factor hDREF is a novel SUMO E3 ligase of Mi2α. *J. Biol. Chem.* **291**, 11619–11634 (2016).
22. Varejão, N. *et al.* DNA activates the Nse2/Mms21 SUMO E3 ligase in the Smc5/6 complex. *EMBO J.* **37**, 1–16 (2018).
23. Cohen, I. & Ezhkova, E. Cbx4: A new guardian of p63's domain of epidermal control. *J. Cell Biol.* **212**, 9–11 (2016).
24. Subramaniam, S. *et al.* Rhes, a physiologic regulator of sumoylation, enhances cross-sumoylation between the basic sumoylation enzymes E1 and Ubc9. *J. Biol. Chem.* **285**, 20428–20432 (2010).
25. Guervilly, J. H. *et al.* The SLX4 complex is a SUMO E3 ligase that impacts on replication stress outcome and genome stability. *Mol. Cell* **57**, 123–137 (2015).
26. Pungaliya, P. *et al.* TOPORS functions as a SUMO-1 E3 ligase for chromatin-modifying proteins. *J. Proteome Res.* **6**, 3918–3923 (2007).
27. Schellenberg, M. J. *et al.* ZATT (ZNF451)–mediated resolution of topoisomerase 2 DNA-protein cross-links. *Science (80-.).* **357**, 1412–1416 (2017).
28. Ikeuchi, Y. *et al.* TIF1γ protein regulates epithelial-mesenchymal transition by operating as a Small Ubiquitin-like Modifier (SUMO) E3 ligase for the transcriptional regulator SnoN. *J. Biol. Chem.* **289**, 25067–25078 (2014).
29. Bouvier, N. M. & Palese, P. The biology of influenza viruses. *Vaccine* **26**, 49–53 (2008).
30. Taubenberger, J. K. & Kash, J. C. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe* **7**, 440–451 (2010).
31. Benjamin O. Fulton, Weina Sun, Nicholas S. Heaton, P. P. The Influenza B Virus Hemagglutinin Head Domain Is Less Tolerant to Transposon Mutagenesis than That of the Influenza A Virus. **92**, 1–2 (2018).
32. Heldt, F. S. Mathematical models of influenza A virus infection: from intracellular replication to virus growth in cell populations. *PhD Diss. Fac. Process Syst. Eng. Otto von Guericke Univ. Magdebg.* (2014).
33. Jahn, R. & Sudhof, T. RECEPTOR BINDING AND MEMBRANE FUSION IN VIRUS ENTRY: The Influenza Hemagglutinin. *Annu. Rev. Biochem.* 777–810 (1999) doi:10.1128/JVI.06147-11.
34. Samji, T. Influenza A: Understanding the viral life cycle. *Yale J. Biol. Med.* **82**, 153–159 (2009).
35. Shen, X., Zhang, X. & Liu, S. Novel hemagglutinin-based influenza virus inhibitors. *J. Thorac. Dis.* **5**, (2013).
36. Walker, A. P. & Fodor, E. Interplay between Influenza Virus and the Host RNA

Polymerase II Transcriptional Machinery. *Trends Microbiol.* **27**, 398–407 (2019).
37. Ramon Flick, Gabrielle Neumann, Erich Hoffmann, Elisabeth Neumeier, G. H. PROMOTER ELEMENTS IN THE INFLUENZA VRNA PROMOTER.pdf. 1046–1057 (1996).
38. Crow, M., Deng, T., Addley, M. & Brownlee, G. G. Mutational Analysis of the Influenza Virus cRNA Promoter and Identification of Nucleotides Critical for Replication. *J. Virol.* **78**, 6263–6270 (2004).
39. Deng, T., Vreede, F. T. & Brownlee, G. G. Different De Novo Initiation Strategies Are Used by Influenza Virus RNA Polymerase on Its cRNA and Viral RNA Promoters during Viral RNA Replication. *J. Virol.* **80**, 2337–2348 (2006).
40. Nayak, D. P., Balogun, R. A., Yamada, H., Zhou, Z. H. & Barman, S. Influenza virus morphogenesis and budding. *Virus Res.* **143**, 147–161 (2009).
41. Chanturiya, A. N. *et al.* PB1-F2, an Influenza A Virus-Encoded Proapoptotic Mitochondrial Protein, Creates Variably Sized Pores in Planar Lipid Membranes. *J. Virol.* **78**, 6304–6312 (2004).
42. Košík, I. *et al.* The ubiquitination of the influenza a virus PB1-F2 protein is crucial for its biological function. *PLoS One* **10**, 1–21 (2015).
43. Krumbholz, A. *et al.* Current knowledge on PB1-F2 of influenza A viruses. *Med. Microbiol. Immunol.* **200**, 69–75 (2011).
44. Wise, H. M. *et al.* A Complicated Message: Identification of a Novel PB1-Related Protein Translated from Influenza A Virus Segment 2 mRNA. *J. Virol.* **83**, 8021–8031 (2009).
45. Hu, J. *et al.* PA-X Decreases the Pathogenicity of Highly Pathogenic H5N1 Influenza A Virus in Avian Species by Inhibiting Virus Replication and Host Response. *J. Virol.* **89**, 4126–4142 (2015).
46. Jagger, B. W. *et al.* An Overlapping Protein-Coding Region in Influenza A Virus Segment 3 Modulates the Host Response. *Science (80-.).* **337**, 199–204 (2012).
47. Wise, H. M. *et al.* Identification of a Novel Splice Variant Form of the Influenza A Virus M2 Ion Channel with an Antigenically Distinct Ectodomain. *PLoS Pathog.* **8**, (2012).
48. Selman, M., Dankar, S. K., Forbes, N. E., Jia, J. J. & Brown, E. G. Adaptive mutation in influenza A virus non-structural gene is linked to host switching and induces a novel protein by alternative splicing. *Emerg. Microbes Infect.* **1**, 0 (2012).
49. Wong, S. S. & Webby, R. J. Traditional and new influenza vaccines. *Clin. Microbiol. Rev.* **26**, 476–492 (2013).
50. Jefferson, T., Demicheli, V., Di Pietrantonj, C. & Rivetti, D. Amantadine and rimantadine for influenza A in adults. *Cochrane Database Syst. Rev.* (2006) doi:10.1002/14651858.CD001169.pub3.
51. Beigel, J. H. *et al.* Oseltamivir, amantadine, and ribavirin combination antiviral therapy versus oseltamivir monotherapy for the treatment of influenza: a multicentre, double-blind, randomised phase 2 trial. *Lancet Infect. Dis.* **17**, 1255–1265 (2017).
52. Lumby, C. K. *et al.* Favipiravir and Zanamivir Cleared Infection with Influenza B

in a Severely Immunocompromised Child. *Clin. Infect. Dis.* 1–4 (2020) doi:10.1093/cid/ciaa023.
53. Aoki, F. Y. Antiviral Drugs for Influenza and Other Respiratory Virus Infections. *Mand. Douglas, Bennett's Princ. Pract. Infect. Dis.* **1**, 531–545 (2014).
54. Mckimm-Breschkin, J. L. Influenza neuraminidase inhibitors: Antiviral action and mechanisms of resistance. *Influenza Other Respi. Viruses* **7**, 25–36 (2013).
55. Hayden, F. G. & Shindo, N. Influenza virus polymerase inhibitors in clinical development. *Curr. Opin. Infect. Dis.* **32**, 176–186 (2019).
56. Hayden, F. G. *et al.* Baloxavir marboxil for uncomplicated influenza in adults and adolescents. *N. Engl. J. Med.* **379**, 913–923 (2018).
57. Chan-Tack, K. M., Murray, J. S. & Birnkrant, D. B. Use of ribavirin to treat influenza. *N. Engl. J. Med.* **361**, 1713–1714 (2009).
58. Friesen, R. H. E. *et al.* New class of monoclonal antibodies against severe influenza: Prophylactic and therapeutic efficacy in ferrets. *PLoS One* **5**, 1–7 (2010).
59. Zeng, L. Y., Yang, J. & Liu, S. Investigational hemagglutinin-targeted influenza virus inhibitors. *Expert Opin. Investig. Drugs* **26**, 63–73 (2017).
60. Shen, Z., Lou, K. & Wang, W. New small-molecule drug design strategies for fighting resistant influenza A. *Acta Pharm. Sin. B* **5**, 419–430 (2015).
61. Eyer, L. & Hruska, K. Antiviral agents targeting the influenza virus: A review and publication analysis. *Vet. Med. (Praha).* **58**, 113–185 (2013).
62. McKimm-Breschkin, J. *et al.* Neuraminidase Sequence Analysis and Susceptibilities of Influenza Virus Clinical Isolates to Zanamivir and Oseltamivir Neuraminidase Sequence Analysis and Susceptibilities of Influenza Virus Clinical Isolates to Zanamivir and Oseltamivir. *Antimicrob. Agents Chemother.* **47**, 2264–2272 (2003).
63. Furuta, Y. *et al.* T-705 (favipiravir) and related compounds: Novel broad-spectrum inhibitors of RNA viral infections. *Antiviral Res.* **82**, 95–102 (2009).
64. Binh, N. T., Wakai, C., Kawaguchi, A. & Nagata, K. Involvement of the N-terminal portion of influenza virus RNA polymerase subunit PB1 in nucleotide recognition. *Biochem. Biophys. Res. Commun.* **443**, 975–979 (2014).
65. Pielak, R. M. & Chou, J. J. Influenza M2 proton channels. *Biochim. Biophys. Acta - Biomembr.* **1808**, 522–529 (2011).
66. Samson, M. *et al.* Characterization of drug-resistant influenza virus A(H1N1) and A(H3N2) variants selected in vitro with laninamivir. *Antimicrob. Agents Chemother.* **58**, 5220–5228 (2014).
67. Shiraki, K. & Daikoku, T. Favipiravir, an anti-influenza drug against life-threatening RNA virus infections. *Pharmacol. Ther.* **209**, 107512 (2020).
68. Nair, H. *et al.* Global burden of respiratory infections due to seasonal influenza in young children: A systematic review and meta-analysis. *Lancet* **378**, 1917–1930 (2011).
69. Thompson, W. W. *et al.* Estimates of US influenza-associated deaths made using four different methods. *Influenza Other Respi. Viruses* **3**, 37–49 (2009).
70. Kuiken, T. *et al.* Host Species Barriers to Influenza Virus Infections. *Science*

(80-.). **312**, 394–397 (2006).
71. Pal, S., Santos, A., Rosas, J. M., Ortiz-Guzman, J. & Rosas-Acosta, G. Influenza A virus interacts extensively with the cellular SUMOylation system during infection. *Virus Res.* **158**, 12–27 (2011).
72. Han, Q. *et al.* Sumoylation of Influenza A Virus Nucleoprotein Is Essential for Intracellular Trafficking and Virus Growth. *J. Virol.* **88**, 9379–9390 (2014).
73. Xu, K. *et al.* Modification of Nonstructural Protein 1 of Influenza A Virus by SUMO1. *J. Virol.* **85**, 1086–1098 (2011).
74. Wu, C.-Y., Jeng, K.-S. & Lai, M. M.-C. The SUMOylation of Matrix Protein M1 Modulates the Assembly and Morphogenesis of Influenza A Virus. *J. Virol.* **85**, 6618–6628 (2011).
75. Santos, A. *et al.* SUMOylation affects the interferon blocking activity of the influenza A nonstructural protein NS1 without affecting its stability or cellular localization. *J. Virol.* **87**, 5602–20 (2013).
76. Pichler, A., Fatouros, C., Lee, H. & Eisenhardt, N. SUMO conjugation - A mechanistic view. *Biomol. Concepts* **8**, 13–36 (2017).
77. Pal, S., Rosas, J. M. & Rosas-Acosta, G. Identification of the non-structural influenza A viral protein NS1A as a bona fide target of the Small Ubiquitin-like MOdifier by the use of dicistronic expression constructs. *J. Virol. Methods* **163**, 498–504 (2010).
78. Killip, M. J., Fodor, E. & Randall, R. E. Influenza virus activation of the interferon system. *Virus Res.* **209**, (2015).
79. García-Sastre, Adolfo; Egorov, Andrej; Matassov, Demetrius; Brand,Sabine; Levy, David E.; Dubrin, Joan E., Palese, Peter; Muster, T. Influenza A Virus Lacking the NS1 Gene Replicates in Interferon-Deficient Systems. 324–330 (1998).
80. Zhao, N. *et al.* Influenza virus infection causes global RNAPII termination defects. *Nat. Struct. Mol. Biol.* **25**, (2018).
81. Jiang, L. *et al.* Internal calibration förster resonance energy transfer assay: A real-time approach for determining protease kinetics. *Sensors (Switzerland)* **13**, 4553–4570 (2013).
82. Yavuz, a. S. & Sezerman, U. SUMOtr: SUMOylation site prediction based on 3D structure and hydrophobicity. *Heal. Informatics Bioinforma. (HIBIT), 2010 5th Int. Symp.* 93–97 (2010) doi:10.1109/HIBIT.2010.5478899.
83. Green, J. R., Dmochowski, G. M. & Golshani, A. Prediction of protein sumoylation sites via parallel cascade identification. http://bioinf.sce.carleton.ca/SUMO/start.php (2006) doi:10.13140/2.1.1621.3446.
84. Zhao, Q. *et al.* GPS-SUMO: A tool for the prediction of sumoylation sites and SUMO-interaction motifs. *Nucleic Acids Res.* **42**, 1–6 (2014).
85. Malik-Chaudhry, H. K., Saavedra, A. & Liao, J. A linker strategy for trans-FRET assay to determine activation intermediate of NEDDylation cascade. *Biotechnol. Bioeng.* **111**, 1288–1295 (2014).
86. Walker, J. M. *Enzyme Engineering, Methods and Protocols*. *Life Sciences* vol. 531 (2009).
87. Fodor, E. *et al.* Rescue of Influenza A Virus from Recombinant DNA. *J. Virol.* **73**,

9679–9682 (1999).
88. Marsh, G. A., Hatami, R. & Palese, P. Specific Residues of the Influenza A Virus Hemagglutinin Viral RNA Are Important for Efficient Packaging into Budding Virions. *J. Virol.* **81**, 9727–9736 (2007).
89. Hoffmann, E., Neumann, G., Kawaoka, Y., Hobom, G. & Webster, R. G. A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc. Natl. Acad. Sci.* **97**, 6108–6113 (2000).
90. Martínez-Sobrido, L. & García-Sastre, A. Generation of Recombinant Influenza Virus from Plasmid DNA. *J. Vis. Exp.* 5–9 (2010) doi:10.3791/2057.
91. Marsh, G. A., Hatami, R. & Palese, P. Specific Residues of the Influenza A Virus Hemagglutinin Viral RNA Are Important for Efficient Packaging into Budding Virions. *J. Virol.* **81**, 9727–9736 (2007).
92. Chaimayo, C., Hayashi, T., Underwood, A., Hodges, E. & Takimoto, T. Selective incorporation of vRNP into influenza A virions determined by its specific interaction with M1 protein. *Virology* **505**, 23–32 (2017).
93. Dou, D., Revol, R., Östbye, H., Wang, H. & Daniels, R. Influenza A virus cell entry, replication, virion assembly and movement. *Front. Immunol.* **9**, 1–17 (2018).
94. König, R., Stertz, S., Zhou, Y., Inoue, A. & Hoffmann, H. Human host factors required for influenza virus replication. *Nature* **463**, 813–817 (2009).
95. Ayllon, J. & García-sastre, A. The NS1 Protein : A Multitasking Virulence Factor. 73–107 (2015) doi:10.1007/82.
96. Osula, O., Swatkoski, S. & Cotter, R. J. Identification of protein SUMOylation sites by mass spectrometry using combined microwave-assisted aspartic acid cleavage and tryptic digestion. *J. Mass Spectrom.* **47**, 644–654 (2012).
97. Nogales, A., Martinez-Sobrido, L., Chiem, K., Topham, D. J. & DeDiego, M. L. Functional Evolution of the 2009 Pandemic H1N1 Influenza Virus NS1 and PA in Humans. *J. Virol.* **92**, 1–19 (2018).
98. Tzarum, N. *et al.* Structure and receptor binding of the hemagglutinin from a human H6N1 influenza virus ACCESSION NUMBERS Atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB) under accession codes 4XKD for Taiwan2 H6 HA in apo form an. *Cell Host Microbe March* **11**, 369–376 (2015).
99. Smith, G. J. D. *et al.* Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza a epidemic. *Nature* **459**, 1122–1125 (2009).
100. Domingues, P. *et al.* Global Reprogramming of Host SUMOylation during Influenza Virus Infection. *Cell Rep.* **13**, 1467–1480 (2015).
101. Han, J. *et al.* Genome-wide CRISPR / Cas9 Screen Identifies Host Factors Essential for Influenza Virus Replication Resource Genome-wide CRISPR / Cas9 Screen Identifies Host Factors Essential for Influenza Virus Replication. *CellReports* **23**, 596–607 (2018).
102. Liao, J., Way, G. & Madahar, V. Target Virus or Target Ourselves for COVID-19 Drugs Discovery ? ― Lessons learned from anti-influenza virus therapies. *Med. Drug Discov.* **5**, 100037 (2020).

103. Martínez-Sobrido, L. & García-Sastre, A. Generation of Recombinant Influenza Virus from Plasmid DNA. *J. Vis. Exp.* 1–6 (2010) doi:10.3791/2057.
104. Pal, S., Santos, A., Rosas, J. M., Ortiz-Guzman, J. & Rosas-Acosta, G. Influenza A virus interacts extensively with the cellular SUMOylation system during infection. *Virus Res.* **158**, 12–27 (2011).
105. Way, G. *et al.* A novel SUMOylation site in the influenza a virus NS1 protein identified with a highly sensitive FRET assay. *J. Biotechnol.* **323**, 121–127 (2020).
106. Sangita Pal1, Juan M. Rosas, and G. R.-A. Identification of the non-structural influenza A viral protein NS1A as a bona fide target of the Small Ubiquitin-like MOdifier by the use of dicistronic expression constructs. **75**, 1781–1791 (2009).
107. He, X. *et al.* Probing the roles of SUMOylation in cancer cell biology by using a selective SAE inhibitor. *Nat. Chem. Biol.* **13**, 1164–1171 (2017).
108. Sun, H. *et al.* Inhibition of neddylation pathway represses influenza virus replication and pro-inflammatory responses. *Virology* **514**, 230–239 (2018).
109. Boulanger, M., Paolillo, R., Piechaczyk, M. & Bossis, G. Molecular Sciences The SUMO Pathway in Hematomalignancies and Their Response to Therapies. doi:10.3390/ijms20163895.
110. Influenza (Seasonal). https://www.who.int/en/news-room/fact-sheets/detail/influenza-(seasonal).
111. CDC. Estimated Influenza Illnesses, Medical visits, Hospitalizations, and Deaths in the United States — 2019–2020 Influenza Season | CDC. https://www.cdc.gov/flu/about/burden/2019-2020.html#references (2020).
112. Harding, A. T. & Heaton, N. S. Efforts to improve the seasonal influenza vaccine. *Vaccines* **6**, (2018).
113. Zhuang, X. *et al.* MRNA vaccines encoding the HA protein of influenza a H1N1 virus delivered by cationic lipid nanoparticles induce protective immune responses in mice. *Vaccines* **8**, 1–17 (2020).
114. Galvin, H. D. & Nutsford, A. N. Drug resistance in influenza A virus : the epidemiology and management. 121–134 (2017).
115. Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M. & Belshaw, R. Viral Mutation Rates. *J. Virol.* **84**, 9733–9748 (2010).
116. Milholland, B. *et al.* Differences between germline and somatic mutation rates in humans and mice. (2017) doi:10.1038/ncomms15183.