Pathogen host interactions: proteomics of Influenza NEP during infection reveals an antagonistic role in the formation of tight junctions

Garrett Robert Poshusta

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2015

Reading Committee:

John Aitchison, Chair

Wilhelmus Hol

Kenneth Stuart

Program Authorized to Offer Degree:

Biochemistry

University of Washington

Abstract

Pathogen host interactions: proteomics of Influenza NEP during infection reveals an antagonistic role in the formation of tight junctions

Garrett Robert Poshusta

Chair of the Supervisory Committee:
Affiliate Professor John Aitchison
Department of Biochemistry

Host-pathogen interaction networks are key to understanding the molecular mechanisms driving disease and can provide new targets for therapeutic intervention. We discovered new host factors interacting with the influenza nuclear export protein (NEP) using an engineered influenza virus expressing NEP with an N-terminal 3xFLAG tag. We collected immunopurification mass spectrometry data for 3xFLAG-NEP during an active infection and during plasmid expression in HEK293T cells. Network analysis of these complementary datasets revealed an enrichment of tight junction proteins in the NEP interactome. Expression of NEP in MDCK cells results in inhibition of tight junction formation as measured by transepithelial electrical resistance and inulin diffusion across the polarized cell monolayer. These findings reveal a new role for NEP as a tight junction antagonist.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank Dr. John Aitchison for advising and supporting the research presented in this thesis. I would also like to thank the members of my supervisory committee Dr. Alan Aderem, Dr. Kenneth Stuart, Dr. Valerie Daggett, Dr. Wilhelmus Hol, and Dr. Kenneth Stuart. I am also thankful for the encouragement, guidance and assistance provided by the members of the Aitchison and Aderem labs with whom I worked over the past 6 years. I would also like to recognize the hard work of the administrative and support staff in the Department of Biochemistry and at the Center for Infectious Disease Research. Prior to graduate school, Dr. Jennifer Shepherd, Dr. Rhea Coler and Jeff Guderian supported and encouraged my growth as a scientist and I am grateful for their mentorship.

I would also like to thank my family and friends for their unfailing support. My mom and dad, Sharrol and Paul, for fostering my curiosity from an early age. My sisters, Dayne Brown, JD, and Naysa Poshusta, DVM, for believing in me and setting an incredibly high bar. And Dr. Colin Correnti for being my partner in crime. I would also like to thank my sons, Wilder and Brooks, for giving up weekends with Dad so he can work on his pieces. Lastly and mostly I want to thank my wife of 10 years, Cathy, for going through grad school with me. The journey to a PhD is long and I took it with the best traveling companion.

Dedication

This thesis is dedicated to
Wilder, Brooks and those yet to come

This above all: to thine own self be true,
And it must follow, as the night the day,
Thou canst not then be false to any man
- William Shakespeare

And to Cathy
for your incredible support and patience

Chapter 1: Introduction

1.1 Influenza Virus

1.1.1 Overview

Influenza, the study subject of this thesis, is a member of the orthomyxovirus family of viruses that includes at present, 6 genera: *influenzaviruses A, B,* and *C*, *Thogotovirus, Isavirus* (Palese and Shaw 2007), and the recently proposed *Quarjavirus* (Presti et al. 2009). The newest member of the family, Wellfleet Bay virus, was characterized this year and is most closely related to the Cygnet River virus of the *Quarjavirus* genera (Allison et al. 2015). Orthomyxovirus genomes are composed of negative-sense single stranded RNA segments. Influenza A has 8 segments encoding 10 proteins with some strains expressing combinations of 7 additional accessory proteins (Palese and Shaw 2007; Vasin et al. 2014). The virus has many subtypes that are identified by the combination of the two viral glycoproteins studding the viral envelope, hemagglutinin (HA) and neuraminidase (NA) (eg. H1N1) (Figure 1). Currently, 16 different HA and 9 different NA variants have been reported in wild aquatic birds which are the virus' natural host and serve as a reservoir of viral genetic diversity (Alexander 2007). The recent discovery of 2 new HA and NA subtypes from highly diverged strains of influenza in bats has added an interesting new chapter to the influenza story (Tong et al. 2012; Tong et al. 2013). Phylogenetic analyses suggest that bats have also served as a longstanding and perhaps important host for influenza and calls into question previous dogma regarding wild birds as the main reservoir host. Current evidence, however, points to a low probably for zoonosis as bat strains have not been shown to productively infect human cells (Juozapaitis et al. 2014).

In order to replicate, influenza must carry out a diverse array of tasks with limited genomic coding capacity. Entry into the host cell, evasion of host antiviral responses, replication and transcription of the viral genome, nucleo-cytoplasmic trafficking of viral components, and

assembly and budding of new viral particles must all be accomplished in some cases with only 10 proteins (Figure 2). This necessitates extensive interaction with host components and also means that individual viral proteins must participate in multiple processes.

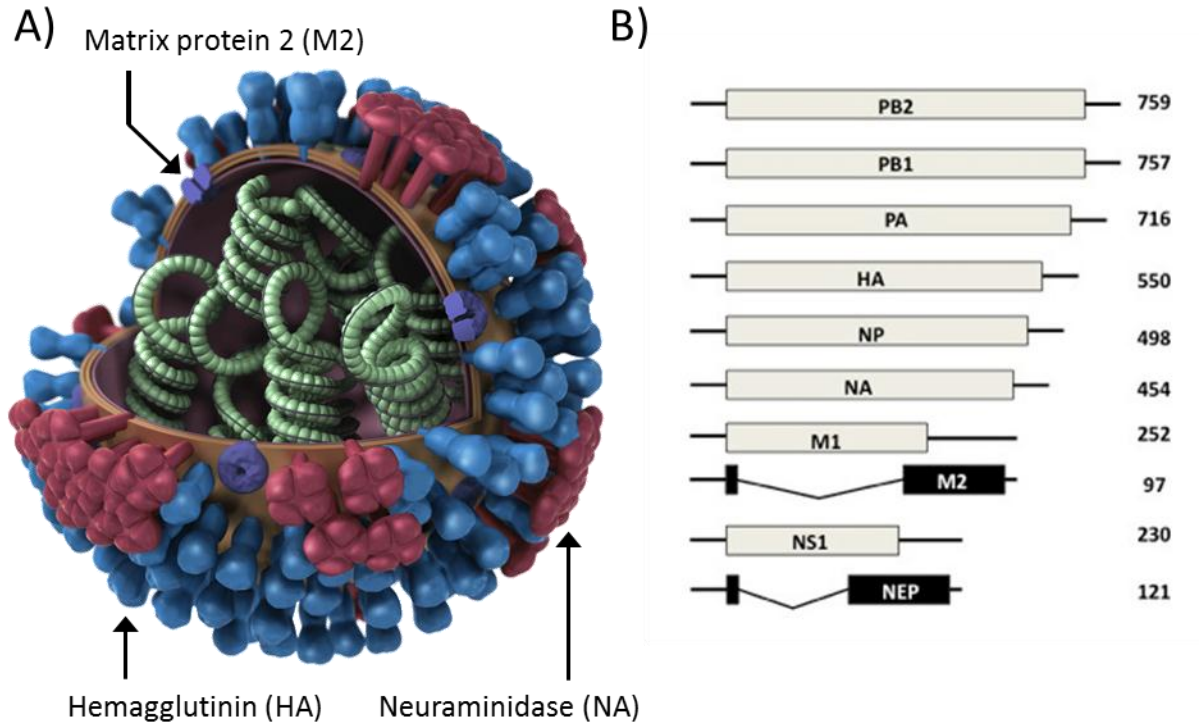


Figure 1. Schematic diagram of an influenza virion and genomic segments
In (A) the viral surface proteins hemagglutinin (HA), neuraminidase (NA), and matrix protein 2 (M2) and are shown on the surface of the virion in blue, red and purple respectively. The -ssRNA genomic segments shown in (B) exist as members of a ribonucleoprotein complex (vRNP) in the virion (green helical structures). The picture was adapted from the Centers for Disease Control.

The epidemiology of influenza is complex as the virus infects numerous animal species around the globe including agriculturally important fowl and livestock. These domesticated animals can serve as an intermediary mixing vessel since they often interface directly or indirectly (eg. use the same water source) with wild birds (Alexander, 2007; Webster et al. 1992). In a process called reassortment, shuffling of whole gene segments from two different strains during co-infection can lead to the emergence of new, immunologically novel strains

that move rapidly through human populations (Desselberger et al. 1978; Nelson et al. 2008; Smith et al. 2009; Vijaykrishna et al. 2010). As a result, strains currently circulating in humans are able to access the deep genetic diversity present in the avian reservoir much to our detriment. Indeed the most deadly disease outbreak in the history of mankind, the Spanish influenza of 1918, is attributed to an influenza virus resulting from reassortment between an avian strain and a strain already circulating in humans (Johnson and Mueller 2002; Worobey, Han, and Rambaut 2014).

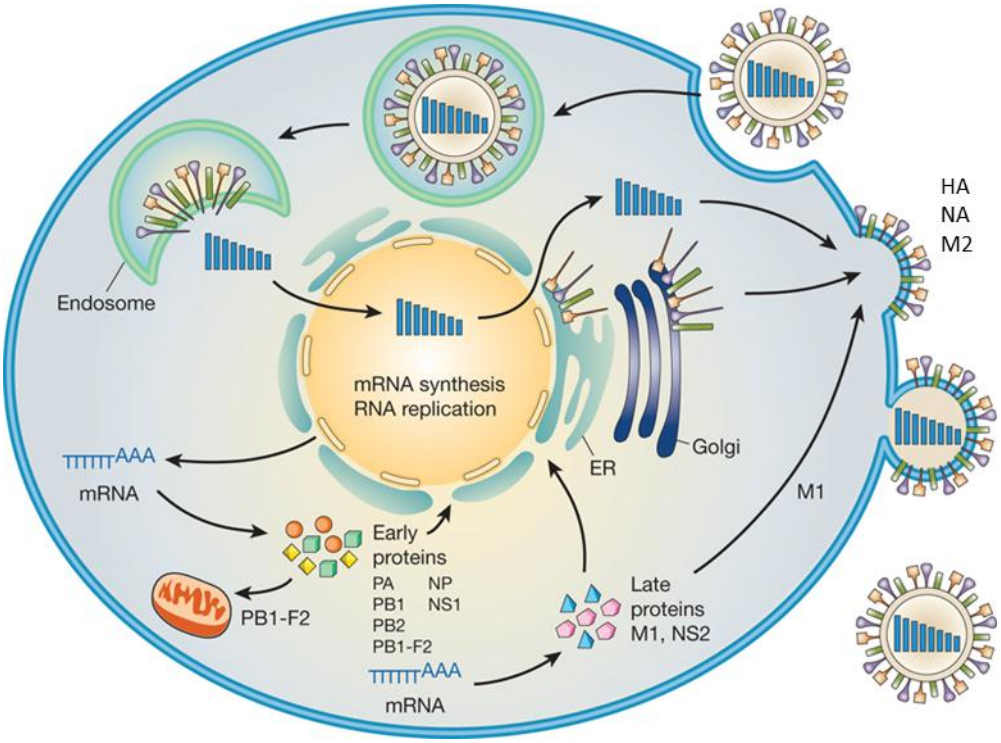


Figure 2. Schematic diagram of the influenza life cycle
Entry into the host cell is initiated by interaction of HA with sialic acids on the cell surface. Following internalization, acidification of the endosome leads to fusion of the viral and host lipid membranes and release of the vRNPs into the cytosol. The vRNPs are actively transported into the nucleus where transcription and replication of the viral genome are carried out by the viral polymerase complex proteins (PA, PB1, PB2). NS1 antagonizes host defenses. PB1-F2 localizes to mitochondria and activates apoptosis. NP is the scaffold protein for vRNPs. M1 and NEP form an export complex with vRNPs and host factors for egress from the nucleus. NEP also regulates the polymerase complex. Figure adapted from Neumann et al. 2009.

Antigenic shift, where an immunologically novel HA is introduced into a human influenza strain by reassortment, is also responsible for two other pandemics in the 20th century (Nguyen-Van-Tam 2003). Once a strain with a new HA becomes wide-spread, it may evolve into a seasonal influenza strain as happened after the H3N2 "Hong Kong Flu" pandemic of 1968-69. Seasonal influenza strains undergo antigenic drift due to immunoselective pressure on HA which is the major target of the antibody response. Since influenza is constantly evolving to evade the immune response, the vaccine must be reformulated on an annual basis to protect against the drifted strains (Gatherer 2009). Despite the availability of a vaccine, seasonal influenza is a major global health burden infecting up to 10% of adults and resulting in about 500,000 deaths annually (Stöhr 2002).

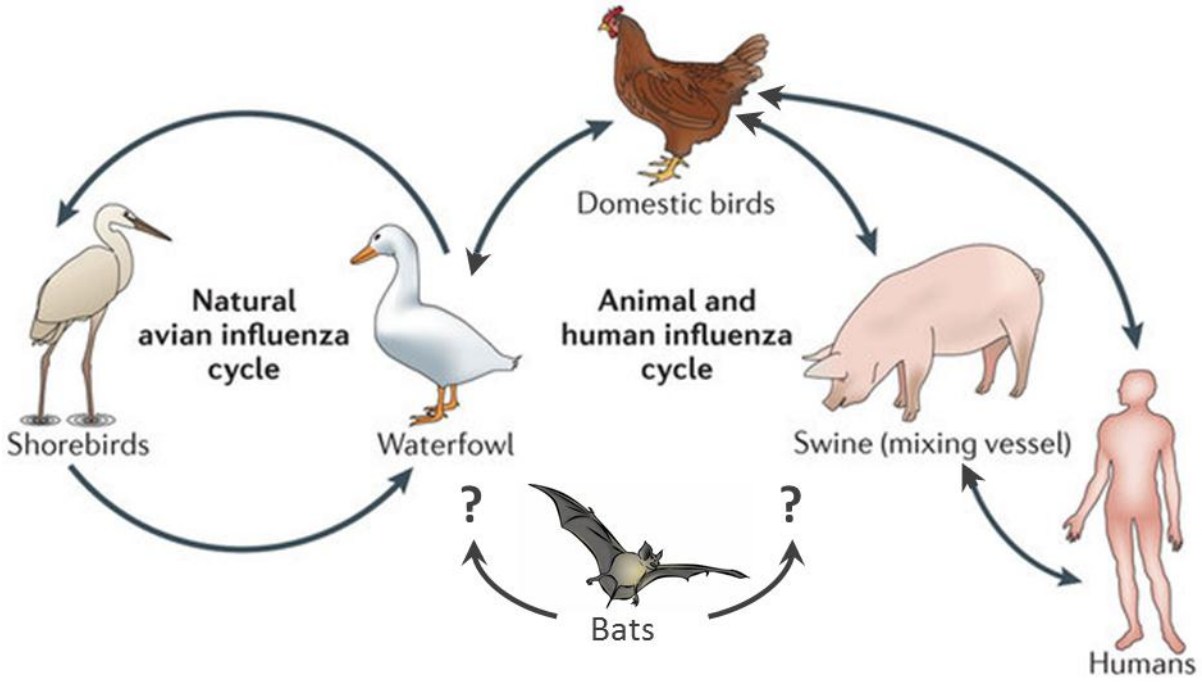


Figure 3. The ecological cycle of influenza A viruses
The natural hosts of influenza A viruses are wild aquatic birds and infection is typically asymptomatic or results in mild disease. The virus infects domesticated poultry when they share space with wild birds and adaption to this new host can lead to high virulence. Pigs can also be infected with several influenza strains. These intermediary hosts serve as mixing vessels by facilitating reassortment between human and avian viruses. At this time, the ecology of the bat influenza viruses is unclear. Figure was adapted from Shi et al. 2014.

1.1.2 Avian Influenza

Some of the important outstanding questions for influenza researchers relate directly to the ongoing process of viral evolution in different hosts that can lead to highly variable outcomes upon infection (Alexander, 2007; Caron et al. 2009; de Wit et al. 2010; Ducatez, Webster, and Webby, 2008; Ito, 2000; Webster et al. 1992; Webster et al. 2007). Since wild birds are the viruses' natural host and the presumed origin for all currently circulating strains, I will briefly discuss avian influenza in context of recent zoonosis with high pathogenicity. The emergence of highly pathogenic influenza (HPAI) in the 1990's upended the traditional model of unidirectional flow of viral strains from wild birds into other species. In wild birds, influenza is a gastrointestinal virus spread through feces and infection typically results in very mild or asymptomatic disease, a situation proposed to be the result of long-coevolution (Webster et al. 1992). The sustained circulation of HPAI strains in wild bird populations and the ensuing die-offs resulting from outbreaks, have introduced new elements into our understanding of the ecology of influenza. Phylogenetic analyses suggest that HPAI strains (currently comprising H5, H7 and to a lesser extent H9 strains) arise from low pathogenicity (LPAI) precursors during adaptation to new hosts such as poultry(Alexander 2007; Banks et al. 2001) and that these HPAI strains can infect and circulate in wild birds causing significant morbidity (Ellis et al. 2004; J. Liu et al. 2005; Sturm-Ramirez et al. 2004). HPAI strains have also spilled-over into humans with deadly consequences. H5N1 influenza happens to be highly pathogenic in both birds and humans with the first reported zoonosis in 2003. Since then (through April 2015), the World Health Organization (WHO) reports a total of 840 cases of H5N1 with an overall case fatality ratio (CFR) of 53%. However, it is important to note that the designation of an influenza strain as "highly pathogenic" refers only to disease in birds, particularly poultry (Alexander 2007; Caron et al. 2009) and human infection with LPAI strains can also result in a high CFR. The first human infections with LPAI H7N9 were documented in 2013 (Gao et al. 2013) and the CFR thus far is slightly lower than H5N1's at 37% (212 deaths from 579 documented infections (WHO)). Since infected poultry are often asymptomatic, outbreaks of H7N9 are more difficult to track than those of H5N1 which result in rapid death (Uyeki and Cox 2013). It remains an open question whether either of these strains will gain the ability to sustainably transmit between human

hosts resulting in a pandemic or if the next pandemic strain will be completely novel. So far human infection with HPAI or LPAI is closely associated with exposure to infected poultry and while cases of human-to-human transmission have been documented, they remain an exception rather than a rule. Other than creating a universal influenza vaccine, the greatest challenge in the field remains predicting the virulence and transmissibility of novel strains based on genetic sequence, and H5N1 and H7N9 top the WHO's threat list for pandemic potential (Pena et al. 2013). New strains of influenza will continue to emerge and early information about their potential virulence is key to implementing appropriate response measures.

Considerable effort has been directed toward understanding the molecular determinants underlying the viral properties of influenza. Host tropism, transmissibility, and pathogenicity are all important characteristics that are incompletely understood and are the result of interactions between multiple viral and host factors (Shi et al. 2014). These interactions have been studied using high throughput technologies leading to a more comprehensive understanding of the components utilized by influenza and how these components affect viral properties (Brass et al. 2009; Hao et al. 2008; Karlas et al. 2010; König et al. 2010; Shapira et al. 2009; Sui et al. 2009; Watanabe et al. 2014). However, poor overlap between screens suggests that no single study is comprehensive and that many interactions are missed due to the limitations inherent to any single experiment (T. Watanabe, Watanabe, and Kawaoka 2010).

1.1.3 Influenza Nuclear Export Protein

Influenza segment 8 encodes two proteins, non-structural proteins 1 and 2 (NS1/2). NS2 is the alternatively spliced product and is also referred to as the nuclear export protein (NEP) due to its role in the egress of influenza ribonucleoparticles (vRNP) from the nucleus. The vRNP consists of multiple copies of nucleoprotein (NP), a scaffolding protein that coordinates the genomic RNA into an elongated helical particle, a single negative-sense ssRNA, and the polymerase proteins (polymerase basic 1 and 2, and polymerase acidic (PB1, PB2, and PA)) (Palese and Shaw 2007). NEP is a small protein containing 121 amino acids and is able to

passively diffuse through the nuclear pore (Gao et al. 2014). The N-terminus is proteolytically sensitive while the C-terminal residues 59-116 have been crystallized and form a stable helical hairpin (Akarsu et al. 2003). Figure 4 shows a schematic representation of NEP and the location of functionally significant epitopes. O'Neill et al. (1998) first ascribed the nuclear export function to NEP based on observations that it interacted with nucleoporins in a Y2H assay and that NEP fusions could facilitate the export of heterologous proteins. They also identified the first of two consensus nuclear export signals. The precise composition of the export complex is still under investigation but it has been established that NEP is essential for this process and interacts with both the host export factor CRM1, (Huang et al. 2012; Neumann, Hughes, and Kawaoka, 2000) and components of the vRNP (matrix protein 1 (M1) (Akarsu et al. 2003; Bui et al. 2000; Shimizu et al. 2010), and polymerase basic protein 2 (PB2) (Mänz et al. 2012) in a bridging fashion. NEP has also been shown to regulate the activity of the viral polymerase complex (Bullido, Gomez-Puertas, and Saiz 2001; Mänz et al. 2012; Reuther et al. 2014; Robb et al. 2009). Influenza must produce three different RNA species during infection: viral mRNAs (vmRNAs), a complementary RNA intermediate (cRNA), and genomic replication products (vRNA). *In vitro* assays indicate that in the presence of NEP, the abundance of the viral RNA species produced by the polymerase complex more closely match those found during infection (less vmRNA, and more vRNA) (Robb et al. 2009). Recent work indicates that the aforementioned interactions and functions are interrelated. Chua et al. propose that NEP functions as a molecular timer whereby gradual accumulation of NEP leads the polymerase complex to produce more vRNA in preparation for vRNP formation and export late in the infectious cycle (Chua et al. 2013). Brunotte et al. (2014) demonstrated that NEP stabilizes the interaction of M1 with vRNP using the same epitope (C-terminal 3 residues) associated with polymerase-stimulatory activity and proposed a refined model of the vRNP export complex shown in Figure 5.

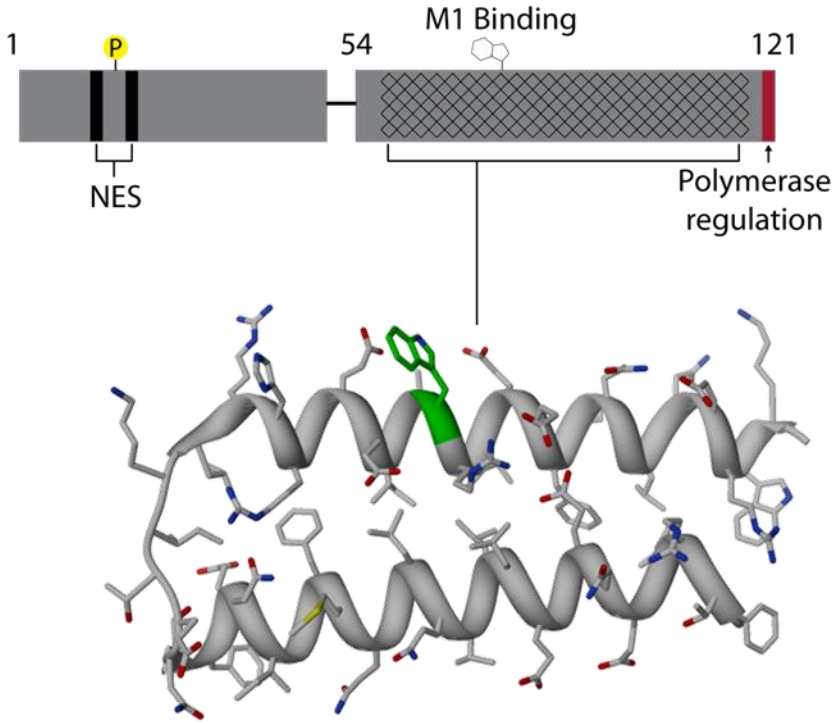


Figure 4. Schematic diagram of influenza NEP
The N-terminal domain is flexible and subject to proteolysis. It contains a triple serine motif (S23-25) that is phosphorylated between two NES signals (12-ILMRMSKMQL-21, and 31-MITQFESLKL-40). A crystal structure of the C-terminal domain shows a helical hairpin. The C-terminal domain contains interaction sites for M1 and PB2. A key residue for interaction with M1, W78, is highlighted in green.

These studies are important because targeting host factors such as CRM1 is a promising antiviral strategy (Perwitasari et al. 2014). Additionally, NEP plays a role in the adaptation of avian influenza strains to mammalian hosts. One widely observed mutation in human isolates of avian influenza is PB2-E627K (Hatta et al. 2001; Song et al. 2009; Song et al. 2014; Steel et al. 2009) which results in more favorable enzyme kinetics at lower temperatures of the respiratory tract (Aggarwal et al. 2011; Massin, van der Werf, and Naffakh, 2001). However some isolates from fatal human infections including KAN-1 (A/Thailand/1(KAN-1)/2004), retain the avian glutamate residue and are still highly pathogenic in mammals. Manz et al. (2012) showed that a single change (M16I) in the KAN-1 NEP could rescue polymerase activity in mammalian cells. Follow-on work identified increased conformational flexibility at lower temperatures for NEP-

M16I as the reason for the enhancement of polymerase activity (Reuther et al. 2014). The duck body temperature is approximately 41°C. In contrast, the human upper trachea averages 32°C while the subsegmental bronchi average 36°C (McFadden et al. 1985). So not only does the virus have to adapt to interact with mammalian proteins, it must also adapt to a new thermal environment. It remains an outstanding question whether the sequence/function relationships described above are the result of direct interactions between the viral proteins or are influenced by the recruitment of host factors.

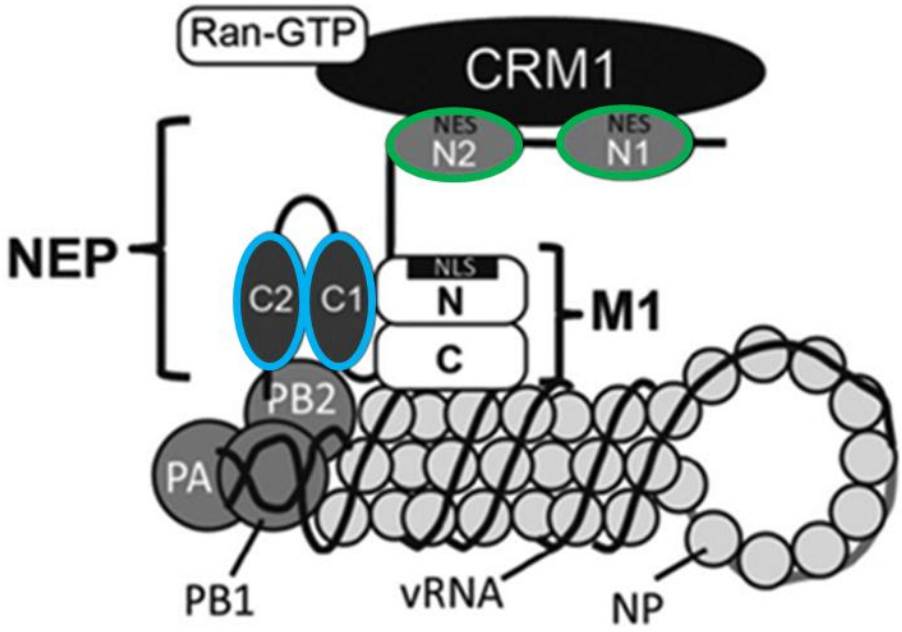


Figure 5. Graphic model of the influenza vRNP export complex
NEP links the vRNP to the host export factor CRM1. The N-terminus of NEP (outlined in green circles) interacts with CRM1 through two nuclear export signals (NESs). The C-terminal helical hairpin (highlighted in blue) stabilizes the interaction of M1 with the vRNP and binds to the polymerase complex subunit PB2. Figure adapted from Brunotte et al. 2014.

Host interaction partners of NEP have been discovered using yeast two-hybrid (Y2H) and immunopurification mass spectrometry (IP-MS) approaches (Table 1). NEP has been shown to interact with Nup98 but the role this plays in the viral lifecycle has not been identified. Gorai et al. (2014) used FLAG-tagged NEP to identify new host interactors of NEP including ATP5B. This

interaction in turn has implicated NEP in the process of viral budding which represents a new role for NEP during infection. A recent Y2H study found an AIMP2-NEP interaction that promotes SUMOylation of M1 (Gao et al. 2015). This modification on K242 of M1 stimulates nuclear export of the vRNP complex (Wu, Jeng, and Lai 2011). Knowing the interaction partners of NEP has contributed to our understanding of how influenza is replicating in host cells and we wanted to expand the catalogue of NEP interactors using sensitive quantitative proteomics approaches. With only one exception NEP interactors have been identified sans infection. We developed a system for introducing tags on the NEP N-terminus that are genetically encoded in an engineered influenza strain. This allows high throughput interaction discovery techniques like IP-MS to be applied during the course of an active infection and is presently the only methodology that uses NEP exclusively as bait in this context.

Table 1. Previously discovered host proteins interacting with NEP

Host ID	Method	Localization	Viral Process	Reference
CRM1	Y2H	Nucleus	Nuclear export	(ONeill et al. 1998)
ATP5B	IP-MS	Plasma membrane	Budding	(Gorai et al. 2012)
Nup98	Y2H	Nucleolus	Unknown	(Chen et al. 2010)
AIMP2	Y2H	Cytosol	Sumoylation	(Gao et al. 2015)

Abbreviations in the "Method" column: yeast two-hybrid (Y2H) and immunopurification mass spectrometry (IP-MS).

1.2 Systems Biology

Living organisms exist as a result of the emergent properties of complex systems of interactions at many different scales—atoms on up to cells and organisms. Systems biology embraces this theme and seeks to understand, predict, and control biological phenomena through an iterative process of experimentation, data integration, modeling, and hypothesis generation. As a formal field of study, systems biology is relatively new but mathematical modeling of biological processes has been explored for over a century starting with the quantitative modeling of enzyme kinetics in the early 1900's. One of the first mathematical models describing a physiological process (action potentials in axons) was presented by Alan Lloyd Hodgkin and Andrew Fielding Huxley in 1952 (Hodgkin and Huxley 1952), and systems biology was highlighted as a distinct field by Mihajlo D. Mesarović at the 3rd Systems Symposium in 1966. However, systems biology remained a minor field until new technologies enabled the collection and processing of larger, more comprehensive datasets around the turn of the century. Systems approaches have yielded valuable knowledge in diverse areas including developmental biology (Peter and Davidson 2011), structural biology (Alber et al. 2007), cell biology (Saleem et al. 2008), immune function (Gilchrist et al. 2006) and host-pathogen interactions (Carpp et al. 2014). Key features of these studies are the collection of quantitative high throughput datasets, integration of multiple data types, and visual and mathematical representations to guide experimentation (Figure 6). The end goal is a predictive model that facilitates rational control of the structure or process.

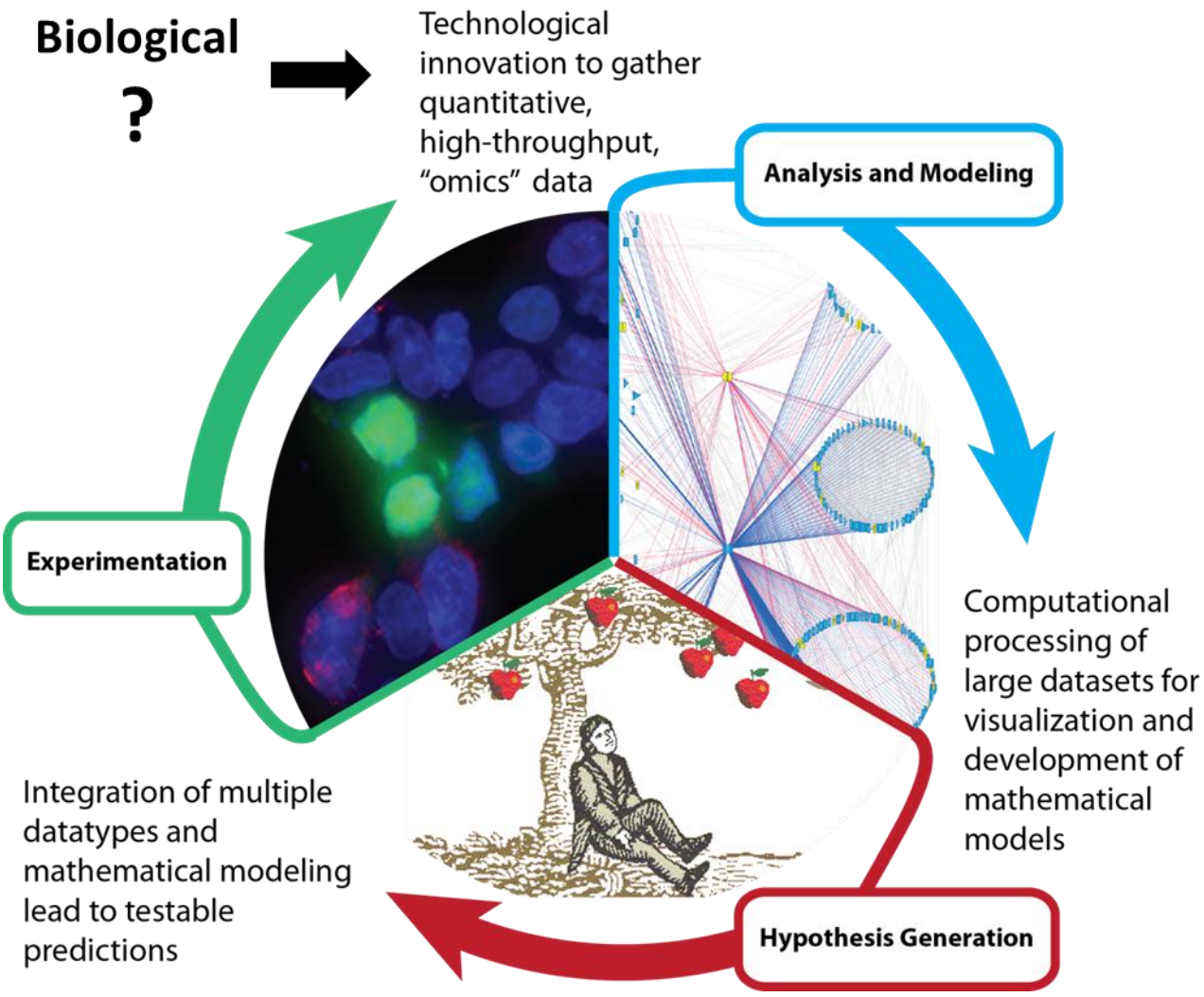


Figure 6. The systems biology cycle
Biological inquiry drives the development of new technologies which in turn enable quantitative measurement of phenotypes under many conditions. Models representing these large scale datasets lead to new hypotheses and then experimental validation. Subsequent iterations of the cycle are informed by the previous results. The goal is a predictive model that facilitates control of the process under investigation.

In the present work we applied the principles of systems biology to study influenza. Our approach was driven by the desire to comprehensively catalogue which host proteins physically interact with influenza proteins. This knowledge is an essential starting point for systems studies the results of which can lead to new therapeutic strategies and also new insights into cell biology.

Our strategy for identifying new host factors interacting with NEP employs a high throughput, quantitative, mass spectrometry-based discovery technology—I-DIRT. Isotopic differentiation of interactors as random or targeted (I-DIRT) was developed in yeast and requires labeling cellular proteins with isotopically heavy amino acids *in vivo.* It also relies on the expression of a tagged version of the target protein and so we engineered the influenza virus to express a tag on the NEP during infection. We used the recombinant virus in IP-MS experiments to identify host proteins interacting with NEP. We produced a network model of these data which showed an enrichment of tight junction proteins. NEP had not previously been reported to associate with tight junction proteins nor had it been shown to affect tight junction functioning. We subsequently tested whether the presence of NEP had an effect on tight junctions and found that expression of NEP led to impaired functioning. Thus, this work represents one full turn of the cyclical process represented in Figure 6 and demonstrates the utility of the systems biology approach.

1.3 Methodologies for Studying Host Pathogen Interactions

The ability to visualize and subsequently identify proteins is essential to studying interactions between hosts and pathogens and is a key starting point for systems biology investigations. Scientists have been developing technologies to separate, visualize and identify proteins for over a century. Very early studies used colorimetric reactions to quantify the presence of different amino acids in protein preparations (Folin and Ciocalteau 1927; McChesney and Swann 1937). In the 1930's Arne Tiselius developed electrophoresis, the separation of charged components with an electric field, for the study of biomolecules (Tiselius 1930). For his work Tiselius was awarded a Nobel Prize in 1948, and the modern implementation of the technique, polyacrylamide gel electrophoresis (PAGE), is fundamental for studying proteins still today. Around this same time antibody technologies, including labeling and immunopurification, were being developed. Labeling antibodies with colorimetric and then fluorescent compounds provided a way to mark antigenic proteins in tissues (Coons et al. 1942). This work culminated in 1950 with the publication of work by Coons and Kaplan (1950) that established the foundation of immunofluorescence, another essential technique widely used to visualize the localization and co-localization of specific proteins. With these technologies in hand, scientists could now begin to look at the proteins of pathogens.

At first these techniques were exploited for the characterization of whole bacteria and viruses (Evans and Kingsbury, 1969; Harris and Kline, 1956). Differential protein bands were detected in gels comparing strains of bacteria and were proposed as means of taxonomic classification (Cann and Willox 1965; Fowler et al. 1963; Lund 1965; Smithies 1955). Similarly, fluorescent antibodies directed against viruses or bacteria were used as a means of detection allowing for both clinical diagnostics (Biegeleisen et al. 1966; Thomason, Cherry, and Edwards 1959), and localization studies of pathogens and their components in cells and tissues (Nayak et al. 1964). The ability to study proteins and their interactions took a great step forward with the development of the western blot by Harry Towbin (Towbin, Staehelin, and Gordon 1979). This allowed scientists to readily detect very small amounts of individual proteins present in complex mixtures such as cell lysates and immunopurifications. Western blots were employed

broadly to immunogenically profile pathogens (Biggar et al. 1985; Gulig and Hansen 1985; van Wyke et al. 1984) and were also used to study the interactions between proteins (Akerström et al. 1985; Lane et al. 1985). These techniques—immunopurification, immunofluorescence, gel electrophoresis, and western blotting—are still foundational for studying proteins and their interactions, but their scalabilities are limited and new technologies were needed to expand the throughput of interaction studies.

Phage display and the yeast two-hybrid (Y2H) system were developed in the 1980's to systematically interrogate thousands of possible interactions with any given target protein/peptide. Phage display technology was first detailed by George Smith in 1985, and has the important characteristic of evolvability (Smith and Petrenko 1997). When a library of peptide sequences is expressed as a fusion with the phage coat protein, the heterogeneous phage population can be screened for interaction with a target. Phages displaying peptides that interact with the target can then re-infect bacteria to create many clones which are again subjected to selection. The breakthrough represented by this technology is the physical linkage of phenotype and genotype thus enabling identification of a single high affinity interaction from billions possible in the library. Phage display has been used extensively to discover medically useful interactions especially through the display of antibody variable regions. In one example, Kashyap et al. (2008) prepared a library of single-chain variable fragments (scFVs) and fragment antigen-binding (Fabs) sequences from survivors of an H5N1 influenza outbreak and screened the displayed antibody regions against H5N1 virions. They were able to identify over 100 monoclonal antibodies (mAbs) binding to the influenza surface protein hemagglutinin, three of which were cross neutralizing with other influenza strains. Such cross neutralizing antibodies are key to broadly protective vaccines that generate immunity against multiple viral strains. Phage display has also been used to study host-pathogen interactions. Lauterbach et al. (2003) identified *Plasmodium falciparum* proteins interacting with human erythrocyte spectrin and protein 4.1 using a *P. falciparum* cDNA bacteriophage library. Random hexamer phage libraries have also been used to identify interaction sites between pathogens and host proteins (Chaemchuen et al. 2011). While phage display is a powerful technology, notable limitations include size constraints for displayed sequences, issues with expression of certain proteins

especially in the case of genome-wide cDNA libraries and the *in vitro* binding conditions for selection which may not accurately reflect the *in vivo* environment necessary for some interactions.

Like phage display, the yeast two-hybrid technique was developed as a way to detect and study protein-protein interactions and it also offered a way to systematically interrogate thousands of possible interactions with any given target protein/peptide. The system was first described by Stanley Fields and Ok-kyu Song in 1989 and is founded on the separable nature of the functions of a transcription factor—DNA binding and transcriptional activation (Fields and Song 1989). Fields and Song based their assay on the GAL4 protein, a yeast regulatory factor essential for growth on galactose. It has both DNA-binding and activating domains that, only when co-localized, lead to transcription of downstream genes. Co-expression of the independent GAL4 domains themselves is not sufficient to induce transcription. By fusing a known target protein to one domain and a query protein to the other domain, transcription occurs only if the two proteins interact to bring the GAL4 domains into close proximity. The transcription product generally encodes a reporter protein or selectable marker so that yeast expressing the two interacting proteins can be readily identified. In this manner whole cDNA libraries can be screened for interaction with target proteins. Though the technique has been widely adopted, it is characterized by high false positive rates, perhaps even greater than 50% (Goldberg and Roth 2003). Protein-protein interactions are concentration dependent and expression levels of the GAL4 domain fusions may be higher (or lower) than those for the native proteins. These artificial conditions can enhance spurious interactions between proteins that would otherwise be spatially segregated *in vivo.* False negatives may occur if the fusion does not behave or fold properly or does not localize to the nucleus. These drawbacks notwithstanding, Y2H is a powerful tool for discovering new interactions and researchers have applied it to study both intra-pathogen interactions (LaCount et al. 2005; Rain et al. 2001; Uetz et al. 2006) and also interactions with the host (Chen et al. 2010; de Chassey et al. 2013; Elton et al. 2001; Huang et al. 2009; Khadka et al. 2011; Lee et al. 2010; X. Liu et al. 2009; Mitzner et al. 2009; Reinhardt and Wolff, 2000; Shapira et al. 2009). Since viruses have relatively small genomes, they interact extensively with cellular machinery making them attractive targets for

Y2H host-pathogen studies. Additionally, entire proteomes for some viruses can be screened since they only encode tens of proteins, as has been done for vaccina virus, hepatitis C virus, dengue virus, Epstien-Barr virus and influenza (de Chassey et al. 2008; Le Breton et al. 2011; Shapira et al. 2009; Zhang et al. 2009).

Immuno or affinity purification with shotgun mass spectrometry (IP/AP-MS) is a more recently developed technology for discovering new protein-protein interactions. Conceptually the technique is straight forward. An antibody or other capture agent coupled to a solid support is used to pull out a known protein (a "bait") from a complex mixture (typically a cell lysate). The lysate is washed away from the solid support leaving the target bait bound to the antibody as well as any other interacting proteins ("prey"). The proteins are eluted from the antibody and identified by mass spectrometry. An example workflow is shown in Figure 7. A key advantage of this technique over phage display and Y2H is that native interactions, including whole complexes established *in vivo*, can be preserved through the IP procedure and need not be reconstituted in an artificial environment. Also proteins are expressed under native conditions allowing for proper post-translational modification which may be necessary for some interactions. In the case of host-pathogen studies, the ability to capture interactions directly from active infections is a unique feature of this methodology. The cellular environment can be drastically altered during disease and some interactions, especially those with the immune system, will only be observed during an infection.

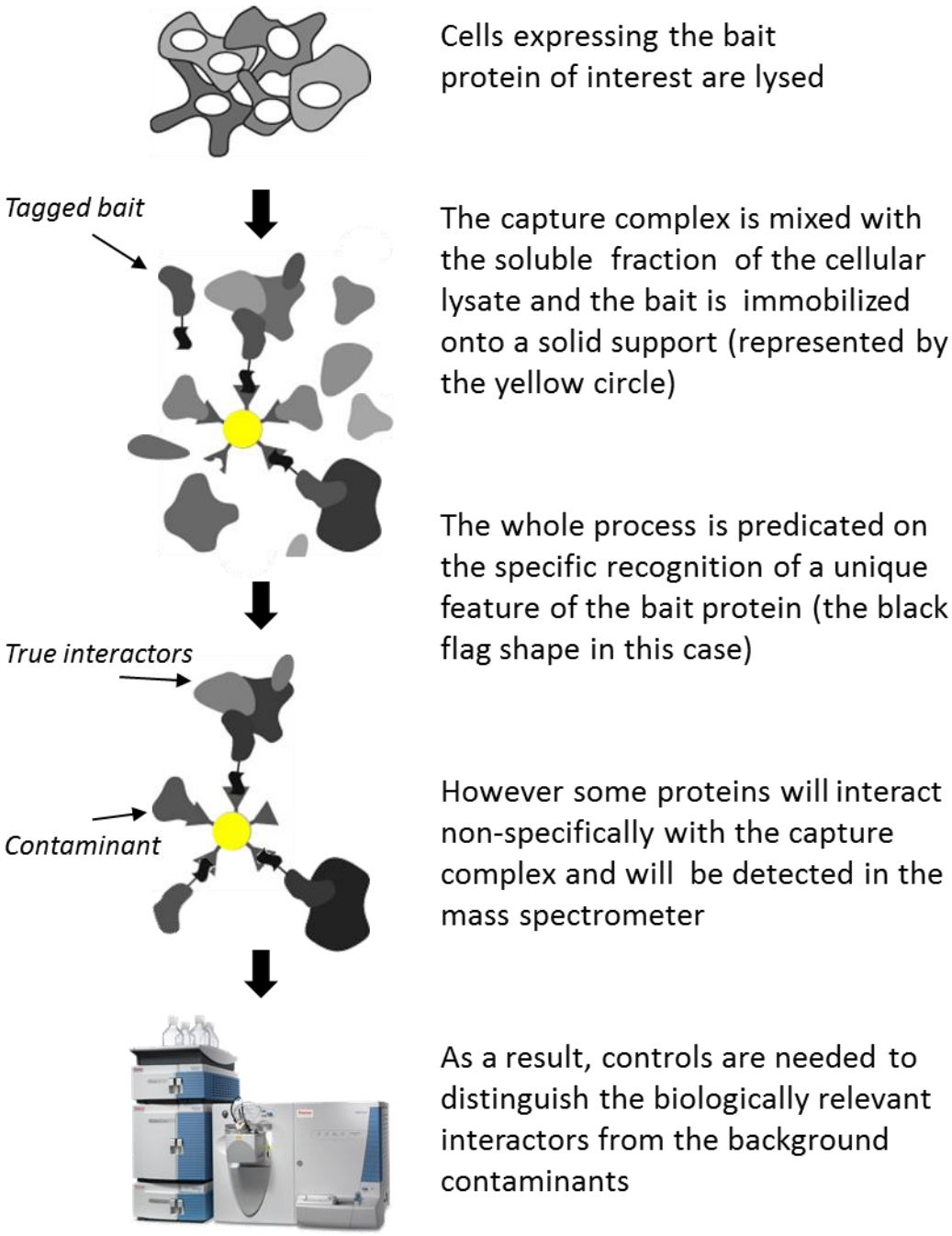


Figure 7. Workflow for an IP/AP-MS experiment

Proteins in solution will associate non-specifically with the capture complex used for IP experiments so controls must be used to distinguish contaminants from true interactors.

Mass spectrometry (MS) has been around for nearly 100 years but it was not used to study large biomolecules until the invention in the 1980's of modern ionizing techniques. Matrix-assisted laser desorption-ionization (MALDI) and electrospray ionization (ESI) gave scientists powerful new tools for the detection and structural identification of whole-proteins and peptides (Fenn et al. 1989; Karas and Hillenkamp, 1988). The technology has experienced rapid improvement over the last two decades as methods for the identification of peptides from their mass spectra evolved in parallel with increasing computational power. Today liquid chromatography tandem mass spectrometry (LC MS/MS) is widely used to identify proteins present in complex mixtures like cell extracts and immunopurifications—generally referred to as proteomics. Datasets encompassing whole proteomic surveys of model organisms are now available (Brunner et al. 2007; Gavin et al. 2006; Krogan et al. 2006; Schrimpf et al. 2009) and a draft of the human proteome was recently published that contains over 1 billion peptide spectrum matches that map to 18,097 of 19,629 annotated human genes (Wilhelm et al. 2014).

Since mass spectrometry is highly sensitive, hundreds or thousands of proteins can be identified in a single IP-MS experiment and as a result, the major drawback of the IP-MS approach is that it can be difficult to distinguish truly interacting proteins from contaminants that are binding non-specifically (Nesvizhskii 2012; Trinkle-Mulcahy et al. 2008).The simple approach of subtracting proteins identified in a control IP from those identified in the target IP can work in some cases but lacks quantitative information about peptide/protein abundance which is essential for establishing dynamic system properties and guiding selection of interaction candidates for follow-up studies. Proteomic measurements by mass spectrometry are not inherently quantitative because the heterogeneous physiochemical properties of peptides lead to differential behavior in the mass spectrometer. As a result, the signal intensity of two different peptides present in equal amounts can vary significantly. Techniques developed to address this problem fall into two categories: those that use isotopic labels to derive quantitative information about protein abundance and label-free strategies that extract quantitative information from spectral counts or peptide signal intensities. Generally the isotopic labeling strategies give higher accuracy while the label-free methods yield higher coverage (i.e. more identified peptides) and more dynamic range (Bantscheff et al. 2012). One

example of a label-free approach is the significance analysis of interactome (SAINT) computational tool (Choi et al. 2011). SAINT establishes a probability that any given bait and prey are true interactors by modeling the spectral counts for each pair with a mixture distribution and is especially powerful for label-free datasets that include multiple baits and purifications which boost the robustness of the model. We used isotopic labeling for quantitation in this work since we were focused on a single bait and desired maximal accuracy from few IP-MS runs.

Isotopic labeling strategies utilize the mass shift inherent to peptides containing heavy isotopes to distinguish them from the otherwise identical "light" version. Since each heavy and light peptide pair behaves identically in the mass spectrometer, their signal intensities can be used for relative quantification. Utilization of isotopes as "internal calibrants" for proteomics was first reported by three groups in 1999 and the strategies differ mainly based on *in vitro* or *in vivo* incorporation of the isotopes (Gygi et al. 1999; Oda et al. 1999; Paša-Tolić et al. 1999). The isotope-coded affinity tag (ICAT) was the first of several tag-based approaches that chemically modify proteins *in vitro* with an isotopically heavy or light affinity tag linker (Gygi et al. 1999). Labeling with tags continues to be a useful technique but it does have notable drawbacks. Since tag-based approaches rely on specific reactive groups the detected peptides will be biased for those groups limiting coverage and side reactions can degrade data quality. Additionally, initial steps prior to mixing the heavy and light labeled proteins must be carefully controlled to minimize variability that could influence downstream quantitation.

In contrast to the tag-based labeling, isotopes can be incorporated into proteins *in vivo* by providing isotopically "heavy" food, in the case of whole animals, or "heavy" growth media for cell cultures. Stable incorporation of labeled amino acids in cell culture (SILAC) is a popular methodology for generating cells that have heavy L-lysine ($^{13}C_6$, $^{15}N_2$) and L-arginine ($^{13}C_6$, $^{15}N_4$), incorporated into almost all proteins and this was the technique we used to generate heavy HEK293T cells for our proteomics work (Ong 2002). Proteins with heavy arginine and lysine are particularly advantageous for proteomics experiments that utilize a tryptic digest as each peptide will contain a heavy arginine or lysine at the N and C termini making spectrum

matching more tractable. *In vivo* labeling is advantageous because samples from the conditions under investigation are mixed at the earliest possible point, even prior to cell lysis. As a result, any experimental error associated with downstream processing and data collection affects the heavy and light peptides equally. The high degree of internal control inherent to this design should enable the confident detection of proteins that are less abundant and that would be difficult to quantify by other methods.

The availability of comprehensive proteomics data has lagged behind transcriptional and genomic data and now that these data can be readily collected, fundamental insights into biological processes are being revealed (Cravatt, Simon, and Yates 2007). In one example that challenges prior dogma, a series of recent papers have used new statistical analyses of proteomic and transcriptomic data to show that protein abundance is controlled primarily at the level of transcription rather than translation (Battle et al. 2015; Jovanovic et al. 2015; Li and Biggin 2015). Proteomics has also provided important information about how pathogens interact with and affect host cells (Carpp et al. 2014; Jäger et al. 2011; Selbach et al. 2009; Shui et al. 2009; Tawaratsumida et al. 2014; Weekes et al. 2014; York, Hutchinson, and Fodor 2014) and these insights can directly inform vaccine and drug development strategies (Dyer, Murali, and Sobral 2008; Roy et al. 2009). Analysis of multiple protein-protein interactions from diverse pathogens indicates that regulatory host proteins at the center of multiple connected processes are preferentially targeted (Franzosa and Xia 2011). This is not surprising since many pathogens, especially viruses, have a limited number of proteins with which to execute their replicative programs.

Influenza is one of these viruses with small genomes and few proteins that, despite years of vigorous research effort to better understand the virus and to develop therapies and vaccines, still significantly impacts human health. We utilized quantitative proteomics to gain further insight into the pathophysiology of influenza infection. In order to facilitate rapid experimentation under biosafety level 2 containment, this work was carried out using an attenuated strain of influenza A/Puerto Rico/8/1934(H1N1) (PR8) (Beare, Schild, and Craig 1975). We produced recombinant PR8 virus with a tag on the influenza protein NEP. IP-MS

experiments identified new host interactors for NEP. Based on the enrichment tight junction proteins in the NEP interactome we hypothesized that NEP is affecting cellular tight junctions and carried out experiments to test whether the presence of NEP has a functional consequence for tight junction integrity. Our assays indicate that NEP does antagonize the barrier function of tight junctions pointing to a new role for this influenza protein during infection.

Chapter 2: Discovering new host interactors of NEP using IP-MS

2.1 Summary

Our strategy for discovering new host interactors of NEP is modeled after an approach first developed in yeast called I-DIRT (isotopic differentiation of interactors as random or targeted). I-DIRT is an IP-MS strategy that uses isotopic labeling of proteins *in vivo* to distinguish contaminants from stable interactors. In order to conduct an I-DIRT experiment a tagged version of the target "bait" protein must be available and can be expressed from a plasmid or may be encoded in the genome. In our case genomic expression of the tagged bait has a distinct advantage since we are assured essential functionality is preserved if infectious virus is recovered. We overcame several challenges related to creating influenza virus that expresses tagged NEP during infection and produced a virus suitable for quantitative proteomics studies.

2.2 Introduction

2.2.1 Control strategies for IP-MS experiments

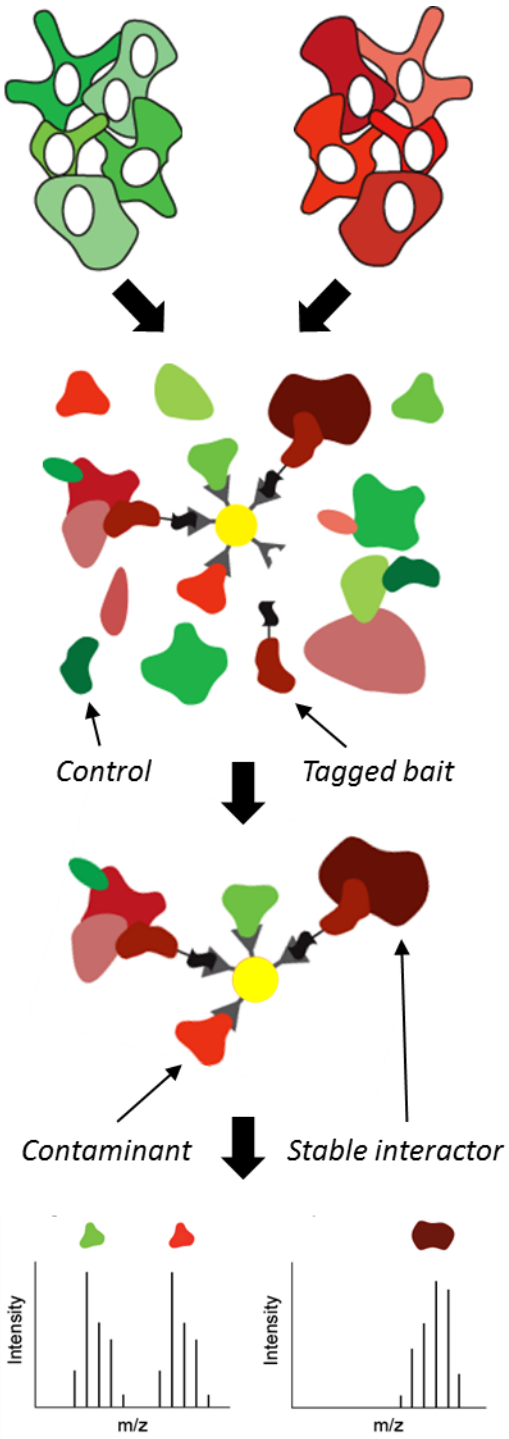
A comprehensive examination of protein interaction partners requires high throughput methodologies (discussed in 1.3). Shotgun mass spectrometry is commonly used to identify proteins in complex mixtures like IP eluates. Modern instruments, like the LTQ Orbitrap Velos, collect high resolution, high mass accuracy data that allow for detection and identification of thousands of proteins across several orders of dynamic range in a single sample. The high sensitivity of mass spectrometry means that irrelevant background proteins are invariably detected thus making identification of specifically interacting partners difficult. As a result, when using an immunopurification-based strategy to discover new protein-protein interactions, controls must be incorporated to distinguish true interactors from proteins that non-specifically

associate with the capture complex (discussed in 1.3). This issue is illustrated in Figure 7 which outlines the steps of a basic IP-MS experiment. Control IPs can be conducted in a number of different ways including removing the tag from the target (Jorba et al. 2008), altering the capture complex (Eisfeld et al. 2011), or placing the tag on a control protein like green fluorescent protein (GFP) (Carpp et al. 2014). These types of control IPs are suboptimal because their conditions are all biochemically distinct from those in the experimental IP and thus the captured interactions will be different. In addition, the high sensitivity of the mass spectrometer means that even proteins with very low abundance in the control will be detected and subtracting the control identifications from the experimental can lead to a high false negative rate.

There are also post acquisition methods for identifying background contaminants including analytical tools like SAINT (discussed in 1.3) and curated databases of common IP contaminants like CRAPome (contaminant repository for affinity purification data) (Mellacheruvu et al. 2013). The CRAPome is based on the principle that proteins identified in negative controls are bait independent. This permits aggregation of negative control results according to experimental conditions (i.e. all control IPs that used magnetic beads). In this manner contaminants can be identified by the frequency and abundance with which they are represented in the repository.

I-DIRT (isotopic differentiation of interactors as random or targeted) is another way to identify stably interacting components in IP-MS experiments. I-DIRT was initially developed in yeast by Tackett et al. and has recently been applied to study host-virus interactions in dengue and influenza (Carpp et al. 2014; Tackett et al. 2005; Tawaratsumida et al. 2014). The technique relies on the incorporation of isotopically heavy versions of lysine and arginine into the full cellular proteome by SILAC (discussed in 1.3 and detailed in 3.3). This leads to the establishment of “heavy” and “light” cell cultures which are differentiated only by the lysine and arginine isotopes in their proteins. All other aspects of the cultures are the same and thus subsequent perturbations are executed on a state of equivalence. Expression of a tagged version of the target protein in either the heavy or light culture is required for an I-DIRT

experiment. Tagged proteins may be expressed from a plasmid (Carpp et al. 2014) or genomically tagged as can be done in yeast (Tackett et al. 2005) and pathogens (Tawaratsumida et al. 2014). Once the tagged target is present in one of the cultures, equal amounts of both heavy and light cells are mixed, lysed, and an IP is performed for the tag. The eluate is then processed for MS. Based on the isotopic equivalence principle, heavy and light peptides should respond congruently in the mass spectrometer and their signal intensities can be used for relative quantitation. Figure 8 is a graphical representation of an I-DIRT experiment. If the tagged bait was expressed in the heavy culture, stably interacting proteins will be identified by primarily heavy peptides while the inverse is true if the tagged protein was from the light culture. Contaminating proteins that interact non-specifically are present in the lysate in equal parts heavy and light and so should also be present in equal amounts in the eluate. Thus proteins identified by an equal number of heavy and light peptides represent background. However an important exception exists. Proteins that interact specifically with the target but rapidly cycle on and off will have heavy/light peptide (H/L) ratios proportional to the exchange rate and IP time. In order to identify specific interactors with fast on/off rates, the time period after mixing and lysis through elution should be minimized (Figure 9).



Cells are split into two cultures: one is supplemented with isotopically heavy Lys and Arg (red) while the other is grown with regular "light" versions (green)

A tagged version of the target is introduced into either culture (heavy in this example) while the control (untagged protein) is expressed in the other culture

The infected cells are combined in equal parts and lysed. The IP is performed in the mixed lysate

Proteins stably interacting with the target bait protein *in vivo* and through the IP process will be identified by isotopically heavy peptides

Proteins that randomly associate with the capture complex will be detected by equal numbers of heavy and light peptides

Figure 8. Graphical representation of an I-DIRT experiment

We used an engineered virus to introduce a tagged version of NEP into isotopically heavy cells and infected the normal cell culture with wild-type virus.

We used the I-DIRT technique to discover new host-influenza interactions for several reasons. First, it features a high degree of internal control because once the cells are mixed in the initial step; all subsequent manipulations are performed on a single sample. This minimizes the effect of technical variability on the quantitative read out. Second, because background interactors are identified by H/L ratios not absolute abundance, less stringent IP conditions can be used which may enable the detection of more weakly interacting partners. Lastly the quantitative readout (H/L ratio) allows the identified proteins to be ranked by likelihood of a stable interaction with NEP and provides guidance for prioritizing follow-up studies.

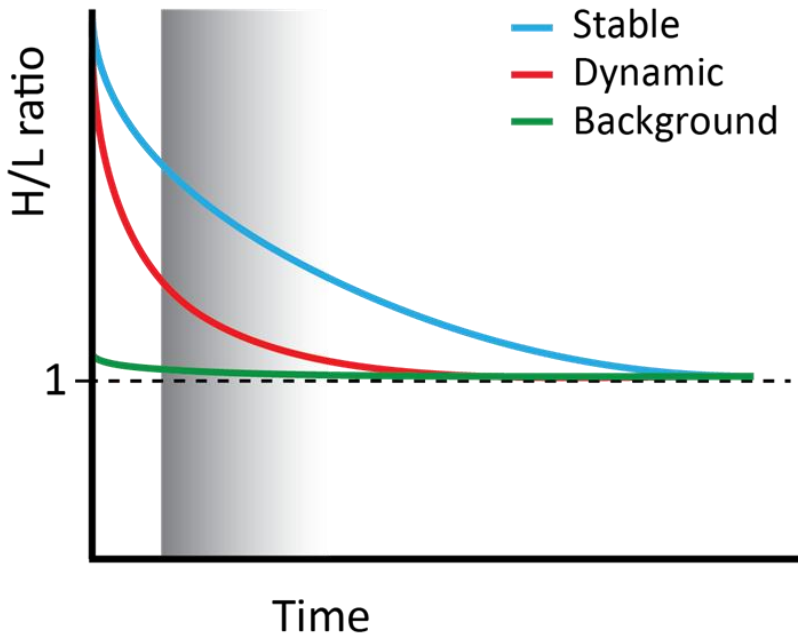


Figure 9. Theoretical representation of H/L ratios for different classes of interactors
The I-DIRT technique uses the ratio of heavy to light peptides to distinguish true interactors from contaminants. Proteins that stably associate with isotopically heavy bait *in vivo* and through the affinity capture procedure should be identified predominantly by heavy peptides and garner a high H/L ratio (blue). Proteins that interact specifically *in vivo* but rapidly cycle on and off the bait will have an intermediate H/L ratio (red). Contaminants that randomly associate with the capture complex should have H/L ratios close to 1 (green). The H/L ratios for all species will approach 1 as time increases. The shaded area represents the target window for execution of the IP procedure. Faster IPs will increase the difference between ratios from contaminants vs. stable and dynamic interactors.

2.2.2 Strategies and considerations for tagging NEP

Epitope tagging of proteins was developed in the 1980's in parallel with molecular biology techniques for specifically manipulating DNA sequences. Early applications of epitope tags include work by Hopp et al. who developed the FLAG tag in 1988 which is still widely used (Hopp et al. 1988). Also in 1988, Geli et al. used a colicin A epitope to study the localization of the immunity protein to colicin A and the researchers noted that "[w]ith the availability of monoclonal antibodies directed against known epitopes, this technique of using an epitope as a tag to label a protein should prove useful for many purposes" (Geli, Baty, and Lazdunski 1988). This has indeed proven to be the case and a diverse array of commercial reagents are now available to streamline workflows involving tags.

In the most general sense, a tag serves as a known component from which to operate. Once in place within the experimental system, the tag can be specifically targeted by reagents that provide information about localization, confer new biochemical properties, or that immobilize the tagged target. Choice of tag will necessarily relate to the desired operation and we focused on the physical capture of target influenza proteins and interacting host factors from lysates of infected cells. It has been reported that antagonistic interactions between virus-host proteins tend to be more transient in nature (Franzosa and Xia 2011). Capture of weaker, more fleeting interactions will be maximized with the use of tags that interact with their capture agents with high affinity. Binding (k_{on}) of monoclonal antibodies to their target epitopes occurs at approximately the rate expected for diffusion-limited bimolecular interactions (10^6 $M^{-1}s^{-1}$) (Raman et al. 1992). As a result high affinity interactions between epitope and antibody are characterized by low rates of dissociation (k_{off}). For the rapid purification of targets with intermediate or low abundance, the equilibrium dissociation rate constant (K_D) between the tag and capture agent should be at or below 10nM (LaCava et al. 2015). Antibody-antigen interactions are generally characterized by equilibrium dissociation rate constants in the nanomolar range (10^{-7} to $10^{-9}M^{-1}$) (Karlsson, Michaelsson, and Mattsson 1991).

Size is an additional consideration since influenza has limited space for additional genomic sequence therefore smaller tags may minimize effects on viral fitness. It is also

important to consider the functional consequences of tagging since this represents a physical alteration to the target protein's structure. Ideally a tag will not interfere with any interactions nor adversely affect protein folding but this is not always the case. When tagged proteins are expressed from a plasmid, additional functionality testing must be done to ascertain whether or not native properties are retained by the chimeric version. In contrast, when an essential protein is tagged genomically and expression of the native form is abolished, this information is self-evident. Two different tags were successfully integrated into the influenza genome to produce N-terminally tagged NEP, the streptavidin binding peptide (SBP) tag and the 3xFLAG tag.

The SBP tag is a 38 amino acid optimized version of a peptide that was selected for streptavidin binding using mRNA display. SBP interacts with streptavidin with high affinity (K_D 2.5nM) and can be eluted in mild conditions using biotin (Keefe et al. 2001). This combination allows for high yield, high purity isolations with few steps. However, we experienced high levels of background in affinity purifications using streptavidin so also tried the 3xFLAG tag. 3xFLAG is a triple, tandem repetition of the original 8 amino acid sequence, DYKDDDDK, developed specifically for immunoaffinity detection and purification at Immunex. This epitope was rationally designed for small size, hydrophilicity and antigenicity and has been widely adopted due to its favorable properties (Einhauer and Jungbauer 2001b; Hopp et al. 1988; LaCava et al. 2015). It is worth noting that the FLAG epitope should be used with caution in non-mammalian systems as the Tyr residue is susceptible to sulfation which prevents interaction with antibodies (Schmidt et al. 2012). The 3x version of the FLAG tag was reported in 2000 where the first two repetitions are variations of the original sequence to prevent inclusion of additional enterokinase recognition sights. This construct features improved detection sensitivity over the original (Hernan, Heuermann, and Brizzard 2000). Several antibodies that recognize the FLAG epitope are commercially available, M1, M2, and M5. M1 recognizes N-terminal FLAG epitopes supposedly in a calcium-dependent manner (Hopp et al. 1988; Prickett, Amberg, and Hopp 1989) but this was subsequently called into question (Einhauer and Jungbauer 2001a). M5 preferentially binds N-terminal FLAG epitopes preceded by Met (Chubet and Brizzard 1996).

The most popular antibody, M2, recognizes FLAG independent of location (Brizzard, Chubet, and Vizard 1994; Einhauer and Jungbauer 2001b) and is the one used in this work.

Once suitable tags were identified, a strategy for engineering the tags into the viral genome was devised. Methodologies for heterologous protein expression in influenza have been available since 1989 when Luytjes et al. observed packaging of the chloramphenicol acetyltransferase (CAT) gene into infectious virions (Luytjes et al. 1989). This early system relied on co-infection with a wild-type "helper" virus so the output was a mixed population of recombinant and wild-type particles. The full recovery of influenza from plasmids proved technically challenging and was achieved in 1999 by Neumann et al. and also by Fodor et al. by transfecting 12 different constructs into cultured cells—8 to produce the negative-sense vRNA segments and 4 to produce the polymerase complex proteins and NP (Fodor et al. 1999; Neumann et al. 1999). The key feature of these new systems was the incorporation of RNA polymerase I (Pol I) signals for the production of uncapped, unadenylated vRNAs. A refined version used in this work was published one year later that includes RNA polymerase II (Pol II) sites for the production of vmRNAs in addition to vRNAs thus enabling transfection of just 8 plasmids instead of 12 (Figure 10) (Hoffmann et al. 2000). With this plasmid-based system in hand, researchers can utilize available tools for DNA engineering to readily manipulate the influenza genome. Two primary considerations were accounted for in our tagging strategies. First NEP is encoded on a segment that is alternatively spliced during transcription and we wanted to avoid tagging NS1. Secondly, there are conserved packaging signals that extend into the coding regions that must be preserved to ensure the segments are incorporated into nascent virions (Fujii et al. 2005; Gog et al. 2007; Muramoto et al. 2006).

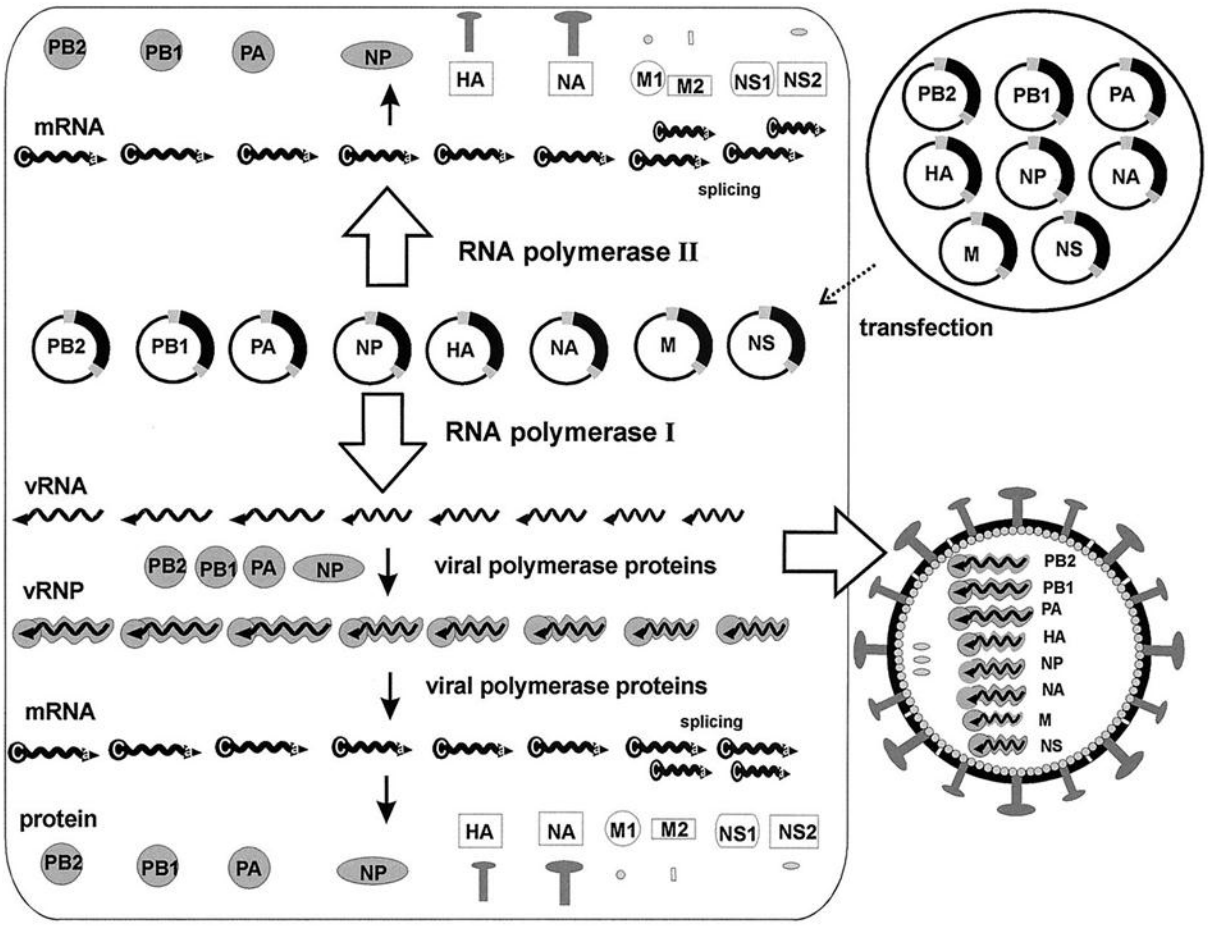


Figure 10. The eight-plasmid Pol I–Pol II system for the generation of influenza A virus

Expression plasmids containing the eight viral cDNAs inserted between the human RNA polymerase I and II (Pol I and Pol II) promoters are transfected into eukaryotic cells. The Pol I transcription product is antisense and serves as template for the viral polymerase complex (vRNA). The Pol II product is an mRNA which is translated into influenza protein. Once the polymerase complex proteins have been translated, the infectious cycle is initiated. Figure adapted from Hoffmann et al. 2000.

One key feature of our tagging strategy was the use of the 2A self-cleaving peptide sequence from foot-and-mouth disease virus (Ryan, King, and Thomas 1991). This 17 amino acid sequence promotes ribosomal skipping during translation leading to the production of two independent proteins from a single transcript (Donnelly et al. 2001). There are alternative versions of the 2A peptide derived from other viruses and processing of these cleavage signals

is context dependent (influenced by both cell-type (Kim et al. 2011) and adjacent sequence (Ryan, King, and Thomas 1991)) and inefficient skipping leads to production of a fusion protein. We tested 2 other self-cleaving sequences, T2A, and E2A and also a triple tandem repeat of 2A (32A) and were able to significantly reduce the amount of fusion with both 32A and E2A (a third variant, P2A failed during the cloning process). The self-cleaving sequences are aligned in Figure 11. The 2A sequence has been employed as a tool in a variety of contexts including gene therapy (Fang et al. 2005; Szymczak et al. 2004) and to study influenza. The 2A sequence was first used for expression of a foreign sequence in influenza in 1994. The chloramphenicol transferase (CAT) reporter gene was stably incorporated into the influenza genome by separating the CAT and NA coding sequences with 2A (Percy et al. 1994). Tawaratsumida et al. (2014) inserted 2A and a tag onto the NS segment and used the recombinant virus in an I-DIRT experiment to find new interactors of NS1 and NEP. However this virus produces 3xFLAG-StrepOne tags on both NS1 and NEP making it difficult to distinguish unique interactors for either protein. In very recent work, the full coding sequences of NS1 and NEP were separated on NS segment with the P2A sequence. Using this strategy fluorescent reporter viruses were produced that express GFP-tagged NEP during infection (Reuther et al. 2015).

We successfully recovered infectious virus expressing NEP N-terminally tagged with either SBP or 3xFLAG by relocating the coding sequence to segment 6 which codes for NA. We also mutated the splice acceptor site on the NS segment to prevent alternative splicing. A schematic diagram of the NS and NA segments of wild type and our first generation virus NA are presented in Figure 12.

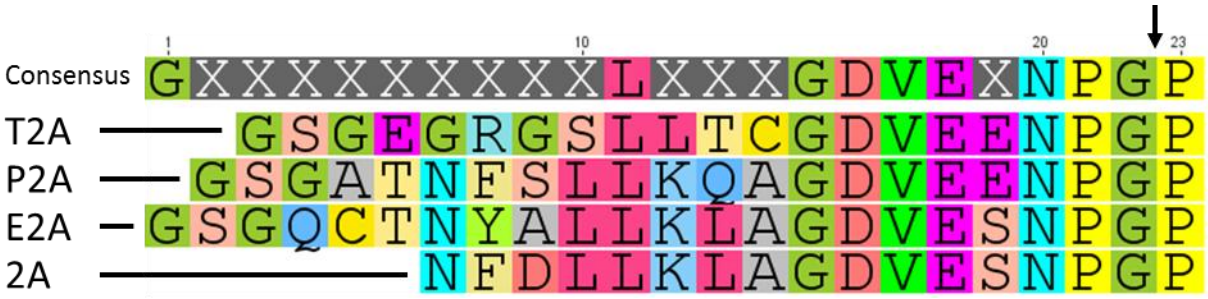


Figure 11. Alignment of self-cleaving 2A variant sequences

The sequences of the 2A variants tested were aligned in Geneious v7.1.9 using the Geneious method with default parameters. Polypeptide chain break occurs between the C-terminal Gly and Pro residues (arrow).

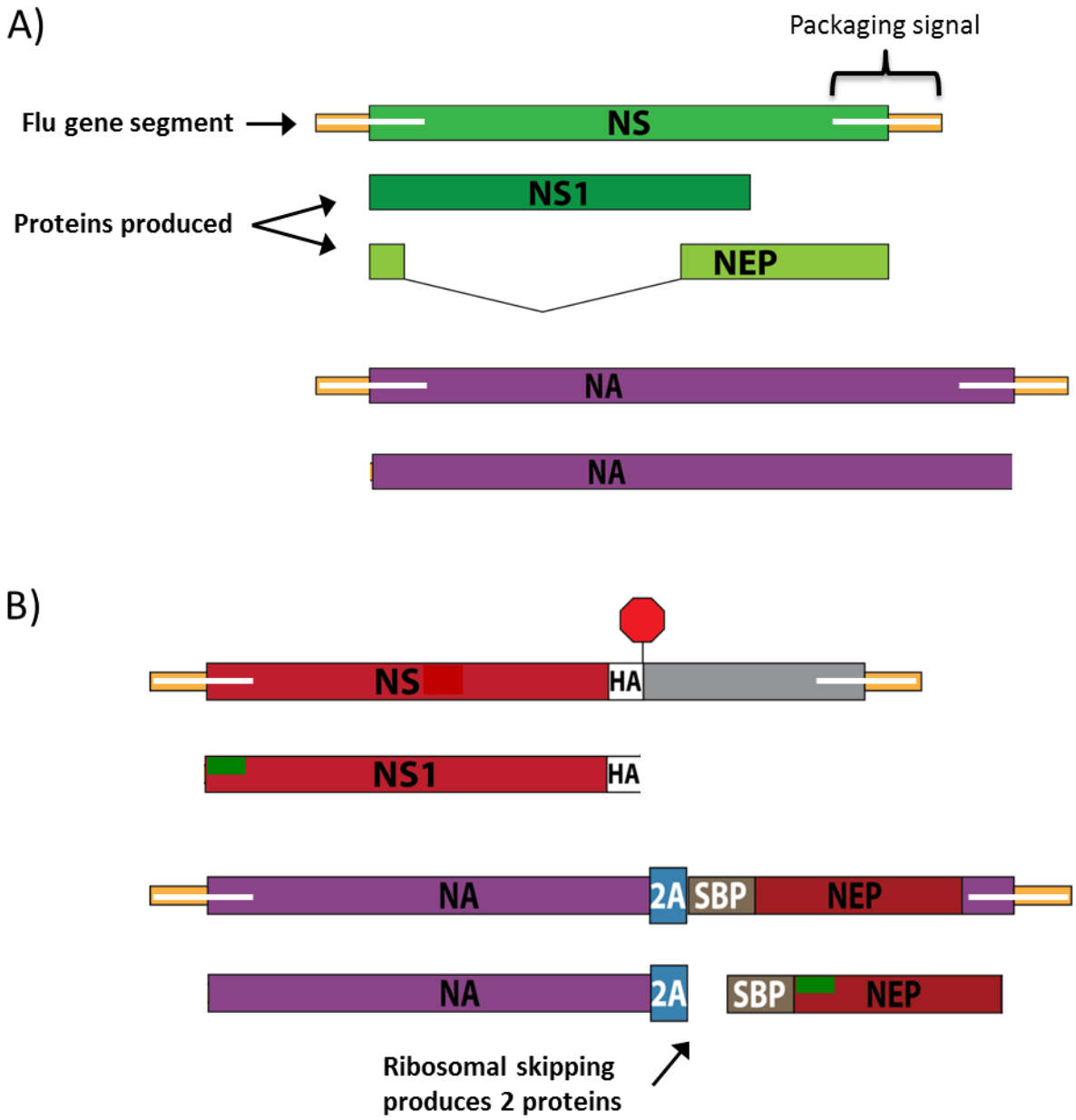


Figure 12. Schematic diagram showing the native and engineered NS and NA segments from PR8 and N2SN viruses and the proteins produced

Wild type NS and NA segments are presented in (A). The NS segment is alternatively spliced to produce NS1 and NEP while the NA segment produces NA protein. (B) The first generation recombinant virus (NA-2A-SBP-NEP, (N2SN)) features a mutated splice acceptor site to prevent production of NEP mRNAs and a C-terminal HA tag on NS1. Sequences encoding the streptavidin-binding peptide (SBP) and NEP were added to the NA segment. The NA and SBP-NEP sequences were separated by sequence coding for the 2A peptide which generates a peptide chain break during translation. The green boxes on NS1 and NEP denote shared N-terminal sequence.

2.3 Methods

2.3.1 Plasmid Constructs

NEP was cloned into the pGEX-4T1 vector (EcoRI/NotI) for *E. coli* expression. pHW2000 vectors were used to generate infectious PR8 virus (Hoffmann et al. 2000). All polymerase chain reactions (PCRs) were performed with Phusion® polymerase using standard procedures. The influenza segments in the pHW2000 plasmids are all located between SalI and BstEII cloning sites on the 5' and 3' ends respectively. Sequences for the engineered NS and NA segments between the SalI and BstEII cloning sites are presented in Table 2. The accuracies of all plasmid constructs were verified by sequencing. For the construction of the 2A sequence variants, the full NA-T2A-SBP-NEP sequence between the BstEII and SalI cloning sites was built by BlueHeron Inc. This construct features the addition of a StuI recognition site in NA by mutating guanine 1215 (from the initiator adenosine) to thymine (AGGCCG -> AGGCCT) and also the addition of an XhoI recognition site in NEP by mutating thymine 61 and guanine 63 (from the NEP initiator adenosine) both to cytosine (TTGGAG -> CTCGAG). These are silent mutations. The self-cleaving peptide and tag sequences can be readily exchanged by restriction cloning between these two sites.

Table 2. Primer and gene segment sequences

Oligo/sequence ID	Purpose	Nucleic acid sequence
pHW2000_SalI_fwd	Universal pHW2000 forward	caattaGTCGACCTCCGAAGTTGGGG
pHW2000_BsteII_rvs	Universal pHW2000 reverse	caattaGGTGACCGCCGGGAGGGCG
2A	Self-cleaving sequence	AACTTTGACCTTCTCAAGTTGGCCGGCGACGTCGAGTCCAACCCAGGGCCC
E2A	Self-cleaving sequence	GGAAGCGGACAGTGTACTAATTATGCTCTCTTGAAATTGGCTGGAGATGTTGAGAGCAACCCTGGACCT
T2A	Self-cleaving sequence	GGAAGCGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGTCGAGGAGAATCCTGGACCT
P2A	Self-cleaving sequence	GGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTGGACCT
NS segment	N2SN virus	GTCGACCTCCGAAGTTGGGGGGGAGCAAAAGCAGGGTGACAAAAACATAATGGATCCAAACACTGTGTCAAGCTTTCAGGTAGATTGCTTTCTTTGGCATGTCCGCAAACGAGTTGCAGACCAAGA ACTAGGCGATGCCCCATTCCTTGATCGGCTTCGCCGAGATCAGAAATCCCTAAGAGGAAGGGGCAGTACTCTCGGTCTGGACATCAAGACAGCCACACGTGCTGGAAAGCAGATAGTGGAGCGG ATTCTGAAAGAAGAATCCGATGAGGCACTTAAAATGACCATGGCCTCTGTACCTGCGTCGCGTTACCTAACTGACATGACTCTTGAGGAAATGTCAAGGGACTGGTCCATGCTCATACCCAAGCAGA AAGTGGCAGGCCCTCTTTGTATCAGAATGGACCAGGCGATCATGGATAAGAACATCATACTGAAAGCGAACTTCAGTGTGATTTTTGACCGGCTGGAGACTCTAATATTGCTAAGGGCTTTCACCGA AGAGGGAGCAATTGTTGGCGAAATTTCACCATTGCCTTCTCTTCCCGGGCATACTGCTGAGGATGTCAAAAATGCAGTTGGAGTCCTCATCGGAGGACTTGAATGGAATGATAACACAGTTCGAGTC TCTGAAACTCTACAGAGATTCGCTTGGAGAAGCAGTAATGAGAATGGGAGACCTCCACTCACTCCAAAACAGAAACGAGAAATGGCGGGAACAATTAGGTCAGAAGTTGGAGGAGGATACCCATA CGATGTTCCAGATTACGCTTGAAGAAATAAGATGGTTGATTGAAGAAGTGAGACACAAACTGAAGATAACAGAGAATAGTTTTGAGCAAATAACATTTATGCAAGCCTTACATCTATTGCTTGAAGTG GAGCAAGAGATAAGAACTTTCTCGTTTCAGCTTATTTAGTACTAAAAAACACCCTTGTTTCTACTAATAACCCGGCGGCCCAAAATGCCGACTCGGAGCGAAAGATATACCTCCCCCGGGGCCGGGA GGTCGCGTCACCGACCACGCCGCCGGCCCAGGCGACGCGCGACACGGACACCTGTCCCCAAAAACGCCACCATCGCAGCCACACACGGAGCGCCCGGGGCCCTCTGGTCAACCCCAGGACACA CGCGGGAGCAGCGCCGGGCCGGGGACGCCCTCCCGGCGGTCACC
NA segment	N2SN virus	GTCGACCTCCGAAGTTGGGGGGGAGCAAAAGCAGGAGTTTAAAATGAATCCAAATCAGAAAATAATAACCATTGGATCAATCTGTCTGGTAGTCGGACTAATTAGCCTAATATTGCAAATAGGGAAT ATAATCTCAATATGGATTAGCCATTCAATTCAAACTGGAAGTCAAAACCATACTGGAATATGCAACCAAAACATCATTACCTATAAAAATAGCACCTGGGTAAAGGACACAACTTCAGTGATATTAACC GGCAATTCATCTCTTTGTCCCATCCGTGGGTGGGCTATATACAGCAAAGACAATAGCATAAGAATTGGTTCCAAAGGAGACGTTTTTGTCATAAGAGAGCCCTTTATTTCATGTTCTCACTTGGAATGC AGGACCTTTTTTCTGACCCAAGGTGCCTTACTGAATGACAAGCATTCAAGTGGGACTGTTAAGGACAGAAGCCCTTATAGGGCCTTAATGAGCTGCCCTGTCGGTGAAGCTCCGTCCCCGTACAATT CAAGATTTGAATCGGTTGCTTGGTCAGCAAGTGCATGTCATGATGGCATGGGCTGGCTAACAATCGGAATTTCAGGTCCAGATAATGGAGCAGTGGCTGTATTAAAATACAACGGCATAATAACTGA AACCATAAAAAGTTGGAGGAAGAAAATATTGAGGACACAAGAGTCTGAATGTGCCTGTGTAAATGGTTCATGTTTACTATAATGACTGATGGCCCGAGTGATGGGCTGGCCTCGTACAAAATTTTC AAGATCGAAAAGGGGAAGGTTACTAAATCAATAGAGTTGAATGCACCTAATTCTCACTATGAGGAATGTTCCTGTTACCCTGATACCGGCAAAGTGATGTGTGTGCAGAGACAATTGGCATGGTT CGAACCGGCCATGGGTGTCTTTCGATCAAAACCTGGATTATCAAATAGGATACATCTGCAGTGGGGTTTTCGGTGACAACCCGCGTCCCGAAGATGGAACAGGCAGCTGTGGTCCAGTGTATGTTG ATGGAGCAAACGGAGTAAAGGGATTTCATATAGGTATGGTAATGGTGTTTGGATAGGAAGGACCAAAGGTCACAGTTCCAGACATGGGTTTGAGATGATTTGGGATCCTAATGGATGGACAGAG ACTGATAGTAAGTTCTCTGTGAGGCAAGATGTTGTGGCAATGACTGATTGGTCAGGGTATAGCGGAAGTTTCGTTCAACATCCTGAGCTGACAGGGCTAGACTGTATGAGGCCTTGCTTCTGGGTTG AATTAATCAGGGGACGACCTAAAGAAAAAACAATCTGGACTAGTGCGAGCAGCATTTCTTTTTGTGGCGTGAATAGTGATACTGTAGATTGGTCTTGGCCAGACGGTGCTGAGTTGCCATTCAGCAT TGACAAGAACTTTGACCTTCTAAGTTGGCCGGCGACGTCGAGTCCAACCCAGGGCCCATGGATGAAAAGACAACAGGATGGAGAGGTGGACATGTCGTAGAAGGACTTGCAGGTGAACTAGAA CAATTGAGAGCAAGATTGGAACATCACCCTCAAGGTCAAAGAGAACCTATGGACCCCAACACCGTGAGCTCATTCCAGGACATACTGCTGAGGATGTCAAAAATGCAGCTCGAGTCCTCATCGGAGG GACTTGAATGGAATGATAACACAGTTCGAGTCTCTGAAACTCTACAGAGATTCGCTTGGAGAAGCAGTAATGAGAATGGGAGACCTCCACTCACTCCAAAACAGAAACGAGAAATGGCGGGAACAA TTAGGTCAGAAGTTTGAAGAAATAAGATGGTTGATTGAAGAAGTGAGACACAAACTGAAGATAACAGAGAATAGTTTTGAGCAAATAACATTTATGCAAGCCTTACATCTATTGCTTGAAGTGGAGCA AGAGATAAGAACTTTCTCGTTTCAGCTTATTTAGTAAGAATTAATCAGGGGACGACCTAAAGAAAAACAATCTGGACTAGTGCGAGCAGCATTTCTTTTTGTGGCGTGAATAGTGATACTGTAGATT GGTCTTGGCCAGACGGTGCTGAGTTGCCATTCAGCATTGACAAGTAGTCTGTTCAAAAAACTCCTTGTTTCTACTAATAACCCGGCGGCCCAAAATGCCGACTCGGAGCGAAAGATATACCTCCCCC GGGGCCGGGAGGTCGCGTCACCGACCACGCCGCCGGCCCAGGCGACGCGCGACACGGACACCTGTCCCCAAAAACGCCACCATCGCAGCCACACACGGAGCGCCCGGGGCCCTCTGGTCAAC CCCAGGACACACGCGGGAGCAGCGCCGGGCCGGGGACGCCCTCCCGGCGGTCACC
NS segment	32A, 32A-3xFLAG virus	GTCGACCTCCGAAGTTGGGGGGGAGCAAAAGCAGGGTGACAAAAACATAATGGATCCAAACACTGTGTCAAGCTTTCAGGTAGATTGCTTTCTTTGGCATGTCCGCAAACGAGTTGCAGACCAAGA ACTAGGCGATGCCCCATTCCTTGATCGGCTTCGCCGAGATCAGAAATCCCTAAGAGGAAGGGGCAGTACTCTCGGTCTGGACATCAAGACAGCCACACGTGCTGGAAAGCAGATAGTGGAGCGG ATTCTGAAAGAAGAATCCGATGAGGCACTTAAAATGACCATGGCCTCTGTACCTGCGTCGCGTTACCTAACTGACATGACTCTTGAGGAAATGTCAAGGGACTGGTCCATGCTCATACCCAAGCAGA AAGTGGCAGGCCCTCTTTGTATCAGAATGGACCAGGCGATCATGGATAAGAACATCATACTGAAAGCGAACTTCAGTGTGATTTTTGACCGGCTGGAGACTCTAATATTGCTAAGGGCTTTCACCGA AGAGGGAGCAATTGTTGGCGAAATTTCACCATTGCCTTCTCTTCCCGGGCATACTGCTGAGGATGTCAAAAATGCAGTTGGAGTCCTCATCGGAGGACTTGAATGGAATGATAACACAGTTCGAGTC TCTGAAACTCTACAGAGATTCGCTTGGAGAAGCAGTAATGAGAATGGGAGACCTCCACTCACTCCAAAACAGAAACGAGAAATGGCGGGAACAATTAGGTCAGAAGTTTGAAGAAATAAGATGGTT GATTGAAGAAGTGAGACACAAACTGAAGATAACAGAGAATAGTTTTGAGCAAATAACATTTATGCAAGCCTTACATCTATTGCTTGAAGTGGAGCAAGAGATAAGAACTTTCTCGTTTCAGCTTATTTA GTACTAAAAAACACCCTTGTTTCTACTAATAACCCGGCGGCCCAAAATGCCGACTCGGAGCGAAAGATATACCTCCCCCGGGGCCGGGAGGTCGCGTCACCGACCACGCCGCCGGCCCAGGCGA CGCGCGACACGGACACCTGTCCCCAAAAACGCCACCATCGCAGCCACACACGGAGCGCCCGGGGCCCTCTGGTCAACCCCAGGACACACGCGGGAGCAGCGCCGGGCCGGGGACGCCCTCCC GGCGGTCACC
NA segment	32A virus	GTCGACCTCCGAAGTTGGGGGGGAGCAAAAGCAGGAGTTTAAAATGAATCCAAATCAGAAAATAATAACCATTGGATCAATCTGTCTGGTAGTCGGACTAATTAGCCTAATATTGCAAATAGGGAAT ATAATCTCAATATGGATTAGCCATTCAATTCAAACTGGAAGTCAAAACCATACTGGAATATGCAACCAAAACATCATTACCTATAAAAATAGCACCTGGGTAAAGGACACAACTTCAGTGATATTAACC GGCAATTCATCTCTTTGTCCCATCCGTGGGTGGGCTATATACAGCAAAGACAATAGCATAAGAATTGGTTCCAAAGGAGACGTTTTTGTCATAAGAGAGCCCTTTATTTCATGTTCTCACTTGGAATGC AGGACCTTTTTTCTGACCCAAGGTGCCTTACTGAATGACAAGCATTCAAGTGGGACTGTTAAGGACAGAAGCCCTTATAGGGCCTTAATGAGCTGCCCTGTCGGTGAAGCTCCGTCCCCGTACAATT CAAGATTTGAATCGGTTGCTTGGTCAGCAAGTGCATGTCATGATGGCATGGGCTGGCTAACAATCGGAATTTCAGGTCCAGATAATGGAGCAGTGGCTGTATTAAAATACAACGGCATAATAACTGA AACCATAAAAAGTTGGAGGAAGAAAATATTGAGGACACAAGAGTCTGAATGTGCCTGTGTAAATGGTTCATGTTTACTATAATGACTGATGGCCCGAGTGATGGGCTGGCCTCGTACAAAATTTTC AAGATCGAAAAGGGGAAGGTTACTAAATCAATAGAGTTGAATGCACCTAATTCTCACTATGAGGAATGTTCCTGTTACCCTGATACCGGCAAAGTGATGTGTGTGCAGAGACAATTGGCATGGTT CGAACCGGCCATGGGTGTCTTTCGATCAAAACCTGGATTATCAAATAGGATACATCTGCAGTGGGGTTTTCGGTGACAACCCGCGTCCCGAAGATGGAACAGGCAGCTGTGGTCCAGTGTATGTTG ATGGAGCAAACGGAGTAAAGGGATTTCATATAGGTATGGTAATGGTGTTTGGATAGGAAGGACCAAAGGTCACAGTTCCAGACATGGGTTTGAGATGATTTGGGATCCTAATGGATGGACAGAG ACTGATAGTAAGTTCTCTGTGAGGCAAGATGTTGTGGCAATGACTGATTGGTCAGGGTATAGCGGAAGTTTCGTTCAACATCCTGAGCTGACAGGGCTAGACTGTATGAGGCCTTGCTTCTGGGTTG AATTAATCAGGGGACGACCTAAAGAAAAAACAATCTGGACTAGTGCGAGCAGCATTTCTTTTTGTGGCGTGAATAGTGATACTGTAGATTGGTCTTGGCCAGACGGTGCTGAGTTGCCATTCAGCAT TGACAAGAACTTTGACTTACTAAAGCTGGCAGGAGACGTAGAGTCCAACCCTGGACCAGGGAATTTCGATCTCCTAAAACTAGCTGGTGATGTAGAATCAAATCCCGGTCCTGGTAACTTTGACCTT CTCAAGTTGGCCGGCGACGTCGAGTCCAACCCAGGGCCCATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGATATCGATTACAAGGATGACGATGACAAGGGCGGTATGGACCCCAA CACCGTGAGCTCATTCCAGGACATACTGCTGAGGATGTCAAAAATGCAGCTCGAGTCCTCATCGGAGGACTTGAATGGAATGATAACACAGTTCGAGTCTCTGAAACTCTACAGAGATTCGCTTGGA GAAGCAGTAATGAGAATGGGAGACCTCCACTCACTCCAAAACAGAAACGAGAAATGGCGGGAACAATTAGGTCAGAAGTTTGAAGAAATAAGATGGTTGATTGAAGAAGTGAGACACAAACTGAA GATAACAGAGAATAGTTTTGAGCAAATAACATTTATGCAAGCCTTACATCTATTGCTTGAAGTGGAGCAAGAGATAAGAACTTTCTCGTTTCAGCTTATTTAGTAAGAATTAATCAGGGGACGACCTAA AGAAAAAACAATCTGGACTAGTGCGAGCAGCATTTCTTTTTGTGGCGTGAATAGTGATACTGTAGATTGGTCTTGGCCAGACGGTGCTGAGTTGCCATTCAGCATTGACAAGTAGTCTGTTCAAAAAA CTCCTTGTTTCTACTAATAACCCGGCGGCCCAAAATGCCGACTCGGAGCGAAAGATATACCTCCCCCGGGGCCGGGAGGTCGCGTCACCGACCACGCCGCCGGCCCAGGCGACGCGCGACACGG ACACCTGTCCCCAAAAACGCCACCATCGCAGCCACACACGGAGCGCCCGGGGCCCTCTGGTCAACCCCAGGACACACGCGGGAGCAGCGCCGGGCCGGGGACGCCCTCCCGGCGGTCACC

The first column shows the sequence name. The second column indicates the context in which the sequence was used. Column three displays the nucleic acid sequence from 5'-3'.

2.3.2 Protein Expression

pGEX-4T1-NEP was transfected into chemically competent commercially available *E. coli* strain BL21(DE3)pLysS. Following ampicillin selection on agar plates, a single colony was inoculated into a 50mL starter culture (Lauria Bertani media with ampicillin (LBamp)) for overnight growth at 37°C with shaking. The following morning 1L LBamp was inoculated with starter culture (10mL) to an optical density at 600nm (OD600) of 0.1. Cultures were grown at 37°C with shaking to an OD600 of 0.5 (3 hours) then induced with 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown for 3 more hours. Bacterial cells were pelleted by centrifugation and placed at -20°C.

2.3.3 Protein Purification

Frozen bacterial cell pellets (from 500mL culture) were thawed in 30mL phosphate buffered saline pH 7.4 (PBS) and then sonicated 3x for 30 seconds. Lysates were spun at 15,000 x g for 15 minutes. Dithiothreitol (DTT) was added to 1mM and the supernatant was added to 3mL washed glutathione Sepharose-4B (GE Healthcare). The slurry was rocked for 12 hours at 4°C and then added to a gravity flow column and washed with 40 column volumes PBS. Column was eluted with three, 3mL fractions of 10mM reduced glutathione in PBS. Fractions containing GST-NEP were combined and dialyzed against PBS.

2.3.4 Protein Analysis

Protein samples were heated to 90°C for 5 minutes in lithium dodecyl sulfate (LDS) sample buffer with 50mM DTT in preparation for electrophoresis. The heated protein samples were separated by electrophoresis using precast Bis-Tris 4-12% gradient polyacrylamide gels in a 2-(N-morpholino)ethanesulfonic acid (MES) buffer system (SDS-PAGE). After electrophoresis, proteins were visualized with Coomassie blue based Imperial Protein Stain (Thermo Scientific, 24615) or transferred to PVDF membrane for western blot using standard procedures. Western blots were probed with anti-NEP polyclonal rabbit serum (at 1:2000 dilution) produced as

described in 2.4.1 or anti-FLAG M2-HRP (Sigma, A8592) at 1:10,000 dilution in PBS with 0.01% Tween-20.

2.3.5 Cell Culture

HEK293T cells were grown in complete Dulbecco's modified Eagle's medium-cDMEM (DMEM with 10% fetal calf serum, and 1% penicillin/streptomycin). MDCKII cells were grown in modified Eagle's medium (MEM) with 5% FCS, 1% penicillin/streptomycin (Gibco, 15140) and MEM vitamin solution (Gibco, 11120). Cells were maintained at 37°C and 5% CO2.

2.3.6 Viral Production

Influenza virus was produced using reverse genetics with a 5-day procedure as follows: Day 1, co-cultures of HEK293T/MDCKII cells were seeded into a 6 well dish (0.15E06/0.35E06 cells/well). Day 2, pHW2000 influenza plasmids were transfected into the co-cultures (0.6µg of each plasmid except NP (2µg), NS (0.1µg), and M (0.1µg) in 200µL of Opti-MEM® with 12µL Mirus LT1 transfection agent per well). Day 3, cDMEM was replaced with infection media (Opti-MEM® with 0.3% Fraction V Bovine Serum Albumin and 2µg/mL TPCK-trypsin). Day 4, MDCKII cells were plated in 6 well plates (0.35E6/well) in 1xMEM (Modified Eagle's Medium with 5% FCS, 1% pen/strep, and 1x MEM vitamins). Day 5, MDCKII medium was replaced with 1mL of infection media and 1mL of viral supernatant from the co-cultures. Once the MDCKII cells exhibited significant cytopathic effects from viral invasion (48-72 hours), the supernatant was transferred to MDCKII cells in a 15mL plate for amplification. After 48-72 hours, cultures were spun (4,000 RMP for 10min) to pellet cell debris and the viral supernatant was preserved for immediate use at 4°C or stored at -80°C.

2.3.7 Affinity purification for the SBP tag

HEK293T cells in 10cm dishes at 80% confluency were infected with sufficient virus to visibly observe cytopathic effects (CPE) within 20 hours (500-200µL of supernatant from the

MDCKII cell amplifications). Cells were harvested by centrifugation at 20-24 hours post infection, washed in 1mL PBS, and stored at -20°C. Frozen cell pellets were thawed on ice, incubated for 10 minutes in 800µL lysis/binding buffer (PBS + 0.5% NP-40 (IGEPAL CA-630)) and 1x protease inhibitor (Sigma, P8340), and then sonicated for 15sec. Following centrifugation (21,000 x g, 2 minutes), the soluble fraction was added to 150µL washed M280-streptavidin Dynabeads (Invitrogen, 11205D) and bound on ice for 30 minutes with three resuspentions. The beads were washed 3x in 1mL PBS or HEPES buffer as noted and transferred to a new 1.5mL Eppendorf tube with a 4th wash. The beads were eluted with 75µL LDS protein sample buffer. Alternative HEPES buffer (20mM HEPES pH 7.4, 110mM KOAc, 2mM $MgCl_2$,0.5% NP-40 and protease inhibitor) was used for IP optimization as noted in 2.4.2.3., Benzonase® nuclease was used at a concentration of 100U/mL as noted.

2.3.8 Immunopurification for the 3xFLAG-tag

HEK293T cells in 15cm dishes at approximately 80% confluency were infected with sufficient virus for visible CPE within 20 hours (300µL of supernatant from the MDCKII cell amplifications for PR8 or 2mL for 32A-3xFLAG). Cells were harvested by centrifugation at 20 hours post infection, washed in 1mL PBS, and stored at -20°C. Frozen cell pellets were thawed on ice, incubated for 10 minutes in 800µL lysis/binding buffer HB (20mM HEPES pH 7.4, 110mM KOAc, 2mM $MgCl_2$, 0.1%-5% Triton X-100 (TX-100)) and 1x protease inhibitor (Sigma, P8340), and then sonicated for 15 seconds. Following centrifugation (21,000 x g, 2 minutes) the soluble fraction was added to 300µL washed PrG-Dynabeads (Invitrogen, 10003D) covalently coupled to anti-FLAG M2 ab (Sigma, F3165). The Dynabeads were prepared as follows: 900µL washed PrG Dynabeads were rotated with 300µg M2 antibody in citrate-phosphate buffer (25mM citric acid, 50mM Na2PO4, pH 5) for 1 hour. Bead-antibody complexes were washed in 1mL 0.2M triethanolamine (TEA) pH 8.2 and then cross-linked with 20mM dimethyl pimelimidate (DMP) in TEA for 15 minutes. The reaction was stopped by washing in 50mM Tris pH 7.5 for 15 minutes. Cleared cell lysates were bound for 20 minutes at RT and then beads were washed 2x in binding buffer and transferred to a new 1.5mL Eppendorf tube with a 3rd wash. The beads were eluted

with either various concentrations of $MgCl_2$, with 200mM glycine pH 3, 1% Rapigest (Waters, 186002122), or LDS sample buffer as noted.

2.3.9 Mass spectrometry

2.3.9.1 Sample Preparation

Bands were excised and cut into 1-2mm cubes from polyacrylamide gels following staining with Imperial Protein stain. Cubes were covered with 500µL de-stain solution (25mM ammonium bicarbonate (AmBic), 50% acetonitrile (ACN)) and shaken at 37°C for 30 minutes. De-stain was removed and another 500µL was added and incubated as before. The second de-stain was removed and proteins were reduced with 200µL 10mM DTT in 100mM AmBic for 30 minutes at 60°C. After cooling to RT, cubes were covered with 200µL 2-iodoacetamide (55mM in 100mM AmBic) for 1 hour in the dark at RT. Cubes were washed 2x with de-stain as before except with 15 minutes incubations. Gel cubes were desiccated by adding 200µL ACN for 15 minutes at RT. After removal of ACN, cubes were allowed to air dry 10 minutes. Gel pieces were then re-hydrated with trypsin (Promega, V5280) buffer (12.5ng/µL trypsin in 100mM Ambic) and incubated overnight at 37°C. Gel pieces were pelleted by centrifugation and the supernatant containing digested peptides was saved. Gel cubes were rinsed with 200µL 1% trifluoroacetic acid to extract residual peptides. The rinse and digest supernatants were combined and dried in a speed vacuum. Peptides were resuspended in 15µL of 1% formic acid and stored at -20°C.

2.3.9.2 Data acquisition and analysis

Peptide samples were run on a Thermo Linear Trap Quadrupole (LTQ) Velos Orbitrap. Pre-fractionation was performed on an Eksigent ultra-high pressure nanopump. 4µL of each sample was passed first through a trap column (New Objective, IF360-75-50-N-5) packed with 2cm of C18 silica beads (Dr.Maisch ReproSil-Pur C18-AQ, 3µm) and then an analytical column (New Objective, PF5010-150H002-3P) with a 135 minute continuous dual solvent gradient

(0.1% TFA-A/acetonitrile-B): 0-5 minutes 3% A, 95 minutes 75% A, 100 minutes 65% A, 105-110 minutes 20% A, 115 minutes 65% A, 117-135 minutes 97% A. Flow rate was 300nL/minute. Standard shotgun MS-MS settings were used: profile mode full mass spectrometry scans (MS1), m/z 400-2000, at 60,000 resolution, dynamic exclusion enabled. The 10 most intense ions from each scan were selected for CID fragmentation (MS2).

Mass spectrometry .RAW files were searched and analyzed using the MaxQuant proteomics computational environment with the Andromeda search engine with default parameters according to the standard workflow (Cox and Mann 2008; Cox et al. 2009). The UniProtKB/Swiss-Prot database (release 2013_12) with influenza A/Puerto Rico/8/1934(H1N1), AcGFP (gb|AAN41637.1), 3xFLAG peptide, and 2A peptide sequences, was used for peptide identifications.

2.4 Results

2.4.1 Expression and purification of recombinant GST-NEP

In order to effectively detect the NEP protein in infected cell lysates, a polyclonal rabbit antiserum was raised against a glutathione-S-transferase (GST) fusion expressed and purified from *E. coli* (GST-NEP). It is important to note that NEP and NS1 share ten N-terminal amino acids, so the polyclonal serum was cross reactive. A Coomassie-stained gel showing the pre-bound supernatant input, the post-bind flow through, and elution fractions, is shown in Figure 13A. A secondary purification step was not used and fractions 1 and 2 were combined for dialysis. The dialyzed GST-NEP protein was used to inoculate rabbits for the production of polyclonal rabbit serum. The naïve serum from three different animals was screened by western blot against GST-NEP and no pre-existing reactivity was found for any of the animals (data not shown). Figure 13B shows a western blot of serum from rabbits #1649 and #1744. Two amounts (100ng and 20ng) of GST-NEP were either untreated or incubated with thrombin (to cleave off the GST) and then subjected to western blot. The presence of uncleaved GST-NEP, GST alone, and NEP alone were clearly evident indicating that the serum does contain antibodies against NEP. There were also faint bands in the untreated samples likely representing GST dimerization products that became much more prominent after cleavage (discussed further in 2.5).

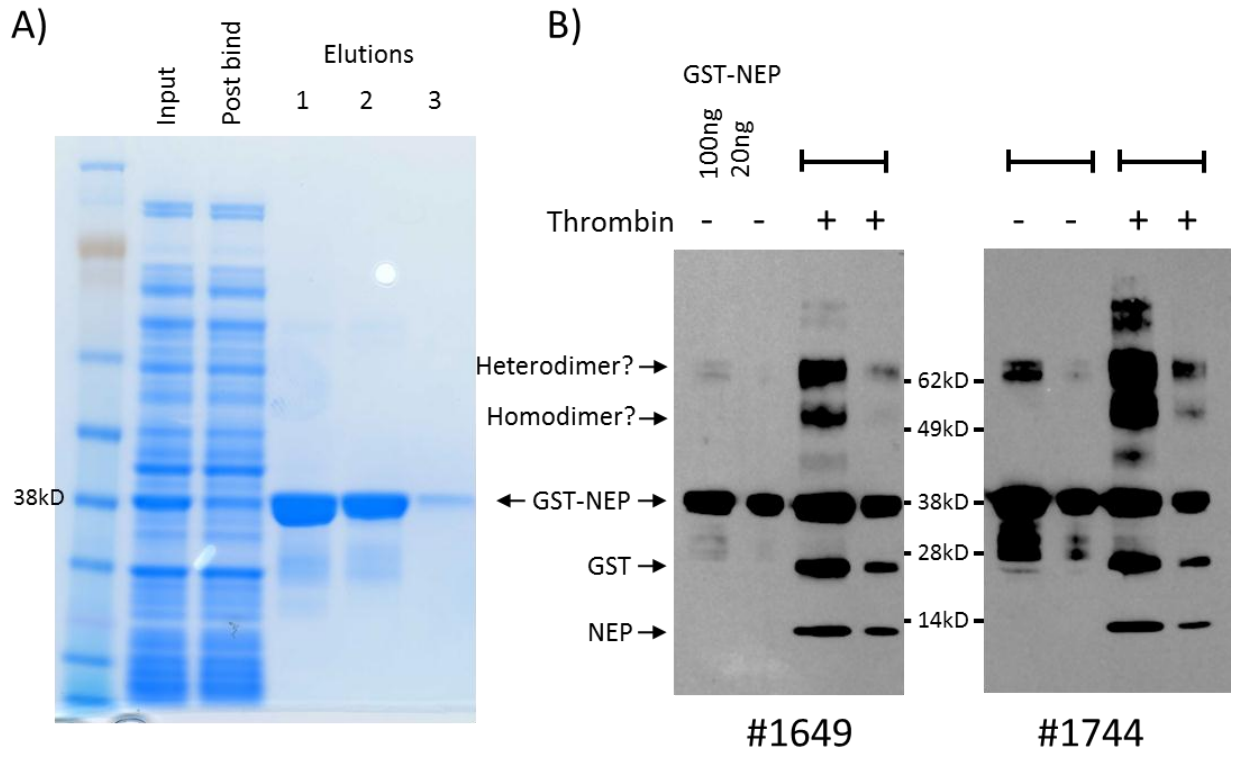


Figure 13. Production of recombinant GST-NEP and polyclonal rabbit serum
Samples from purification of GST-NEP from *E. coli* cell lysate using glutathione Sepharose-4B are shown with Coomassie blue staining in (A). Eluted fractions 1 and 2 were combined and dialyzed and then inoculated into rabbits to produce polyclonal serum. In (B) post-inoculation serum from rabbits #1649 and #1744 were analyzed for reactivity to GST-NEP treated or not with thrombin. Samples were subjected to western blot with a 1:2,000 dilution of serum. The presence of uncleaved GST-NEP, GST alone, and NEP alone was evident in the treated samples while only GST-NEP was present in the untreated. The serum from #1649 was used in all subsequent western blots. The high molecular weight bands in the treated lanes were likely from GST dimerization. The MW of GST-NEP is 41kD, GST is 26kD and NEP is 14.4 kD.

2.4.2 Genetic engineering of influenza for the expression of tagged NEP

2.4.2.1 Recombinant influenza virus NA-2A-SBP-NEP (N2SN)

The first version of the tagged virus encoded a single 2A sequence and was designated N2SN (NA-2A-SBP-NEP). The NS1 protein was also tagged with a C-terminal HA sequence to facilitate detection. Gene segments from the N2SN virus are represented schematically in Figure 12B. Replicative fitness of theN2SN virus was compared to that of wild-type virus by western blotting infected cells lysates for the influenza capsid protein M1 twenty-four hours after infection (Figure 14A). M1 signal was similar between PR8 and N2SN indicating that replication was comparable for the two strains. In Figure 14B infected cell lysates were blotted for NEP. Wild-type NEP and NS1 were detected in the PR8 sample. In the N2SN sample SBP-NEP, NS1-HA, and a high molecular weight band for the uncleaved fusion protein NA-2A-SBP-NEP were observed. When an IP for SBP was performed on infected cell lysates the fusion predominated in the eluate (Figure 14C). These results validated our tagging strategy since we were able to generate replication competent influenza that expressed an N-terminally tagged NEP. However significant production of fusion protein precluded use of this strain for AP-MS and necessitated further development to improve the processing efficiency of 2A.

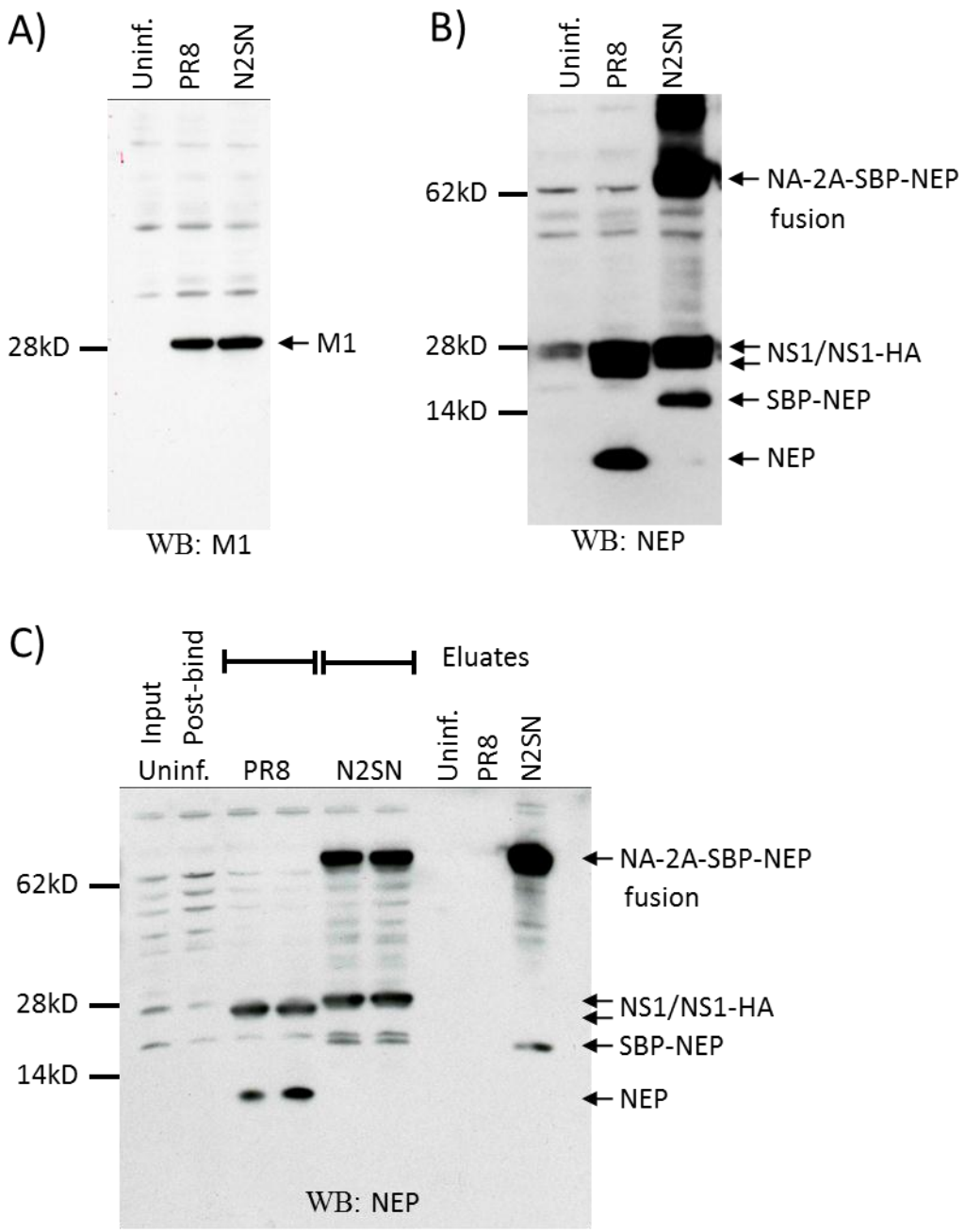


Figure 14. Western blot of HEK293T cells infected with N2SN virus
A comparison of the influenza capsid protein M1 (MW 28kD) signal in cells infected with PR8 or N2SN (A) indicates that the recombinant virus replicated similarly to wild-type. Blotting the same infected cell lysates for NEP showed wild-type NEP and NS1 (MW 26kD) in the PR8 infection and SBP-NEP, NS1-HA and NA-2A-SBP-NEP fusion (MW 71kD) in the N2SN infection (B). In (C) samples from an IP for SBP from uninfected, PR8 infected or N2SN infected cells were western blotted for NEP.

2.4.2.2 Testing of alternative self-cleaving sequences

Based on work by Kim et al. it is clear that the processing of the self-cleaving peptide sequences is cell type dependent so we screened variants of 2A from other viruses to see if they could reduce or eliminate the production of the NA-2A-SBP-NEP fusion protein during infection. Two other self-cleaving peptides were tested: E2A (from equine rhinitis A virus) and T2A (from *Thoseaasigna* virus) while a third, P2A (from porcine teschovirus-1) failed during the cloning process. The ribosomal skipping is also affected by amino acids N-terminal of the 2A sequence (Ryan, King, and Thomas 1991) so two different motifs upstream of the 2A sequence were tested. These motifs were selected based on prior work and had been used to generate tags on the NS1 protein. One was the HA tag and the other was the C-terminal NS1 sequence "RSEV". The last strategy for improving the processing of 2A was repetition of the sequence in tandem. We reasoned that if the first 2A sequence did not induce skipping then a second or third repeat may do so. The double tandem version failed to clone so only the triple tandem repeat was tested. A schematic diagram showing the engineered NS and NA segments for the 2A screening experiment is shown in Figure 15. Restriction sites were introduced to facilitate rapid exchange of 2A and tag sequences on the NA segment as detailed in 2.3.1. The yellow box shows the location of the test sequences (HA-2A, T2A, RSEV-2A, E2A, and 32A).

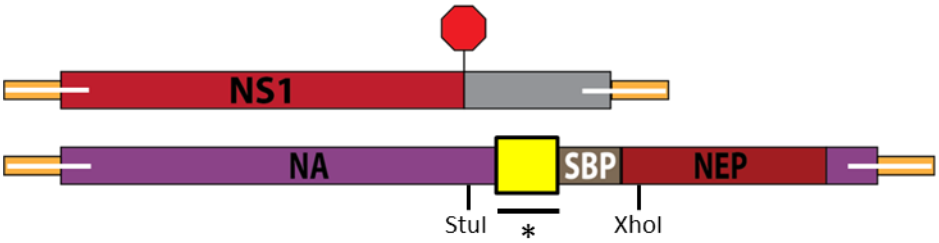


Figure 15. Recombinant NS and NA segments designed for optimization of 2A and tag sequences

Restriction sites (StuI and XhoI) were introduced into the recombinant NA segment for the rapid production of constructs with alternative 2A and tag sequences. The yellow box denoted with * indicates the placement of the alternative sequences tested in Figure 16.

These sequences were all screened in biological quadruplicate by producing 4 independent viral stocks and then infecting HEK293T cells. At 20 hours post infection, the cells were lysed and analyzed by western blotting with anti-NEP polyclonal serum (described in 2.4.1). Since the serum also recognizes NS1, the presence of the NS1 band was used as a proxy for viral replication. The key objectives for this screen were the detection of NS1 to indicate functional virus, and the absence of fusion protein. As a negative control the recombinant segment 6 encoding NA-2A-SBP-NEP was omitted during the viral production step and so bands (boxed in yellow) detected in these samples represent background (Figure 16A lanes 3-4). In the positive control PR8 samples both NS1 and NEP were produced. More NS1 is produced during infection as it is the primary product from segment 8 and the signal intensity of the NS1 band versus the NEP band was in congruence with this. The HA-2A sequence produced significant fusion product. The T2A construct did not produce robust infections as the NS1 band was absent from one replicate (Figure 16A lane 9) and severely attenuated in the other three lanes (Figure 16A lanes 10-12). Interestingly, simply exchanging the HA motif for RSEV drastically reduced fusion production (Figure 16B lanes 1-4) demonstrating the importance of sequences N-terminal to 2A. The triple tandem 2A sequence gave the least amount of fusion (Figure 16B lanes 5-8) and the E2A sequence was processed with an intermediate efficiency (Figure 16B lanes 9-12).

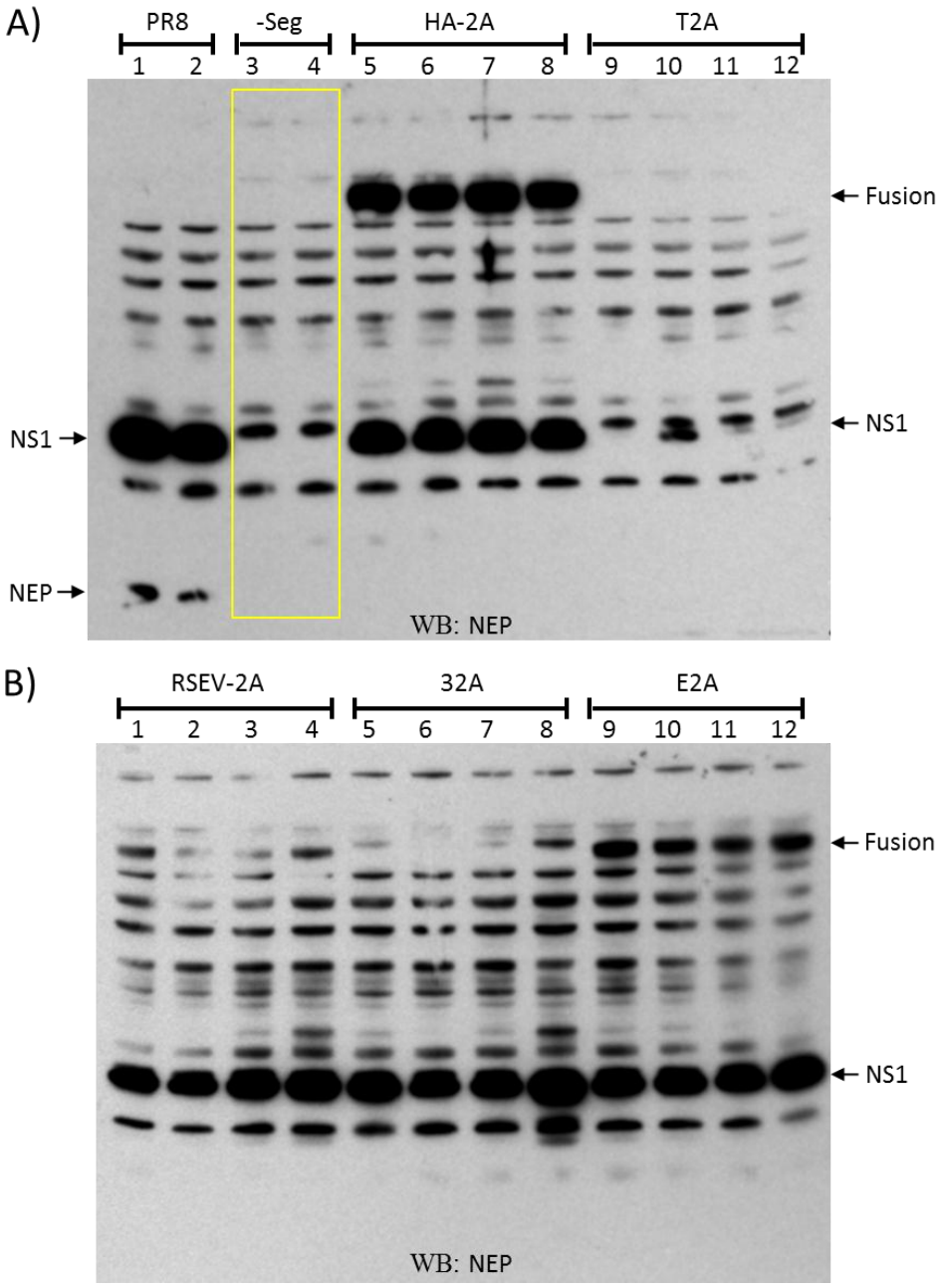


Figure 16. Western blot of self-cleaving peptide variants in infected 293T cells
A series of self-cleaving peptide sequences were compared to 2A. Viruses encoding the alternate sequences in place of 2A were produced in quadruplicate in the N2SN background and then used to infect 293T cells. Western blots of cell lysates were probed for NEP. The –Seg lanes (A3-4) had segment 6 omitted during viral production as a negative control so the bands in the yellow box represent background.

2.4.2.3 Recombinant influenza virus NA-32A-SBP-NEP (32A)

Based on the results from the screening of alternative 2A sequences in 2.4.2.2, test immunopurifications for the E2A and 32A viruses were performed. Cell lysates from HEK293T cells infected for 20 hours with the N2SN, E2A and 32A viruses were subjected to affinity purifications of the SBP tag using commercially available streptavidin magnetic beads according to the procedure detailed in 2.3.7. A western blot of the IP is shown in Figure 17. Two different exposures (A and B) are presented to illustrate different features. The NS1 band just below 28kD was used to gauge the relative degree of infection in the samples. The N2SN infection was more robust than the other two and this was likely due to a larger initial inoculum. There has been no published evidence for an interaction between NS1 and NEP and, as expected, the NS1 band was absent in the E2A and 32A eluates (Figure 17 lanes 8, 9). It is hard to conclude whether or not NS1 was present in the N2SN eluate. There was a band at 28kD in lane 7 but this could have been a breakdown product of the fusion. Signal from the fusion protein band just above 62kD was reduced in the E2A and 32A samples which is consistent with the screening results presented in Figure 16. The large smear of bands prevalent in the N2SN eluate was attributed to degradation productions of the fusion protein. The slight shift in relative size evident between the E2A and 32A fusions in Figure 17 lanes 8, 9 was due to the presence of two additional 2A sequences (+4 kD) in 32A. SBP-NEP was detected in all of the eluate samples but only in the input and post-bind from N2SN (Figure 17B lanes 1, 2) and 32A (Figure 17B lanes 5, 6). Interestingly, 32A had essentially no detectable fusion in input and post-bind fractions but SBP-NEP was observed (Figure 17B lanes 5, 6) while the opposite was true for E2A (Figure 17B lanes 3, 4). SBP-NEP from the 32A infection was detected as doublet. This is most likely because two of the three 2A sequences were processed during translation leading to a lower MW species and a higher MW species (with an N-terminal 2A sequence still attached). See discussion in 2.5 for more detail. Since the ratio of SBP-NEP to fusion was better in 32A versus E2A, we proceeded with the 32A virus.

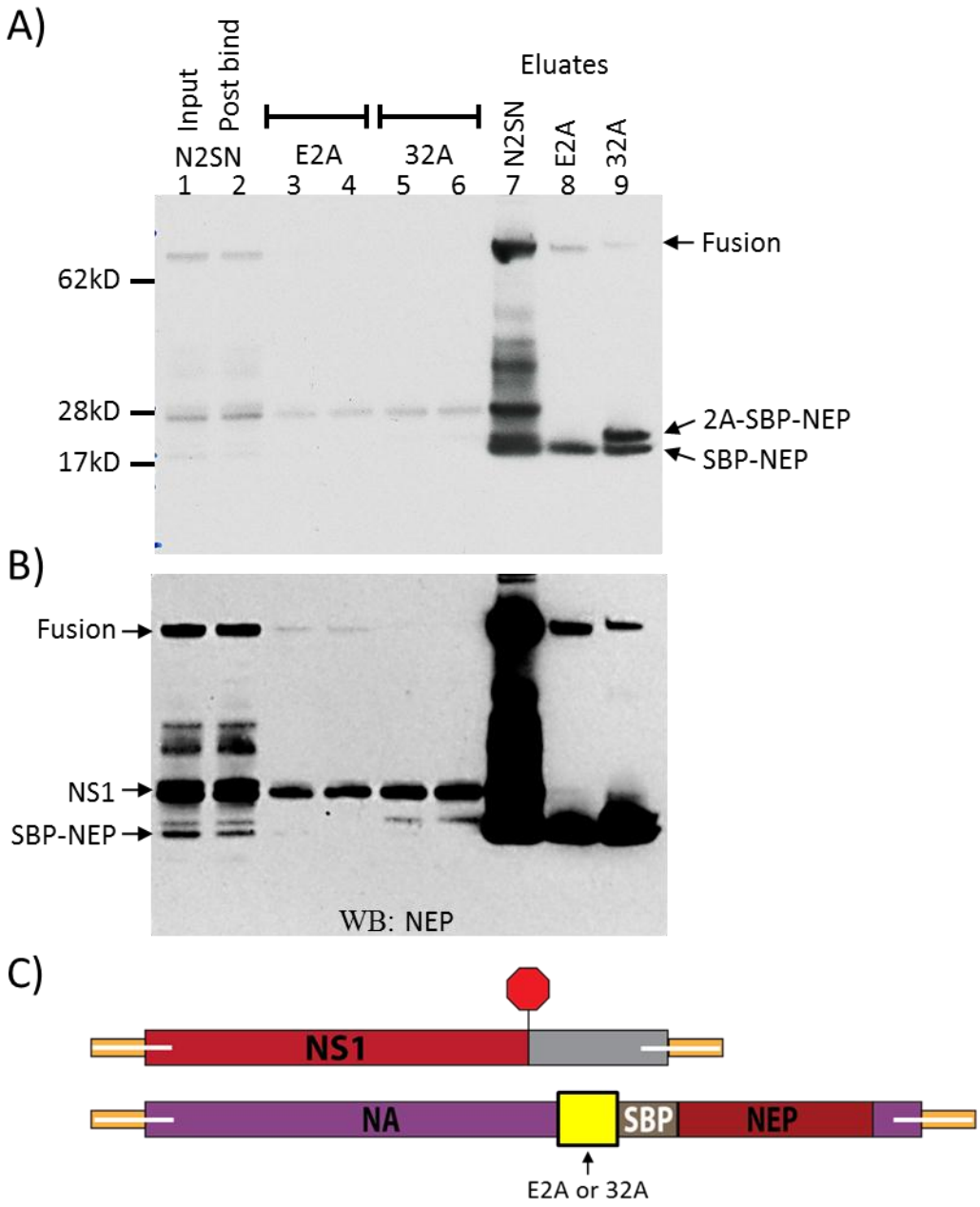


Figure 17. Western blot of an AP for SBP from cells infected with N2SN, E2A and 32A viruses

Cells were infected for 20 hours, frozen, then lysed in PBS + 0.5% NP-40 with protease inhibitors. Cleared lysates were added to M280 streptavidin Dynabeads for 30 minutes on ice. Beads were washed 4x with lysis buffer and then eluted with 1x LDS sample buffer. Input, post bind and eluates from the IP were analyzed by western blotting for NEP. Blots (A) and (B) are two different exposures of the same western blot. The first 6 lanes are input and post bind samples for N2SN, E2A, and 32A affinity purifications while the last three lanes are corresponding eluates. The NS and NA segments for the E2A and 32A viruses are shown in (C).

Immunopurification conditions were examined for the SBP-NEP pulldown by first comparing two different buffers. The manufacturer recommended buffer for affinity purifications with M280 Dynabeads is PBS so we compared a PBS buffer with a HEPES buffer (HB) (formulation detailed in 2.3.7). Duplicate cultures of HEK293T cells were infected with either PR8 or 32A virus to create a total of 4 samples. Lysis, binding and washing were done in parallel in the two buffers and the eluates were compared by Coomassie blue staining. In Figure 18A, the first two protein samples are from the HB IP and the other two are from the PBS IP. Ideally, all proteins interacting with SBP-NEP would be captured with minimal contamination from non-specific background but in practice there is a tradeoff. The stringency of the IP conditions must be balanced with the amount of permissible contaminants. Results shown in Figure 18A indicate that HB allowed too much background binding (lane 1) while the PBS buffer seemed to only capture a few interactions (lane 4). In an attempt to find middle ground, a combination of the buffers was tested: the lysis, binding and first wash were performed with HB or PBS and the second and third washes were performed with the other buffer. The results from this mixed buffer IP are shown in Figure 18B. Lanes 1 and 2 are eluates from affinity purifications in HB followed by PBS washes while lanes 3 and 4 are from PBS affinity purifications with HB washes. Results from Figure 18(A and B) suggest that HB facilitated the capture of more interactions than PBS. Washing the beads from the HEPES IP with PBS seemed to remove a number of background interactors in the PR8 IP eluate (lane A1 versus B1) while the protein composition of the 32A eluate was largely maintained (lane A2 versus B2). When beads were bound in PBS buffer, no significant difference was apparent between the different washing conditions (lanes A3 and A4 versus B3 and B4). Subsequent affinity purifications were performed with the mixed buffer procedure (HB bind with PBS wash). Since the capture of direct interactions was desired, the banding patterns from affinity purifications with or without a nuclease wash were compared. Benzonase is a pan nuclease with activity across wide ranging conditions and was added to beads (or not) in the second wash (100U/mL) for 5 minutes. There was no difference between the banding patterns of the treated versus untreated eluates detectable by Coomassie blue staining (Figure 18C) and therefore we reasoned that the majority of the proteins captured were not bridged through DNA or RNA. Last we sought to

assess the inherent variability in the IP eluates by comparing two sets of affinity purifications in parallel under the same conditions. In Figure 18D, eluates from the first AP (lanes 1 and 2) had slightly more total protein and some low MW bands were present in the first PR8 sample compared to the second PR8 eluate (lane 1 versus 3). Overall the eluates remained relatively consistent even when compared between different experiments (eg. Figure 18B lanes 1 and 2 compared to Figure 18D lanes 1 and 2).

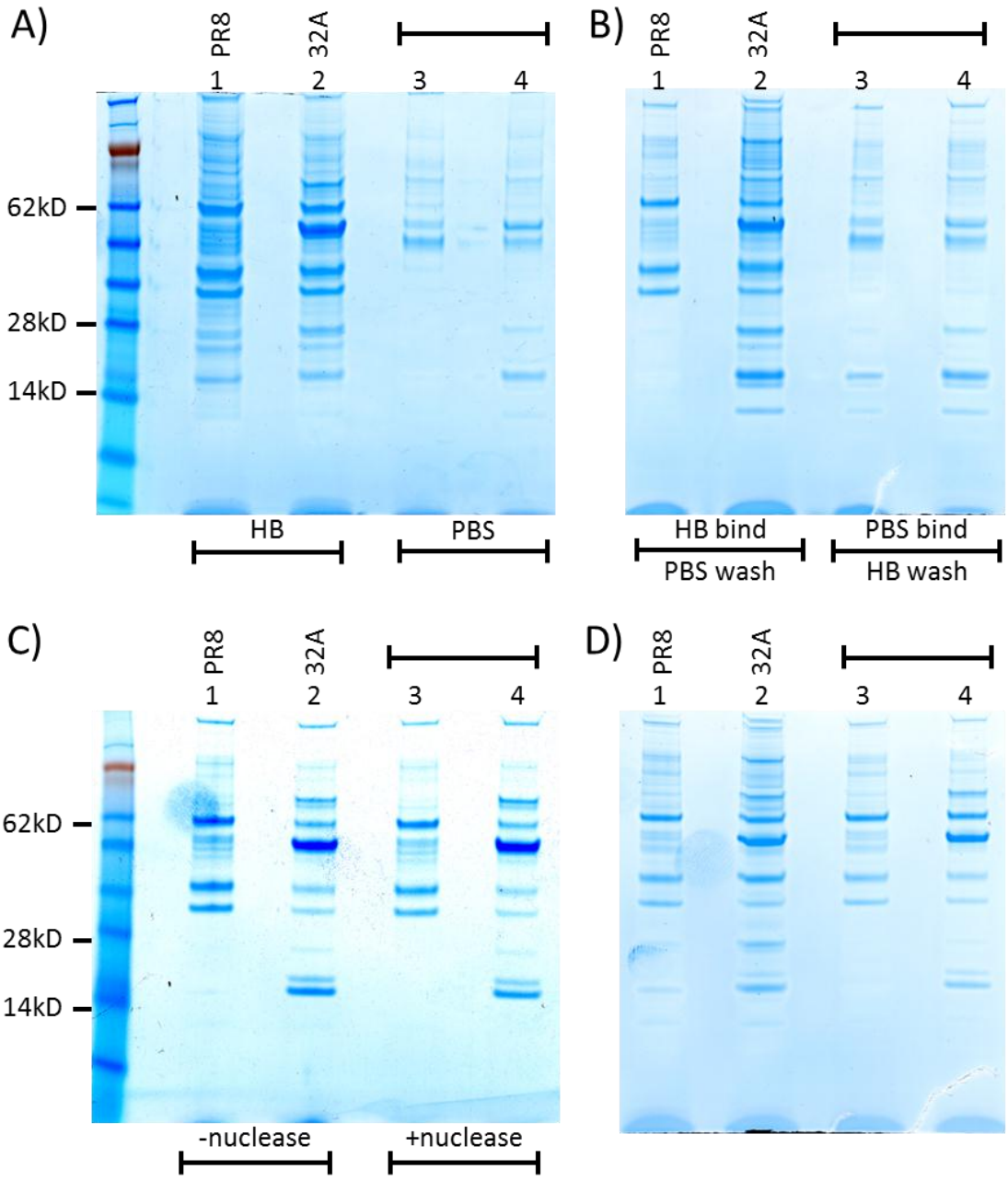


Figure 18. Optimization of SBP affinity purification conditions

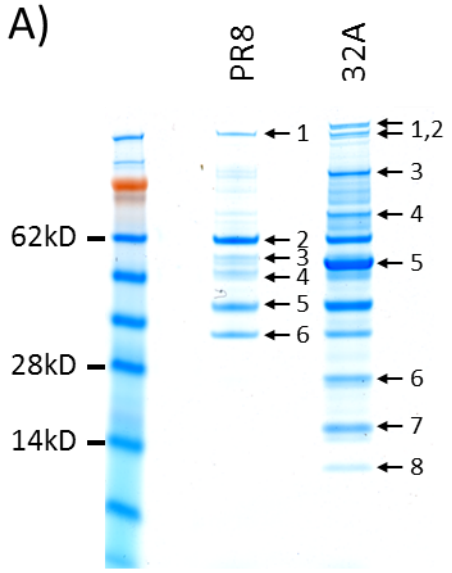
Lysates from cells infected with either PR8 or 32A virus were subjected to AP for the SBP tag using M280 streptavidin Dynabeads under different conditions. For each comparison, a total of four samples were infected (2x PR8 and 2x 32A) and processed in parallel to minimize variability. Beads were eluted with LDS sample buffer and the entire sample was separated by SDS-PAGE and stained with Coomassie blue dye. In (A), PBS and HEPES-based (HB) buffers were compared. In (B), wash buffers were compared. In (C), a 5 minute wash +/- Benzonase was performed to see if the interactions were bridged by nucleic acids. The last panel (D) simply compared two sets of affinity purifications under the same conditions to assess experimental variability. SeeBlue Plus2 was used as the MW marker.

Proteins in the major bands of the PR8 and 32A AP eluates were identified using mass spectrometry. Bands were excised from the gel and processed for mass spectrometry analysis according the procedure in 2.3.9.1. Figure 19A shows which bands were selected for identification from each sample. Figure 19B shows the most abundant protein identified in each band. The most prominent band in the 32A AP eluate (band 5) was influenza NP and it appeared specific. The large excess of NP over SBP-NEP (32A band 7) suggests that whole vRNPs were being captured. It has been shown that vRNP complexes associate tightly with chromatin to gain preferential access to host nuclear export machinery. This may explain the specific presence of histones (32A band 8) which are a common contaminant in immunopurifications (Trinkle-Mulcahy et al. 2008). Also Garcia-Robles et al. found that M1 but not NEP interacted with histones (Garcia-Robles et al. 2005) and histones have been identified as NEP interactors by IP-MS using FLAG-tagged NEP (Gorai et al. 2012). Influenza M1 (32A band 6) interacts with both vRNPs and NEP and was also specific to the 32A AP. Figure 19, 32A band 4 was influenza PB2 which interacts with the vRNP and also with NEP (Brunotte et al. 2014). Aside from histones (and possibly myosin-10, and spectrin (bands 1 and 2)), angiomotin (band 3) was the only host protein prominently enriched in the 32A AP.

Angiomotin (AMOT) is a scaffolding protein important for the maintenance of tight junctions and has also been shown to interact with membranes. Myosin-10 (MYH10) is an unconventional myosin that can interact with both microtubules and F-actin and is involved in anchoring of the mitotic spindle. Spectrin (SPTAN1) is a major component of the cell cortex. AMOT, MYH10 and SPTAN1 could be involved in the budding process since cytoskeletal and plasma membrane components are important for budding (Avalos, Yu, and Nayak 1997; Nayak, Hui, and Barman 2004). The PR8 bands 6, 5 and 3 were all components of the pyruvate dehydrogenase complex and were also present in the 32A AP. PR8 band 2 was also present in the 32A AP and was identified as bovine serum albumin which was present in high concentrations in the cell culture medium and was likely carried through the AP procedure. The major component of PR8 band 4 was a lipoamide acyltransferase component of the mitochondrial branched-chain alpha-keto acid dehydrogenase complex (DBT). The PR8 bands 3 and 4 were probably present in the 32A eluate as well, but were masked by NP.

The results in this section show that processing efficiency of the 2A peptide sequence can be improved by linking multiple copies in tandem. The background in the SBP affinity purifications was high and persistent despite efforts to optimize the procedure. As a result, further development was necessary to reduce the background in the control AP and the next section describes those efforts.

A)



B)

Band	PR8	32A
1	PRPF8	SPTAN1
2	BSA	MYH10
3	PDHB	AMOT
4	DBT	PB2
5	PDHA	NP
6	PDHB	M1
7		3xFLAG-NEP
8		Histones

Figure 19. Identification of major protein bands in the affinity purifications for SBP
Affinity purifications for the SBP tag were performed on infected (PR8 or 32A) HEK293T cell lysates using the mixed buffer procedure described in 2.4.2.3. Coomassie-stained major gel bands (indicated by arrows) from the eluates were excised from the gel and subjected to mass spectrometry. (B) is a list of protein identifications from mass spectrometry analysis of the correspondingly numbered bands in (A).

2.4.2.4 Recombinant influenza virus NA-32A-3xFLAG-NEP (32A-3xFLAG)

Based on results presented in 2.4.2.3, a significant portion of 32A AP eluate was either background or known interactions. To improve the quality of the AP eluates we substituted the 3xFLAG tag for SBP. A western blot of HEK293T cell lysates from 20 hour infections with 32A-3xFLAG (Figure 20) shows that production of native NEP was abolished and that NEP was tagged with 3xFLAG in the 32A-3xFLAG virus. The NEP band in the 32A-3xFLAG lane was shifted higher due to presence of the tag. The NA-32A-3xFLAG-NEP fusion band was significantly reduced while a more prominent band at approximately 40kD was detected. This intermediate sized fusion was recognized by both NEP and FLAG antibodies suggesting that it was a degradation product of the larger fusion protein. In Figure 20B, the same samples were blotted for FLAG. The 3xFLAG-NEP doublet as well as the fusion products were recognized indicating the epitope was present in all of these proteins.

We also performed a time course western blot for the influenza capsid protein M1 to compare the replication kinetics of PR8 and 32A-3xFLAG viruses. For this assay HEK293T cell cultures were infected with sufficient virus to observe visible CPE within 20 hours (as indicated in 2.3.7). Samples were obtained at 2 hour intervals for the first 8 hours then at 4 hour intervals for an additional 16 hours post infection. For both infections the M1 signal was first detected at 12 hours and increased through the 24 hour time point indicating that replication of the 32A-3xFLAG virus was not impaired (Figure 21A, B). 3xFLAG-NEP was also detected at 12 hours and, in contrast to M1, appeared to rapidly reach a steady state (Figure 21C). No discernable difference in FLAG signal intensity was apparent from 12 to 24 hours.

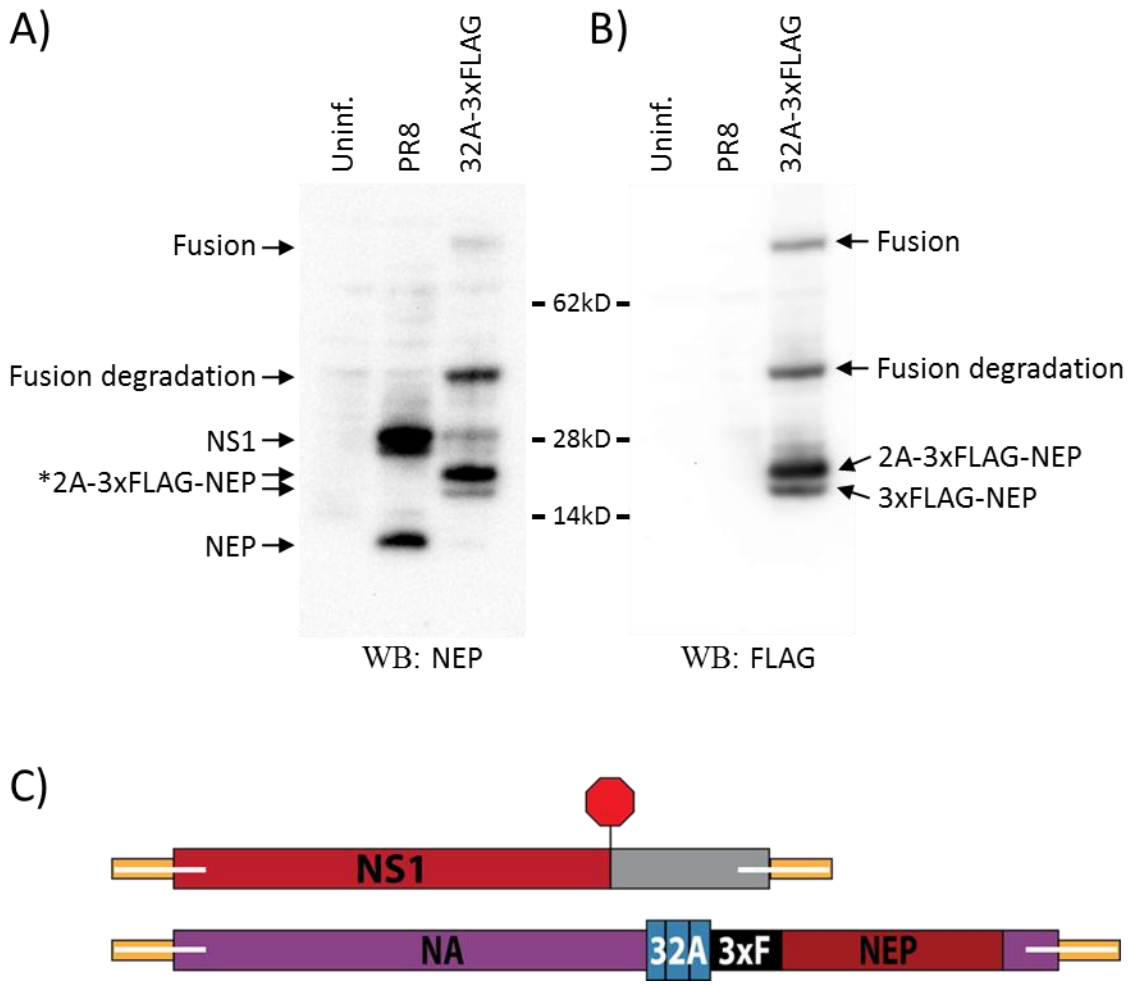


Figure 20. Western blot comparison of PR8 and 32A-3xFLAG infected cell lysates
HEK293T cells were infected for 20 hours and then lysed in protein sample buffer. Equivalent amounts of cell lysate were analyzed by western blot for either NEP (A) or FLAG (B). The MW of 3xFLAG-NEP is 17.4kD. The NS and NA segments from the 32A-3xFLAG virus are shown in (C). In (A) the 3xFLAG-NEP signal was predominantly from the higher MW band in the doublet (indicated by the double arrows).

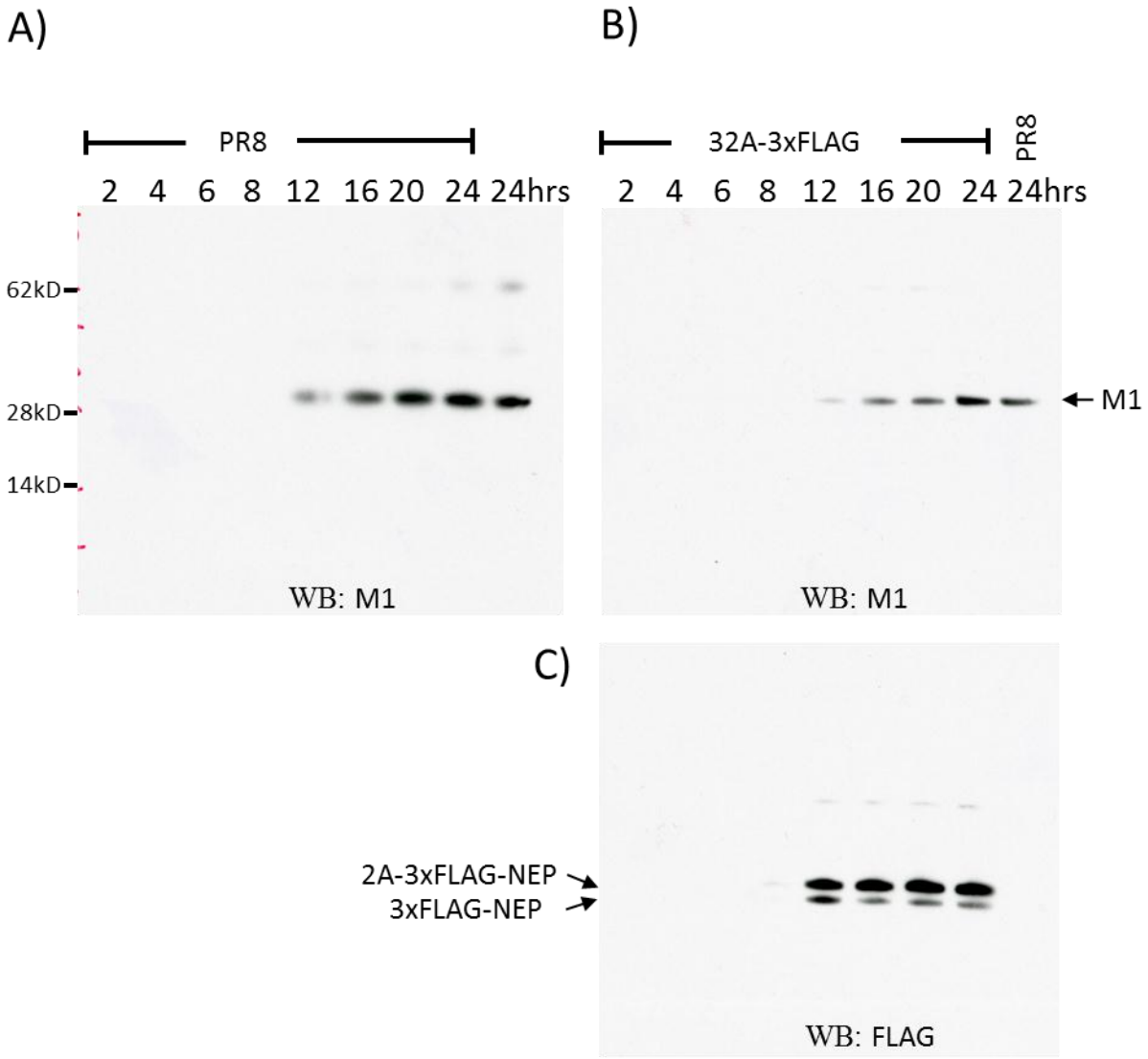


Figure 21. Time course Western blot comparison of PR8 and 321-3xFLAG virus
HEK293T cells seeded 20 hours prior in 12 well plates were infected with 10µL PR8 (A) or 50µL 32A-3xFLAG virus (B) for 2 hours. Media was replaced and the first culture was collected, washed with 1mL PBS then preserved in 150µL LDS sample buffer at -20°C. Samples from the infected cultures were obtained at the indicated times post infection and were blotted for either M1 (A) and (B) or FLAG (C). The 24 hour time sample from the other virus was included in the blots as a control for loading and transfer efficiency. Samples in (C) are the same as in (B).

Once we had established that the 32A-3xFLAG virus produced 3xFLAG-NEP during infection and that it replicated comparably to wild-type PR8 virus, we performed a test IP (according the procedure detailed in 2.3.8) for FLAG on 32A-3xFLAG infected HEK293T cells. After infection for 20 hours, cells were lysed in buffer HB with 0.5% TX-100. Lysates were cleared and supernatants were incubated on Dynabeads-PrG covalently coupled to anti-FLAG M2 antibodies. After washing, an $MgCl_2$ gradient (in binding buffer) was used to elute the interacting proteins. A final elution in LDS sample buffer was used to remove any proteins left on the Dynabeads. Figure 22 shows a western blot of samples from the test IP. Lanes 1 and 2 are from input and post bind samples. Decreased signal intensity in the post bind indicates that the target (3xFLAG-NEP) was being captured from the lysate. The signal from the 40kD fusion degradation band was not attenuated in the post bind whereas signal from the full fusion band was decreased. This suggests that the FLAG epitope from the full fusion protein was recognized by M2 better than FLAG in the intermediate weight fusion. Trace amounts of the higher MW species of 3xFLAG NEP (likely 2A-3xFLAG-NEP) were detected in the 50mM $MgCl_2$ eluate. The 2A-3xFLAG-NEP appears in small amounts in the 100 and 250mM eluates with only a trace amount of 3xFLAG-NEP (the lower MW band from the doublet). The disproportionate signal intensities between 2A-3xFLAG-NEP and 3xFLAG-NEP in these eluates and in the input and post bind may reflect the lower affinity of the M2 antibody for the 3xFLAG with N-terminal 2A sequence. The interaction of 3xFLAG with the M2 antibody was disrupted with 500mM $MgCl_2$ and substantially abolished with 2.5M $MgCl_2$. A ladder of fuzzy bands was observed in the SB eluate (Figure 22 lane 8) and it was unclear what these represented. It is possible that the process used to collect this sample (heating the Dynabeads in SB) altered the way the fusion products migrated during PAGE.

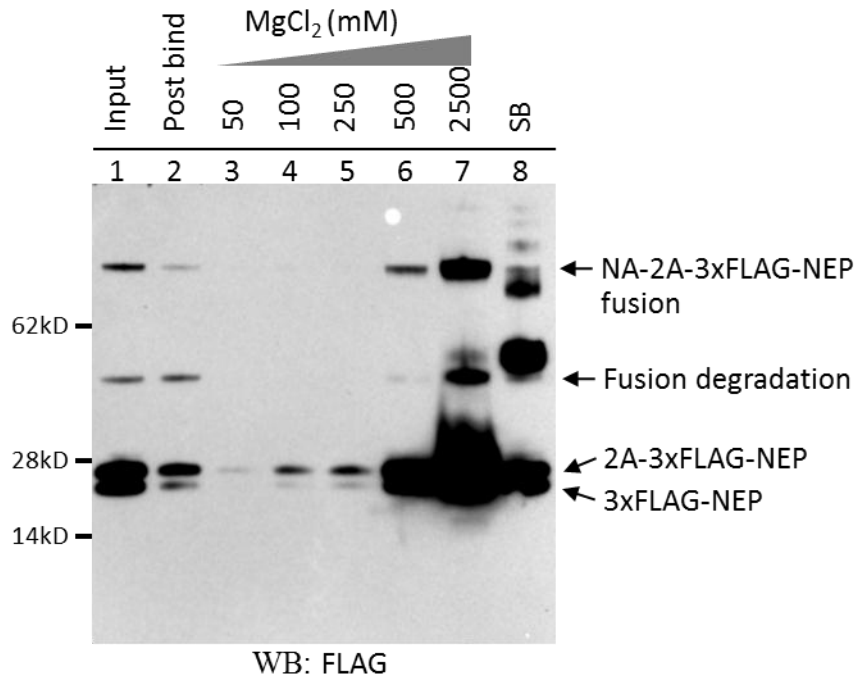
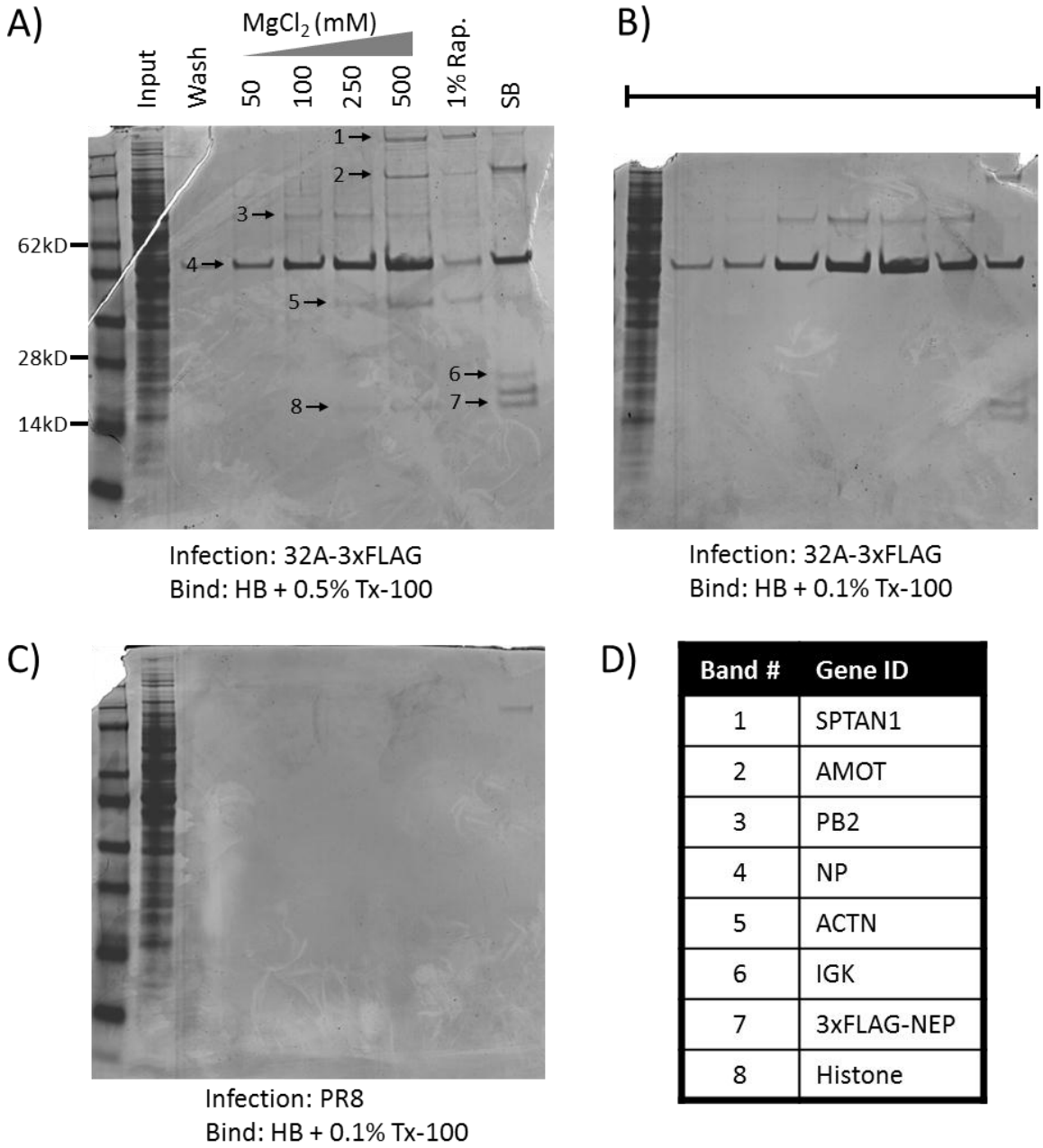


Figure 22. Western blot of a FLAG IP from 32A-3xFLAG infected HEK293T cells
HEK293T cells were infected with 32A-3xFLAG virus and immunopurified for FLAG in HB + 0.5% TX-100. Sequential addition of an $MgCl_2$ gradient was used for elution followed by LDS sample buffer (SB). Western was blotted for FLAG.

Immunopurification conditions for 3xFLAG-NEP were established by comparing detergents and elution conditions. HEK293T cells infected in 15cm plates were lysed in HB with either 0.1% or 0.5% TX-100 (TX-100 is a manufacturer recommended detergent for immunopurifications with M2) and bound to anti-FLAG Dynabeads prepared as indicated in 2.3.8. Beads were washed in binding buffer and a gradient of increasing $MgCl_2$ concentration was used to elute the bound proteins. Following the salt gradient, denaturant (Rapigest) and sample buffer elutions were performed to extract any remaining proteins from the beads. Figure 23A and B compare 0.1% and 0.5% TX-100, while Figure 23B and C compare immunopurifications with 0.1% TX-100 from 32A-3xFLAG or PR8 infected cells. This experiment showed that increasing the amount of TX-100 stabilized many interactions in the 32A-3xFLAG IP and that the background in the control PR8 IP was low.

We identified the major bands (numbered arrows) that were present in Figure 23A by mass spectrometry and the identifications are presented in Figure 23D. SPTAN1 and AMOT (bands 1 and 2) were present again in the FLAG IP. Histones were also identified as before. Actin (band 5) was a new host component in the FLAG IP that was not evident in the SBP AP. The other abundant proteins were all from influenza. NP (band 4) was the most abundant protein in the FLAG IP just as it was in the SBP AP. PB2 was also present again (band 3). In contrast to the SBP AP, there was no visible evidence for M1 in the FLAG IP. A small band from immunoglobulin kappa light chain (band 6) was detected in the sample buffer elution and was probably shed from M2.

We also performed a side-by-side comparison of FLAG immunopurifications from uninfected, PR8 infected, or 32A-3xFLAG infected HEK293T cells (Figure 24). In this case cells were infected and subjected to IP for FLAG as described previously except that the interacting proteins were eluted with 500mM $MgCl_2$ followed by low pH (200mM glycine pH 3), and sample buffer elutions. The interaction between the FLAG epitope and the M2 antibody is disrupted at low pH so the glycine elution should have released 3xFLAG-NEP and any interacting proteins remaining after the high salt elution while contaminants that are non-specifically associating with the capture complex could have remained bound to the beads. Compared to the SBP AP (Figure 18), significantly lower background was observed in the FLAG IP (lanes 1, 2, 4, 5, 7, and 8). Both the $MgCl_2$ and low pH elutates (lanes 3 and 6) contained a significant number of interacting proteins. There were also a number of proteins that remained bound to the captured complex after the glycine elution (lane 9). However the 3xFLAG-NEP band in the LDS elution (lane 9) was less intense compared to the glycine elution (lane 6) indicating that most of the 3xFLAG-NEP had already been liberated.



D)

Band #	Gene ID
1	SPTAN1
2	AMOT
3	PB2
4	NP
5	ACTN
6	IGK
7	3xFLAG-NEP
8	Histone

Figure 23. Coomassie blue stained gels of FLAG IP eluates from different conditions
HEK293T cells were infected for 20 hours with either 32A-3xFLAG or PR8 and then lysed in buffer HB with 0.1% or 0.5% TX-100. Lysates were subjected to IP for FLAG. Interacting proteins were eluted with an $MgCl_2$ gradient followed by denaturant and LDS sample buffer (SB) elutions. Eluates (50% of total) were separated by SDS-PAGE and visualized by Coomassie blue staining. (A) and (B) compare 2 different detergent concentrations (0.5% and 0.1%) on IPs from cells infected with 32A-3xFLAG virus while (B) and (C) compare IPs from cells infected with PR8 or 32A-3xFLAG virus. Protein IDs from mass spectrometry analysis of the correspondingly numbered bands in (A) are shown in (D).

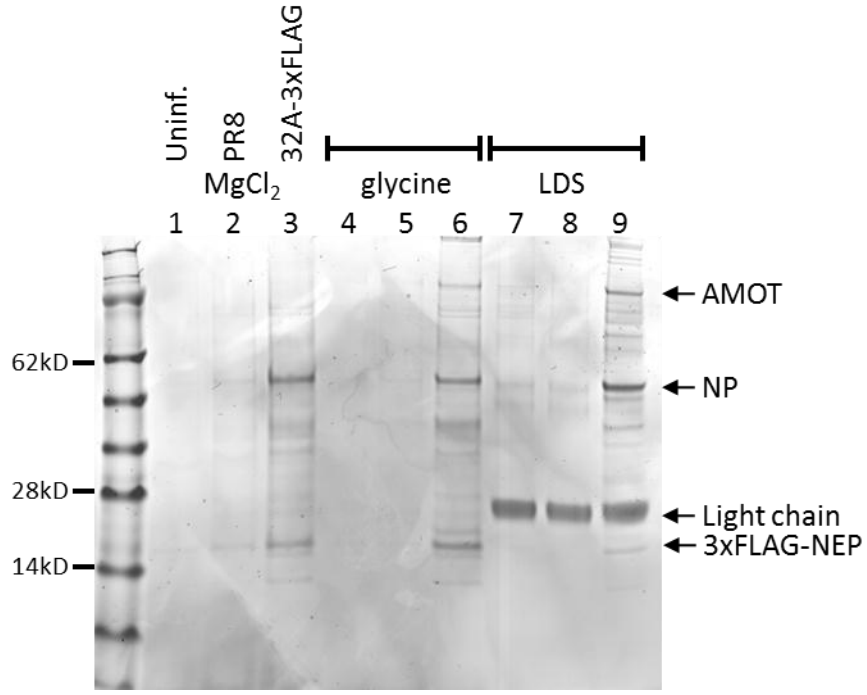


Figure 24. Coomassie blue gel of FLAG IP eluates from HEK293T cells
HEK293T cells were infected (or not) for 20 hours with either 32A-3xFLAG or PR8 and then an IP for FLAG was performed on cleared cell lysates in HB + 0.5% TX-100. Beads were eluted sequentially with 500mM $MgCl_2$, 200mM glycine pH 3, and LDS sample buffer. Proteins in the eluates were separated by SDS-PAGE and visualized by Coomassie blue staining. Bands were identified based on visual similarity to the bands in Figure 23A.

2.5 Discussion

In order to comprehensively identify host cell interactors of influenza NEP, we produced viral strains that express a tagged version of the protein to enable IP-MS experiments. We first expressed and purified a recombinant GST-NEP fusion from *E. coli* and then inoculated rabbits with the fusion in order to generate polyclonal serum. Commercially available antibodies for NEP did not work well and it was important to be able to reliably detect NEP in infected cells by western blotting. When purified GST-NEP was incubated with thrombin, the liberated NEP fragment was detected just below the 14kD marker (Figure 13B). In addition to the expected proteins (NEP, GST and uncleaved GST-NEP fusion), two higher MW species were also detected in the thrombin-treated samples. These were mostly likely GST dimerization products. GST has a well-established propensity to form dimers with equilibrium dissociation values in the sub-nanomolar range (Fabrini et al. 2009; Ji et al. 1992). After thrombin cleavage free GST may have dimerized into homo (GST/GST) and heterodimers (GST/GST-NEP). The relative MW of these species compared to the standard supports this conclusion (52 and 66kD respectively). Protein samples for SDS-PAGE analysis typically included a reducing agent (50mM DTT), but it is possible this was mistakenly omitted from these samples. The serum does recognize NEP in infected cell lysates. Wild-type NEP ran slightly below the 14kD marker (Figure 14B). In some cases the serum produced significant background but the results were interpretable when proper controls were included such as uninfected cell lysates (Figure 16). Because NEP and NS1 share N-terminal residues, the serum also recognized NS1. We used this feature to our advantage since the production of NS1 can be used as a proxy for viral replication and infection. A strong NS1 band in the westerns indicates robust replication while little or no signal may indicate less viral activity.

Influenza genomic segment 8 is alternatively spliced to produce two proteins, NS1 and NEP. In the wild-type arrangement, N-terminally tagging NEP would disrupt NS1 and possibly splicing as well. A C-terminal tag would likely disrupt the packaging signal that must be maintained for proper incorporation of the segment into viral particles. To ensure proper translation of NS1 and abolish production of native NEP, we mutated the splice acceptor site on

segment 8 (the NS segment). Sequence coding for 2A, the SBP tag and NEP was added to segment 6 (the NA segment) downstream of NA. Sequence from the 5' end of segment 6 (160bp) was repeated after 2A-SBP-NEP to ensure packaging signals were intact (influenza gene segments are negative sense therefore they read 3' to 5'). Schematic diagrams of the wild-type and engineered segments for N2SN are presented in Figure 12. By replacing the native pHW-NS and NA plasmids with pHW-NS1-HAstop and pHW-NA-2A-SBP-NEP, recombinant influenza virus (N2SN) was generated that produced an SBP-tagged NEP during infection (Figure 14). This first generation virus validated the engineering strategy and proved that the essential functions of NEP were maintained in the presence of an N-terminal tag. However, N2SN produced a preponderance of NA-2A-SBP-NEP fusion protein which was attributed to poor processing of the 2A peptide.

We tested several new constructs including other self-cleaving peptide sequences for reduced fusion production (Figure 16). We confirmed that sequence upstream of the 2A peptide does affect processing efficiency, as viruses encoding the HA-2A and RSEV-2A sequences produced significantly different amounts of fusion despite having the same 2A and also having robust viral activity as indicated by the strong NS1 bands (Figure 16A lanes 5, 6, 7, 8 versus B lanes 1, 2, 3, 4). NS1 was absent or attenuated from infection of HEK293T cells with recombinant T2A virus as was the fusion (Figure 16A lanes 9, 10, 11, 12). This points to impaired viral fitness and could mean that the T2A sequence does not lead to cleavage. Alternatively, there may be a nucleic acid sequence motif in T2A that is functionally incompatible with the viral lifecycle. The T2A sequence failed to yield virus when tested with the 3xFLAG tag as well (data not shown). E2A virus produced much less fusion (Figure 16B lanes 9, 10, 11, 12) while a triple-tandem repeat of the 2A peptide (32A) resulted in the most complete processing (Figure 16B lanes 5, 6, 7, 8 and Figure 17).

SBP-NEP from the 32A virus was observed as a doublet rather than as a single band (Figure 17). We believe this was due to ribosomal skipping on either the second or third 2A sequence. If the first or second sequences are processed, two or one copies of 2A will remain appended to the N-terminus of SBP-NEP. In Figure 17B only the one slightly higher MW SBP-

NEP band predominates in the 32A input and post-bind (lanes 5 and 6) while two bands are detected in the eluate (lane 9) suggesting that either the first or second 2A sequence led to skipping most frequently, while the second or third repeat was processed less often. This would have led to less abundance of the smaller MW SBP-NEP species in infected cells. The enrichment of the smaller MW SBP-NEP band in the 32A eluate likely reflects a higher affinity interaction with streptavidin due to the absence of N-terminal 2A sequence. Since the first 2A repeat had essentially the same upstream sequence as in the N2SN construct, it seems likely that this one was processed poorly and the second and third sequences produced the bands observed in the 32A eluate. This is corroborated by the fact that the lower MW band in the 32A eluate ran at the same size as SBP-NEP from the E2A virus which only had a single E2A sequence. Infections with 32A seemed to produce more SBP-NEP, less fusion protein and resulted in cytopathic effects similar to PR8 (data not shown) so this virus was used for further studies.

Buffer composition can affect protein-protein interactions in unpredictable ways so affinity purification conditions should be optimized through empirical testing (LaCava et al. 2015). We compered two AP-MS buffers and washes: PBS and HEPES. PBS is an isotonic salt solution and is the buffer recommended by the bead manufacturer for binding. HEPES, a zwitterionic buffering molecule whose pK_a lies very close to physiological pH, is widely used in cell culture and for cryopreservation as it provides maximal buffering at the relevant pH range and across a wide temperature range (Baicu and Taylor 2002; Good et al. 1966). HEPES is also a common buffer for studying protein interactions (LaCava et al. 2015) and we found that lysis and binding in HB followed by washing with PBS balanced the capture of many interactors of SBP-NEP with non-specific background in the PR8 control AP (Figure 18). We also tested whether or not the presence of Benzonase affected the proteins captured in the AP. Benzonase digests all forms of nucleic acid to small oligonucleotides 3-5 bases in length and is active in wide ranging conditions (Liao et al. 2007). No difference in the banding pattern was observed (Figure 18C) indicating that the interactions captured by the AP were not bridged by DNA or RNA fragments accessible to Benzonase. We also tested side by side affinity purifications under

the same conditions to assess the technical variability. There were more proteins in the first set of AP eluates than in the second but the dominant band pattern was consistent (Figure 18D).

Selected bands from each AP were identified by mass spectrometry (Figure 19). Highly prominent bands in the PR8 AP were identified as members of the pyruvate dehydrogenase complex (PDC) and were consistently observed in both eluates as was BSA. The presence of BSA is readily explained since it was present at high concentrations (100mg/mL) during cell culture and could be carried through the AP procedure. However it is less clear why PDC members were so abundant. The PDC is a large multienzyme complex that converts pyruvate to acetyl-coenzyme A and thereby serves as bridge between glycolysis and the tricarboxylic acid cycle. Band intensity of the PDC members was independent of the presence of the SBP tag suggesting that they do not compete for the same interaction site. The PDC was absent from subsequent immunopurifications using a different Dynabeads capture complex (Figure 24) implying that it was interacting specifically with streptavidin.

Half of the bands selected for identification in the 32A IP eluate were influenza proteins and interestingly the most prominent band was NP. NP monomers bind vRNA with high affinity approximately every 24 bases (approximately 70 copies per gene segment) to form the twisted anti-parallel double helix structure characteristic of vRNPs (Eisfeld, Neumann, and Kawaoka 2014). The stoichiometric excess apparent between NP and SBP-NEP suggests that entire vRNP complexes were being isolated. PB2 and M1, also identified, have both been shown to interact directly with NEP and are associated with vRNPs (Figure 5). AMOT, SPTAN1, MYH10 and histones were the identified host proteins specifically present in the 32A AP. SPTAN1, MYH10, and histones have all been identified as common contaminants in Dynabead immunopurifications (Trinkle-Mulcahy et al. 2008). This leaves AMOT as the single host factor identified in the 32A IP that was not a common contaminant. Except for histones, these proteins have not previously been shown to interact with influenza proteins before and have not been implicated as important for the life cycle (T. Watanabe, Watanabe, and Kawaoka 2010). Based on the results from the SBP AP optimization and mass spectrometry we did not feel that the SBP tag in combination with the M280 streptavidin Dynabeads was well suited to

NEP interactome studies. Nearly all of the most prominent bands in the 32A AP were from either known influenza interactors or were from background contaminants. We believed that additional optimization with an alternative tag would lead to lower background and result in an IP better suited to deep exploration of the NEP interactome.

Since the streptavidin AP for SBP contained significant contaminating bands, we generated a third generation of virus with the 3xFLAG tag. The FLAG epitope has been widely adopted and numerous commercial reagents are available that target the tag. Since we had previously included restriction sites for rapid testing of new tag/2A sequences (Figure 15) we were able to exchange the SBP tag for 3xFLAG. The new 32A-3xFLAG virus was qualitatively similar to the previous 32A virus. Western blot analysis of 32A-3xFLAG infected cell lysates indicates that the 3xFLAG-NEP was expressed whereas native NEP was not (Figure 20A). The band for the full fusion protein was less intense in the 32A-3xFLAG virus but a band of approximately 40kD was more prevalent and was recognized by both NEP and FLAG antibodies. This suggests that the NA portion of the fusion was cleaved at a site upstream of the 2A-3xFLAG-NEP sequence. The background was very low in the western blot for FLAG illustrating the high specificity of the M2 antibody (Figure 20B).

The titers of the recombinant viruses were lower than PR8 titers in all cases. This is based on the observation that more recombinant viral supernatant (from the MDCKII cell amplification step detailed in 2.3.6) was needed to achieve the same CPE produced by PR8. In order to assess whether the recombinant 32A-3xFLAG virus had replication defects compared to PR8, we sampled infected cell cultures over time and analyzed them by western blot for either M1 or FLAG (Figure 21). The kinetics of viral replication appeared similar between the two strains. M1 signal was detected at 12 hours in both infections and the rate of increase was comparable as were the final intensities at 24 hours. We concluded that the reduced titers of the recombinant viruses were not due to impaired replication. In the recombinant viruses NA had 16 (or more depending on the construct) extra C-terminal residues from 2A. It is probable that the lower titers are related this extra sequence since the C-terminal tails of NA and HA impact viral morphology (Chen et al. 2008; Jin et al. 1997). The FLAG signal appeared suddenly

at 12 hours and did not seem to increase in intensity over time. It is important to note that the FLAG signal was actually the result of transcription and translation of the NA segment. Trace amounts of the 40kD fusion protein were detected while the full length fusion was not detected indicating that the 32A sequence was being processed effectively.

We used the M2 antibody in combination with PrG-Dynabeads for our FLAG immunopurifications. Streptococcal Protein G (PrG) recognizes immunoglobulins with high affinity and this property has been exploited in a number of ways (Sjöbring, Björck, and Kastern 1991). We used commercially available Dynabeads that had been coated with PrG to bind the M2 antibody and then covalently coupled the PrG-M2 complex with DMP, an amine-reactive 7 atom crosslinking reagent. Crosslinking the PrG-Dynabeads to M2 prevents dissociation of the antibody-bead complex during the IP procedure and reduces M2 antibody contamination in the IP eluates. The FLAG IP procedure (described in 2.3.8) was largely the same as that for SBP except that more cells were used (15cm plate vs. 10cm previously), binding was shortened to 20 minutes and conducted at RT and NP-40 was replaced with Trition X-100 (TX-100). We previously observed more protein in eluates from affinity purifications in HB (Figure 18B) so we also used this buffer for the FLAG immunopurifications.

A western blot of a FLAG IP from 32A-3xFLAG infection was performed and reveals several important features (Figure 22). Depletion of the signal from input to post bind indicates the target was being captured and immobilized on by the Dynabeads-PrG-M2 capture complex (Figure 22 lane 1 versus 2). For our 3xFLAG-NEP immunopurifications, an excess of FLAG epitope over M2 antibody was preferred in order to minimize off-target interactions. This appeared to be the case since FLAG signal was present after binding (Figure 22 lane 2). We observed that the 3xFLAG-M2 interaction was mostly maintained in the presence of 100mM $MgCl_2$, but began to dissociate in 500mM $MgCl_2$, and was substantially disrupted with 2.5M $MgCl_2$. The signal intensity of bands in the various samples is also informative. The 3xFLAG-NEP doublet signal was relatively even in the input. However the lower MW band (3xFLAG-NEP) was depleted in the post bind relative to the slightly higher MW band (2A-3xFLAG-NEP) indicating that 3xFLAG-NEP without N-terminal 2A sequence was preferentially recognized by M2. This

conclusion is consistent with the observation that the 2A-3xFLAG-NEP band was also more intense in the 50 and 100mM $MgCl_2$ eluates indicating a weaker interaction with M2. Interestingly, the high MW fusion (NA-2A-3xFLAG-NEP) was depleted in the post bind while the 40kD fusion was not. Figure 20B indicates that the 40kD fusion band was specific to cells infected with 32A-3xFLAG so this could mean that, for this species, the 3xFLAG epitope was sequestered from M2 in solution.

We tested the effect of different detergent concentrations in HB (Figure 23). For immunopurifications with M2, the manufacturer recommended NP-40 concentration should not exceed 0.1% so we tried an alternative detergent TX-100 which can be included up to 5%. We observed a significant increase in the number of proteins recovered in 0.5% TX-100 compared to 0.1% (Figure 23A versus B). We used an $MgCl_2$ gradient to separate the eluate into fractions and bands were excised from the gel for mass spectrometry identification. Influenza NP was observed in every fraction including the wash (Figure 23A, band 4). As it was in the SBP AP, the interaction with NP appeared specific since no NP band was present in the control IP from PR8 infected cells (Figure 23C). A similarly low background was observed in control IPs from with 0.5% TX-100 (data not shown). PB2 was detected (Figure 23A band 3) but M1 was not observed in this IP. The M1-NEP interaction may have been stabilized by the presence of NP-40 in the SBP AP. SPTAN1 and AMOT were again identified (bands 1 and 2). These two proteins did not appear in the 250mM eluate but did in the 500mM eluate suggesting an electrostatic basis for the interaction. We tested the denaturant Rapigest because 3xFLAG-NEP remained bound to M2 in 500mM $MgCl_2$. Rapigest is an acid-labile surfactant that is compatible with mass spectrometry analysis. However the Rapigest elution failed to liberate much of the 3xFLAG-NEP and did not recover more interacting proteins. The 3xFLAG-M2 interaction was substantially preserved through the salt and surfactant elutions and was only detected after addition of LDS sample buffer (band 7).

We performed a side-by-side comparison of immunopurifications from infected (PR8 and 32A-3xFLAG) and uninfected HEK293T cells (Figure 24). We wanted to observe the background and also to test a low pH glycine elution. Ultimately we were optimizing the IP to

capture host proteins interacting with NEP for identification by mass spectrometry. It was clear that many proteins were liberated from the capture complex with 500mM $MgCl_2$, but also that much of the 3xFLAG-NEP remained bound. We wanted to release 3xFLAG-NEP from M2 to ensure we identified high-affinity interactors and could have done so with an anionic detergent like sodium dodecyl sulfate (SDS) but this would have also released background contaminants interacting non-specifically with the capture complex. Since the FLAG-M2 interaction is disrupted at low pH we followed the 500mM $MgCl_2$ elution with 200mM glycine at pH3. The low pH elution liberated 3xFLAG-NEP and a number of other proteins (Figure 24 lane 6). Although both the 500mM $MgCl_2$ and the glycine eluates from the 32A-3xFLAG IP contained substantial amounts of protein, there were still a number of proteins left on the beads as shown by the LDS eluate. However the band corresponding to 3xFLAG-NEP appeared strongest in the glycine fraction suggesting that most of the specific host interactors were captured in the first 2 elutions. A significant number of proteins were observed in the 32A-3xFLAG LDS elution (Figure 24 lane 9) and it seems paradoxical that less 3xFLAG-NEP was observed. However it could be that the low pH elution increased non-specific interactions between proteins still associated with the capture complex after the high salt elution. The intense band in the LDS eluate in all samples was most likely light-chain from M2.

Ultimately we are seeking new insights into the process of viral infection through the totality of the host-virus interactome. Rather than identify single interacting host proteins by Coomassie blue staining, we would like to identify all of the proteins in the NEP immunopurifications in a high throughput, quantitative manner. Our strategy for doing so, I-DIRT, requires mixing of control and experimental cell lysates into a single combined IP. As a result, Coomassie blue staining does not give insight into the quality of the IP. It is, therefore, highly important to develop optimal IP protocols and conditions as we have done in this chapter before conducting quantitative proteomics experiments.

Chapter 3: Influenza NEP Interactome

3.1 Summary

Currently, there is no publicly available set of host protein interactors of NEP isolated from an active infection. Using a modified version of the I-DIRT technique, we performed quantitative proteomics studies with the 32A-3xFLAG virus to identify host-virus interactions during infection. I-DIRT ratios were used to rank the identified proteins according to their probability of interaction with NEP; however, this approach did not yield a high confidence list of NEP specific interactors. To determine the subset of host factors interacting with NEP during infection, we generated a complementary dataset from plasmid expression of 3xFLAG-NEP. A visual network model was produced incorporating all of the proteomics data, as well as existing knowledge of interactions curated from the literature, and functional clustering calculated using the DAVID bioinformatics resource. The network analysis revealed a novel enrichment of tight junction proteins in the NEP interactome.

3.2 Introduction

High throughput datasets can be analyzed in a variety of ways and specific tools have been developed to sort, cluster, rank and visualize what would otherwise be a long list of identifications and values. Biological processes are driven by groups of related genes/proteins rather than individual actors, and this is the principle behind enrichment analysis of high throughput datasets. There are a number of different resources such as the Gene Ontology (GO) project that ascribe biological processes, molecular functions, and cellular localization to individual gene products, and these resources can be used to examine the relationship between proteins in a given dataset. The enrichment of annotation terms in an experimental dataset can also be quantified by comparison to a relevant background using statistical methods. In this

manner the significance of the enrichment can be assessed (Huang, Sherman, and Lempicki 2009a). Bioinformatics tools have been developed for enrichment analysis of large datasets , one of which is the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Dennis et al. 2003; Huang, Sherman, and Lempicki 2009b). DAVID has been successfully applied in a variety of contexts including proteomics studies of host-pathogen interactions (Brown et al. 2010; Ray et al. 2012; Weekes et al. 2014), and we used the gene function classification tool in DAVID (Alvord et al. 2007) to identify functionally related groups of genes/proteins enriched in our IP-MS data. Based on the assumption that genes with similar annotation profiles will be functionally related to each other, the tool aggregates annotation terms from 14 public databases into a binary gene-term matrix. The annotation profiles of all gene pairs in the dataset are then assessed using *kappa* statistics (a chance-corrected measure of the co-occurrence of two observations) and genes are clustered based on the *kappa* values using heuristic fuzzy multiple-linkage partitioning. This custom method allows genes to participate in more than one grouping and gave better results than typical approaches like K-means. Once the optimal number of gene groups has been established, the significance of the functional representation is calculated from the enrichment scores of the individual annotation terms associated with each cluster. More specifically, a modified Fisher exact probability (EASE score, (Hosack et al. 2003)) is calculated for each individual annotation term in the grouping from a background of all human genes to statistically measure for over-representation. Each cluster's enrichment score is taken as the geometric mean of the EASE scores (both significant and insignificant) for the associated annotations. In this manner, the functional enrichment analysis is focused on the biological modules represented by the clusters rather than individual genes/proteins.

Visualization is another way to gain biological insight into large datasets and is also an important aspect of systems biology approaches (Shannon et al. 2003). Humans are well suited to discerning patterns from complex visual landscapes so programs like Cytoscape (Shannon et al. 2003) have been developed to create graphical representations of large amounts of data. Cytoscape is particularly useful for generating network diagrams that show relationships between entities through connecting edges and attributes. Nodes in biological networks are

typically genes or proteins but could also be cells or individuals. Edges represent attributes that connect nodes such as physical interaction or co-regulation. By layering on additional information through color, shape, or other visual modalities, data from diverse sources can be displayed in a single pictorial representation. We used Cytoscape to create a visual model of our proteomics data which revealed a previously unexplored connection between NEP and tight junctions.

3.3 Methods

3.3.1 Stable incorporation of labeled amino acids in cell culture (SILAC)

SILAC cultures were produced with low passage HEK293T cells grown in SILAC DMEM (Thermo, 89985) supplemented with 10% dialyzed FBS (Cellgro or Thermo Scientific; heat-inactivated in-house), 1x pencillin/streptomycin, 1x non-essential amino acid solution (Gibco, 11140), 0.398mM heavy L-arginine-HCl ($^{13}C_6$, $^{15}N_4$), 0.798mM heavy L-lysine-2HCl ($^{13}C_6$, $^{15}N_2$) (Cambridge Isotope Laboratories or Sigma-Aldrich), GlutaMAX, and 200µg/ml L-proline (Sigma-Aldrich) for a minimum of 5 passages. Light 293T cells were produced as above except with the substitution of L-arginine and L-lysine (Fischer scientific, BP372 and BP386). Cultures were grown in 10cm plates. Cells were maintained at 37°C and 5% CO_2.

3.3.2 Viral I-DIRT Immunopurification

SILAC cultures of HEK293T cells in 10cm dishes at approximately 80% confluency were infected with sufficient virus to achieve modest CPE within 13 hours (100µL Pr8 (light) or 1mL 32A-3xFLAG (heavy)). Cells were harvested by centrifugation at 13 hours post infection, washed in 1mL PBS, and stored at -20°C. Frozen cell pellets were directly lysed in 800µL lysis/binding buffer HB (20mM HEPES pH 7.4, 110mM KOAc, 2mM $MgCl_2$, 0.5% Triton X-100 (TX-100)) and 1x protease inhibitor (Sigma, P8340), and then sonicated for 15 seconds. Following centrifugation (21,000 x g, 2 minutes), the soluble fraction from one heavy and one light culture was combined and immediately added to 300µL washed PrG-Dynabeads covalently coupled to anti-FLAG M2 (Sigma, F3165). Beads were prepared as follows: 600µL washed PrG Dynabeads (Invitrogen, 10003D) were rotated with 80µg M2 antibody in citrate-phosphate buffer (25mM citric acid, 50mM Na2PO4, pH 5) for 1hr. Bead-antibody complexes were washed in 1mL 0.2M triethanolamine (TEA) pH 8.2 and then crosslinked with 20mM dimethyl pimelimidate (DMP) in TEA for 15 minutes. The reaction was stopped by washing in 50mM Tris pH 7.5 for 15 minutes. The mixed cell lysates were bound for 10 minutes at RT with rotation. Beads were washed 2x in binding buffer and transferred to a new 1.5mL Eppendorf tube with a 3rd wash. The beads

were eluted with 500mM $MgCl_2$ (50µL) followed by 200mM glycine pH 3 (50µL) with heating at 42°C for 10 minutes.

3.3.3 Immunopurification for FLAG from plasmid expression

HEK293T cells at 75% confluency in 15cm plates were transfected with either 30µg pAc3xFLAG-NEP or 10µg pAc3xFLAG-GFP using 80 or 30µL PEI in 2mL OptiMEM. On day 3 post transfection, cultures were harvested pipetting into a 15mL conical tube. Pelleted cells were washed with 1mL PBS and froze. Frozen cell pellets were incubated on ice for 10 minutes in 1.6mL buffer HB + 0.5% TX-100 then sonicated for 20 seconds. Following centrifugation (21,000 x g, 2 minutes), the soluble fraction was added to 300µL washed PrG-Dynabeads covalently coupled to anti-FLAG M2 (Sigma, F3165) prepared as in 2.3.8. Three pellets were combined for the 3xFLAG-NEP IP compared to 1 pellet for the 3xFLAG-GFP IP. Lysates were bound for 30 minutes at RT with rotating. Beads were washed 1x in binding buffer, followed by two PBS washes and transferred to a new 1.5mL Eppendorf tube with a fourth PBS wash. The beads were eluted with 200mM glycine pH 3 (60µL) with heating at 42°C for 10 minutes followed by LDS sample buffer.

3.3.4 Sample preparation for Mass Spectrometry by in solution digest

Eluted proteins were subjected to an in solution digest as follows: 75µL 100mM NH4(HCO3) pH8.3 (Ambic) was added to 25µL of eluted proteins. 100µL of trifluoroethanol was added with mixing, followed by 8µL tris(2-carboxyethyl) phosphine (200mM) with 30 minutes incubation at room temperature (RT). Next, 10µL iodoacetamide (250mM) was added with 30 minutes incubation in the dark. Solution was diluted with 1mL Ambic, and proteins were digested with 1µg trypsin (Promega, V5280) overnight at 37°C. Digests were cleaned up on Waters Oasis MCX 1cc prepacked columns (Waters, 186000252) as follows. 100µL 10% trifluoroacetic acid (TFA) was added to each sample. Columns were primed with 800µL methanol and equilibrated with 800µL 0.1%TFA. Samples were loaded onto the columns followed by washing with 1mL 0.1%TFA, 1mL 80% acetonitrile/20% 0.1%TFA, and 1mL H2O.

Peptides were eluted with 800µL 10%NH4OH/90% methanol. All peptide samples were concentrated in a speed-vacuum and resuspended in 10µL 0.1% formic acid.

3.3.5 Proteomics Data Analysis

Mass spectrometry data acquisition and analysis was performed as detailed in 2.3.9.2. .RAW files were analyzed with MaxQuant with the Andromeda search engine. The lists of proteins identified with MaxQuant were filtered to remove contaminants including all keratin proteins. They were further filtered to remove proteins identified in only one replicate, and also those proteins identified by only one unique peptide sequence. The gene function classification tool from the Database for Annotation, Visualization and Integrated Discovery (DAVID) was used to identify groups of functionally related proteins. The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) was used to identify known interactions among the host proteins identified in the IP-MS data (Franceschini et al. 2013). Cytoscape was used to display the IP-MS and associated data.

3.4 Results

3.4.1 Viral I-DIRT IP-MS

Viral I-DIRT immunopurifications were performed as indicated in 3.3.2. A total of 4 immunopurifications were performed on two separate days using the conditions established in chapter 2. Coomassie blue stained eluates (50% of the total volume) from the 4 biological replicates are shown in Figure 25. Samples from the first two biological replicates are shown in Figure 25A. The eluate from one biological replicate had decreased protein concentrations compared to the other likely due to a lower concentration of 3xFLAG-NEP in the IP lysate. In Figure 25B samples of the IP lysates (inputs) and eluates from replicates 3 and 4 are shown. The remaining 50% of the 500mM $MgCl_2$ eluates from both replicates 1 and 2 (Figure 25A) as well as the remaining 50% of the 500mM $MgCl_2$ and glycine eluates from replicates 3 and 4 (Figure 25B) were prepared for mass spectrometry analysis by the in solution digest protocol (3.3.4). The gel lane from Figure 25A IP-2 was fractionated into 16 sections (11-12 and 13-16 were combined for mass spectrometry analysis) and prepared with the in gel digest procedure (2.3.9.1).

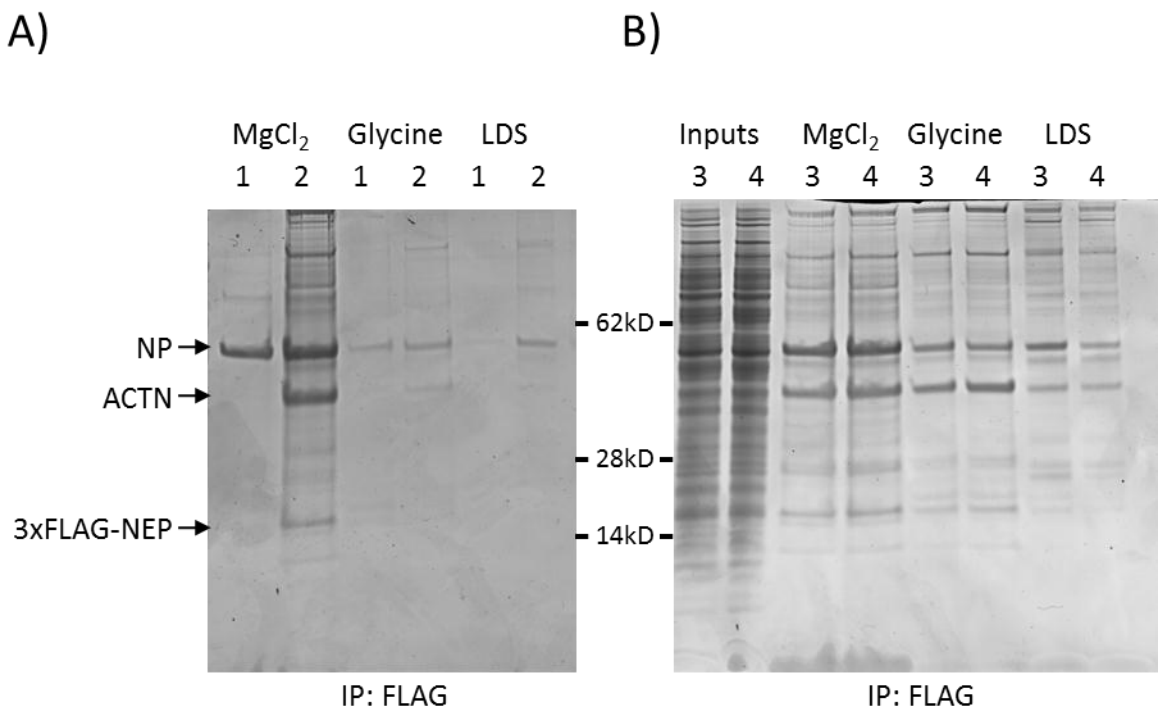


Figure 25. Viral I-DIRT immunopurifications
SILAC cultures were infected with either PR8 (light) or 32A-3xFLAG (heavy). Cleared lysates from the infected cultures were mixed in equal ratio and immunopurified for FLAG. Beads were eluted sequentially with 500mM $MgCl_2$, 200mM glycine pH 3 with heating at 42°C for 10 minute, and LDS sample buffer. The eluates (50%) were analyzed by SDS-PAGE with Coomassie blue staining. (A) shows Coomassie blue stained eluates from 2 biological replicates (1 and 2). (B) shows Coomassie blue stained input samples and eluates from a repeat experiment with two biological replicates (3 and 4).

A total of 17 individual mass spectrometry runs (11 gel fractions and 6 in solution digests) were conducted on an LTQ Orbitrap Velos instrument as detailed in 2.3.9.2. .RAW files were analyzed using MaxQuant with the Andromeda search engine (summary data of the analysis shown in Table 3). Replicates were combined to increase statistical power. A combined total of 184,394 MS/MS spectra were submitted for identification. Of these, 39,173 (21%) were identified and mapped to 3900 unique peptide sequences. For comparison, the success rate of MS/MS identification is generally between 10-30% (Houel et al. 2010). This analysis falls squarely within this range indicating that the underlying mass spectrometry data were of reasonable quality. The average Andromeda peptide identification score was 112 which is also in line with published datasets (Geiger et al. 2012). Using a 1% FDR cutoff for peptide and protein identifications resulted in an overall false positive rate (FPR) for this dataset of <4%. To establish a false discovery rate (FDR), MaxQuant reverses the protein sequences in the search database to create a nonsense decoy peptide set that is also queried against the MS/MS spectra. Scores for both decoy and real peptide/spectra matches are ranked from best to worst. Peptide identifications are accepted until decoys represent a given percentage of the total (typically 1%). The FPR was calculated as follows: $(2 \times n_{decoy}) / (n_{decoy} + n_{real})$ where n_{decoy} was the number of decoy proteins (11 in this dataset) and n_{real} was the number of real proteins identified (576). Based on the reasoning that the number of decoys identified was equal to the number of false positives present, n_{decoy} was doubled. Our FPR is also comparable to published data (Carpp et al. 2014). All subsequent analyses were performed at the protein level.

Table 3. Summary of mass spectrometry data from viral I-DIRT immunopurifications

Sample ID	MS	MS/MS Submitted	MS/MS Submitted (ISO)	MS/MS Identified	MS/MS Identified (ISO)	MS/MS Identified [%]	MS/MS Identified (ISO) [%]	Peptide Sequences Identified
IP(A)-2F1	3169	12096	944	2832	220	23	23	731
IP(A)-2F2	3648	10564	2875	1815	819	17	28	499
IP(A)-2F3	3450	10657	1356	1745	225	16	17	436
IP(A)-2F4	4283	6783	770	1038	139	15	18	268
IP(A)-2F5	3681	8772	1336	2014	335	23	25	587
IP(A)-2F6	4087	7519	1152	1541	238	20	21	429
IP(A)-2F7	2709	13444	2583	2002	362	15	14	285
IP(A)-2F8	3482	9625	1636	2407	372	25	23	430
IP(A)-2F9	2860	13426	1921	2227	396	17	21	178
IP(A)-2F10-13	3520	10781	1346	1585	235	15	17	362
IP(A)-2F14-16	3866	9499	1396	942	209	10	15	235
IP(A)-1MgCl2	3169	11127	2977	1904	660	17	22	352
IP(A)-2MgCl2	2687	13568	2635	4541	621	33	24	1274
IP(B)-3MgCl2	2684	12401	2568	3532	522	28	20	1028
IP(B)-3Glycine	3198	10182	2380	2023	351	20	15	428
IP(B)-4MgCl2	2725	12460	2297	4215	500	34	22	1139
IP(B)-4Glycine	2859	11490	2529	2810	517	24	20	631
Average	3299	10847	1924	2304	395	21	20	547
Standard dev.	494	1935	700	977	182	6	4	313
Total	**56077**	**184394**	**32701**	**39173**	**6721**	**21**	**21**	**3900**

The sample ID corresponds to the labels in Figure 25: IP from (A) or (B), biological replicate (1, 2, 3, or 4), elution method or gel fraction. Samples from the in gel digest of IP(A)-2 are designated by their fraction number since the lane was divided into 16 sections starting with 1 at the bottom. Column 2 indicates the total number of MS1 spectra. Column 3 shows the number of MS2 spectra submitted to the Andromeda search engine. The designator (ISO) refers to those MS2 spectra detected as isotopic patterns due to SILAC pairs.

The protein list was filtered to remove common contaminants, decoy proteins, proteins with fewer than 2 sequences identified, and proteins unique to a single replicate. Proteins without an assigned H/L ratio were also removed. The remaining 265 proteins are presented in Table 4. The proteins are in descending order sorted according to their H/L ratio. The list includes the influenza proteins (bold *) all of which were identified (3xFLAG-NEP/NA, PB2, PB1, PA, NP, M1, HA, NS1, M2). NP was the second most abundant protein by intensity (behind ACTB). NS1, M1, and PB2 were all in the top 30 when ranked by intensity followed closely by PA at 32. The intensity/MW versus H/L was plotted for all proteins in Figure 26. The influenza proteins are represented by yellow dots. The shaded region represents the area of the chart where candidate NEP interactors would be located and indicates the region of the graph where H/L ratios greater than 1 standard deviation would plot. A total of 16 host proteins from our

dataset met or exceeded this statistical cutoff (red dots) and are shaded in blue in Table 4. As expected, the protein with the highest H/L ratio was NEP. PA and PB2 were also present in this group consistent with previously published data as discussed in 1.1.3. ATP5A and ATP5B are known host interactors of NEP (Gorai et al. 2012). The other host factors in this subset were commonly identified contaminants except for EPRS, SSRP1, LRRC49 and UNC45A. Overall, the viral I-DIRT methodology did not yield a distinct separation between stable interactors and contaminants as we had hoped. Reasons for this are explored further in the discussion in section 0.

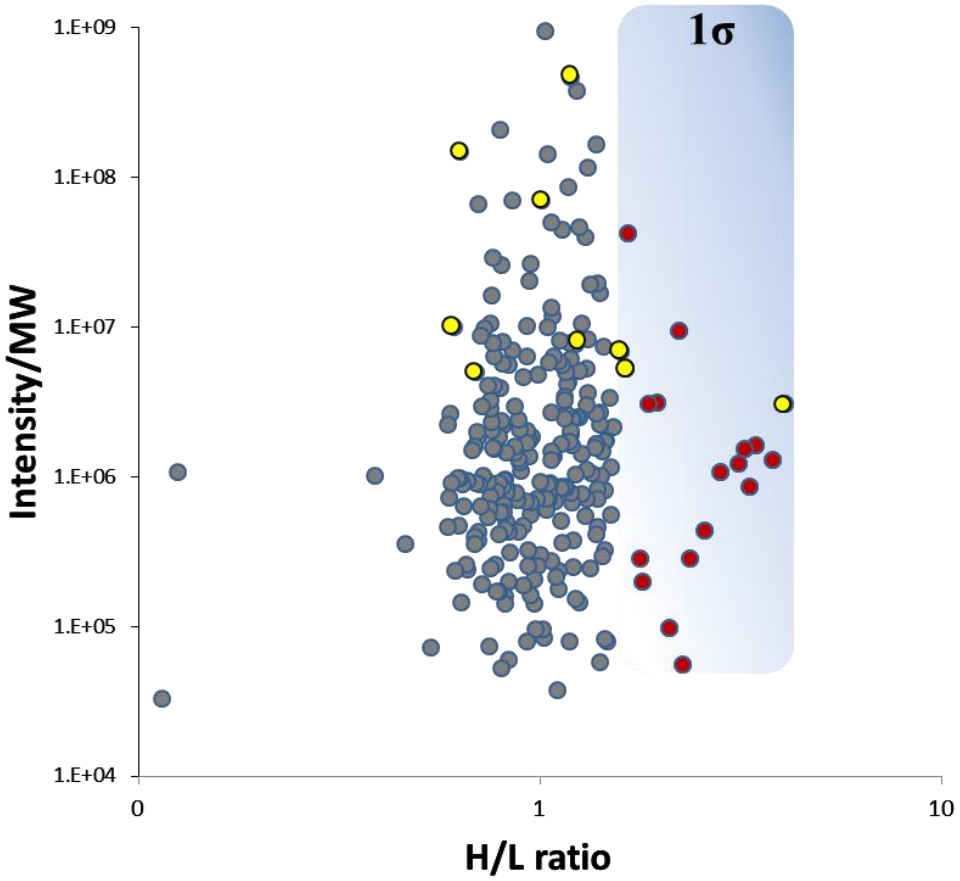


Figure 26. Normalized intensity versus H/L ratio from the I-DIRT experiment
The intensity of each protein passing the filter criteria was normalized for size by dividing by MW and then plotted versus H/L ratio on log scales. Shaded area indicates the region of the graph where H/L ratios are 1 standard deviation greater than the mean (proteins meeting this statistical threshold are shown with red dots). The influenza proteins are represented by yellow dots.

Table 4. Proteins identified in viral I-DIRT immunopurifications and associated data

Gene names	H/L	IP(A) 1	IP(A) 2	IP(B) 3	IP(B) 4	Intensity	MW [kDa]
3xFLAG-NEP*	4.09	5	18	24	33	2.2E+08	73
CCT3	3.81	5	17			7.2E+07	56
HSPD1	3.46	3	26		1	1.0E+08	61
ATP5A1	3.33	1	9			5.2E+07	60
PHGDH	3.24	2	15		1	8.8E+07	57
ATP5B	3.13	3	24			7.0E+07	57
EIF4A1	2.82	1	20			5.0E+07	46
TCP1	2.58		10	1		2.7E+07	60
TUBA1C	2.37		2		1	1.4E+07	50
EPRS	2.27	1	5			9.5E+06	171
EEF1A1	2.23	10	26	6	10	4.8E+08	50
UNC45A	2.10	2	2			9.9E+06	102
HIST1H3A	1.96	2		1		4.8E+07	15
HIST2H2BE	1.87	1	1	1	1	4.3E+07	14
SSRP1	1.80		6		1	1.6E+07	81
LRRC59	1.79		2	1	3	1.0E+07	35
TUBB	1.67	25	121	40	37	2.0E+09	48
PA*	1.64	11	54	6	5	4.4E+08	83
PB2*	1.60	5	92	7	10	5.9E+08	86
XRCC6	1.53	2	31	4	4	1.5E+08	70
XRCC5	1.51	2	19	4	6	9.7E+07	83
RPL37A	1.51			1	2	4.3E+06	8
TUBB4B	1.50	3	11	6	8	1.7E+08	50
NOP2	1.48	1	1		3	7.0E+06	89
RPL7A	1.46	2	7	1	5	5.3E+07	30
TUBB8	1.46	3		2	1	8.6E+07	50
DYNC1I2	1.46	1	2			4.0E+06	12
CDH2	1.45		3		1	8.0E+06	97
RPL3	1.45	1	7	1	2	3.8E+07	46
H2AFV	1.45	2	1	2	4	1.0E+08	14
PABPC1	1.44	1	30	1	5	9.1E+07	61
RPS16	1.43		3		1	4.9E+06	16
RPL6	1.42	3	3	1	2	3.4E+07	33
HIST1H2AA	1.42	3		1	4	2.4E+08	14
RPL8	1.42	3	4	3	4	6.0E+07	22
PCNA	1.42		2			1.7E+06	29
HIST1H2AC	1.42	1	1	2	2	3.1E+07	14
RPL13A	1.40	1	1	1	2	2.0E+07	24

RPS18	1.40	1	3			8.3E+06	18
HIST2H3A	1.40	7	3	10	6	3.0E+08	15
RPL18	1.40		3		1	1.3E+07	18
SET	1.39			1	2	1.3E+07	31
RPL7	1.38	2	5	2	3	4.1E+07	24
RPL4	1.38	1	21	1	5	1.2E+08	45
HIST1H4A	1.38	15	35	33	24	1.9E+09	11
RPL5	1.38	3	5	2	3	4.6E+07	28
RPLP0	1.37	1	6	1	5	4.3E+07	27
RPS24	1.35	1	3	1	2	1.6E+07	15
RPLP2	1.34	6	5	9	10	2.3E+08	12
IRS4	1.34	2	8	2	2	3.3E+07	134
HIST1H2BL	1.32	16	30	22	26	1.6E+09	14
RPL29	1.32	3	4	4	3	6.5E+07	18
PB1*	1.32	13	90	11	13	7.2E+08	87
RPL28	1.32	1	1	1	3	5.7E+06	8
RPS21	1.31	2		2	3	4.7E+07	9
RPL10A	1.31	1	2	1	3	2.8E+07	25
RPL14	1.31	2	3	2	4	4.4E+07	15
RPLP1	1.31	5	1	4	3	4.6E+08	12
RPL9	1.30		2		1	1.1E+07	21
RPS12	1.28	7	3	6	12	1.5E+08	15
RPL35	1.28		3		1	8.1E+06	11
RPS19	1.27	2	2	2	2	2.3E+07	16
SLC25A6	1.26		3	1	1	4.8E+06	33
HIST1H1E	1.26	8	16	19	20	1.0E+09	22
RPS3A	1.26	4	5	3	5	1.2E+08	24
PARP1	1.25	7	32	7	12	2.8E+08	113
HIST2H2AC	1.24	24	19	50	55	5.3E+09	14
RPS8	1.24	1	4	2	4	5.5E+07	22
DOCK7	1.23		12		1	3.6E+07	238
HIST1H1C	1.23	2	3	3	3	1.7E+08	21
RPS2	1.23	1	4	1	3	2.2E+07	21
CAPRIN1	1.22	2	5	1	2	2.9E+07	77
LUC7L2	1.22	1		3	2	1.2E+07	46
RPS5	1.21	2	3	2	5	5.6E+07	23
RPL30	1.21	1	4	4	3	3.1E+07	13
RPL17	1.20	1	1		4	1.4E+07	21
RPL11	1.20	1	2	3	2	3.3E+07	20
RPL12	1.20	8	3	8	10	1.1E+08	18
RPL13	1.20	3	10	4	7	4.7E+07	24

RPL19	1.20	1	2	1	3	1.9E+07	23
NP*	1.19	170	974	226	231	2.6E+10	56
RPL31	1.19	2	2	3	4	2.6E+07	13
RPL23	1.19	2	6	4	6	1.2E+08	15
EEF2	1.19		5	1	1	7.5E+06	95
RPL24	1.19	1	2	2	3	2.5E+07	14
SLC25A5	1.18	4	12	6	11	1.4E+08	33
AMOT	1.18	102	421	111	146	1.0E+10	118
RPL23A	1.17	2	3	4	5	7.4E+07	18
RPS25	1.17	2	2	3	3	3.6E+07	14
DHX9	1.16		31	1	3	1.2E+08	141
RPSA	1.16	3	10	5	5	1.5E+08	30
MYH10	1.16	26	88	2	9	7.9E+08	229
FAU	1.16	1	2	1	2	2.2E+07	7
DDX21	1.16	4	16	6	10	7.5E+07	87
RPS7	1.15	4		2	4	5.4E+07	22
ACTG1	1.14	31	42	47	80	1.9E+09	42
GNB2L1	1.14		6		1	1.3E+07	35
RPS17L	1.14	6	1	4	8	8.8E+07	16
RPS11	1.13	1	1	2	1	9.5E+06	18
RPS13	1.13	1	2	1	3	2.8E+07	17
MATR3	1.13	2	7	2	10	8.1E+07	95
RPS10	1.13	2	1	1	1	1.6E+07	19
RPL38	1.12	2		2	1	6.8E+07	8
LMNA	1.12	1	9		3	1.5E+07	63
AMOTL1	1.12	3	4	2	2	1.8E+07	101
RPS20	1.11	3	5	7	5	7.9E+07	13
SF3B1	1.11		2	1	1	5.5E+06	146
MPDZ	1.10	1	16			4.5E+07	208
SLC25A11	1.09	1	2	1	2	1.9E+07	28
TMOD3	1.09	1	7	1	1	5.4E+07	40
DYNLL1	1.08	4	2	3	7	6.6E+07	10
RPL26	1.08	2	1	2	4	2.5E+07	17
HNRNPC	1.08	8	16	9	12	3.0E+08	25
UBC	1.07	5	17	5	9	1.9E+08	14
ACLY	1.07	1	46	1		1.8E+08	120
RPS6	1.07		4	2		7.8E+06	29
RPS3	1.07	3	7	4	3	7.2E+07	27
HNRNPM	1.07	1	19	4	7	9.7E+07	74
ACTA1	1.07	7	46	10	7	2.1E+09	42
IGF2BP1	1.06		11	1	4	5.8E+07	63

BANF1	1.06	2		5	6	5.9E+07	10
RPS26	1.05		1	2	2	9.3E+06	13
RPS28	1.05			1	3	6.3E+06	8
CSRP2	1.05	14	7	12	16	2.1E+08	21
MPP5	1.05	2	23		1	6.9E+07	77
ACTBL2	1.05	6	25	12	16	6.0E+09	42
MYL12A	1.04	1	3	1	1	1.2E+07	20
HSP90AB1	1.04	1	21		1	5.9E+07	83
ACTB	1.04	144	770	224	313	4.0E+10	42
RPA1	1.02		3	1	1	5.7E+06	68
INADL	1.02		8		1	1.3E+07	130
M1*	1.02	28	63	49	54	2.0E+09	28
ATAD3A	1.01	1	22	2	2	6.2E+07	66
YWHAE	1.00		3		1	8.8E+06	29
ERH	0.99	1	4	4	4	6.0E+07	12
DEK	0.99	1	3		1	1.1E+07	43
SSBP1	0.99	2	1		1	1.0E+07	14
SF3B2	0.98		4	1		9.6E+06	100
ANK3	0.97		8	1	1	2.3E+07	111
ELAVL1	0.97		2		1	5.1E+06	36
NACA	0.96	2	3	1	1	2.8E+07	15
RPS4X	0.95		4	1	1	1.7E+07	30
MYO6	0.95	1	8			2.4E+07	149
YBX1	0.95	1	4	1	2	2.3E+07	34
SPTAN1	0.95	156	451	92	101	7.5E+09	282
MAP4	0.95	4	12	7	6	1.2E+08	91
SPTBN1	0.94	142	308	78	103	5.6E+09	275
TMPO	0.94	3	4	3	7	6.8E+07	39
RBM14	0.94	1	8	1	1	2.2E+07	69
KIAA1671	0.94	3	12		1	5.0E+07	197
MYL6	0.93	4	4	5	5	1.0E+08	16
CPSF6	0.93		2		2	4.2E+06	52
LIMA1	0.93	32	62	34	37	8.7E+08	85
RPL22	0.93	1	2	2	3	3.0E+07	15
PRPF19	0.93	1	3	1	2	4.7E+07	28
SNRPN	0.92		2	1	2	1.6E+07	24
PPP1R9B	0.91	5	8		1	4.2E+07	89
RPL27	0.91		2	2		3.0E+06	16
RBMX	0.91	3	11	5	6	1.5E+08	32
SNRPD3	0.90	1	2	1	1	1.5E+07	13
HNRNPU	0.89	4	15	5	5	1.0E+08	83

PYCR2	0.89	1	5	5	8	7.0E+07	34
SPTANB1	0.89	2	4	2	3	1.1E+08	251
LIN7C	0.89	3	5		2	5.3E+07	22
SRP14	0.89	1		1	2	1.7E+07	13
PPP1CC	0.88	1	3		2	2.4E+07	35
MYH9	0.88	5	32	3	4	1.8E+08	227
AKAP2	0.87	6	22	7	4	1.8E+08	121
ACTN4	0.87	18	45	15	17	3.1E+08	105
HNRNPA3	0.87	2	8	3	4	5.4E+07	34
TPM3	0.86	45	67	35	33	2.0E+09	29
HNRNPK	0.85	6	9	11	17	3.6E+08	51
SYNCRIP	0.85		7	1	1	1.8E+07	59
PLEC	0.84	5	2		1	3.1E+07	513
PYCR1	0.84		4	1	2	1.4E+07	33
STIP1	0.84		6		1	1.2E+07	60
CALD1	0.84	9	39	6	8	3.6E+08	64
YWHAZ	0.83	1	1		2	5.0E+06	6
RPS27	0.83	1	1	2	3	1.4E+07	9
SRSF1	0.83		1	1	4	1.8E+07	22
HMGN2	0.82	2		1	1	4.7E+06	9
SPTBN2	0.82	10	50	3	7	2.6E+08	271
EWSR1	0.82	1	2	3	3	2.8E+07	63
EPB41L2	0.82	1	9			1.8E+07	113
MTHFD1	0.82		7	1	2	1.4E+07	102
SFPQ	0.82	5	7	2	3	7.1E+07	76
RBM33	0.82	2	8	6	3	6.5E+07	100
PCBP2	0.82		1	1	2	9.6E+06	17
FLNA	0.81	52	95	33	48	1.6E+09	280
SERBP1	0.81	6	8	7	5	9.3E+07	42
TPM4	0.81	4	5	3	5	2.5E+07	29
CORO1C	0.81	12	29	13	13	4.3E+08	53
CFL1	0.81	8	8	5	7	4.9E+08	19
CALM1	0.81	1	3	1	1	4.0E+07	17
MYO1C	0.80		3		1	6.3E+06	118
HSPA9	0.80		17	2	3	4.2E+07	72
ALYREF	0.80	3	5	5	7	1.1E+08	27
VIM	0.80	109	254	181	167	1.1E+10	54
TJP1	0.79	19	49	2	6	3.5E+08	187
SNRNP70	0.79	1		1	2	1.8E+07	20
CKAP2	0.79	1	3	1	2	9.5E+06	55
LMO7	0.79	1	20	1	1	6.4E+07	154

UBAP2L	0.78	1	7			1.8E+07	103
CHCHD3	0.78	1	2	2	2	1.2E+07	16
IQGAP1	0.78	1	18			4.9E+07	189
HNRNPD	0.77		4		1	2.0E+07	13
HNRNPA2B1	0.77	8	24	10	6	2.4E+08	37
NCL	0.77	13	25	19	22	3.1E+08	77
PTBP1	0.77		7	3	7	9.0E+07	57
HSPA8	0.77	12	66	11	20	5.6E+08	71
DBN1	0.76	46	75	40	50	2.1E+09	71
FUS	0.76	3	10	3	4	9.1E+07	45
CNN3	0.76	17	21	19	21	5.9E+08	36
HNRNPH1	0.76	3	18	2	4	1.1E+08	47
HNRNPA1	0.76	13	17	10	11	3.1E+08	29
MPRIP	0.76	3	25		2	8.6E+07	117
NONO	0.76	8	18	7	10	1.5E+08	54
TJP2	0.75	1	12		1	3.2E+07	130
FLNB	0.75	36	76	14	32	9.0E+08	276
ILF3	0.75	2	14	2	3	4.6E+07	76
SRSF6	0.75		1		1	2.9E+06	38
HNRNPAB	0.75	2	3	2	2	2.8E+07	30
INA	0.75	2	5	1	4	3.0E+07	55
DDX5	0.74	3	33	7	17	2.8E+08	69
HSPA1A	0.73	17	76	15	26	7.0E+08	70
CAPZA1	0.73	3	7			3.4E+07	33
ABLIM1	0.72		5		1	8.8E+06	46
SRSF3	0.72	3	1	3	2	4.2E+07	14
POLDIP3	0.72	10	46	14	12	4.0E+08	46
NEFM	0.72	5	11	1	2	5.1E+07	79
TPM1	0.71	2	2	1	2	2.6E+07	29
GLUD1	0.71	1	3	1	1	2.3E+07	54
NPM1	0.71	22	31	33	39	2.2E+09	33
DDX3X	0.71	3	13	4	3	1.4E+08	73
HSPA6	0.70	2	3	2	2	2.7E+07	71
PDLIM7	0.70	4	14	2	4	1.0E+08	50
DNAJA1	0.70	1	11	1	2	4.1E+07	45
HA*	0.69	11	31		1	3.2E+08	63
SRSF7	0.69		3	1		6.0E+06	15
IMMT	0.69	2	21	8	11	1.2E+08	73
SRSF2	0.69		1		2	8.3E+06	23
DDX17	0.68	3	15	2	3	1.2E+08	80
KPNA2	0.66	2	5	2	6	5.5E+07	58

CGN	0.66	1	11	1		3.3E+07	136
HNRNPL	0.66		2		4	1.3E+07	51
PA2G4	0.65		14		1	2.6E+07	42
G3BP1	0.64	2	9	1	2	4.7E+07	52
VDAC1	0.64		2		1	4.5E+06	31
NS1*	0.64	36	101	75	79	3.8E+09	26
PPP1R12A	0.63	4	24	3	3	1.1E+08	109
KHDRBS1	0.63	1	3	2	1	2.3E+07	48
CFL2	0.63	1	2		1	1.8E+07	19
DSP	0.62	7	22			7.9E+07	332
M2*	0.61	4	10	6	5	1.1E+08	11
HNRNPH3	0.60			2	1	1.4E+07	15
CORO1B	0.60	4	16	3	3	1.5E+08	54
RAB34	0.60	2	1	1	2	1.5E+07	21
YAP1	0.59	1	7			2.2E+07	48
HSPA5	0.59	3	34	3	8	1.6E+08	72
SUPT16H	0.54	1	2	1	1	8.7E+06	120
SPECC1L	0.46	1	9			4.3E+07	120
DCD	0.39		5	1		1.2E+07	11
S100A9	0.13	1	6			1.4E+07	13
DMD	0.11		2	1	1	8.9E+06	271

Proteins are identified by their official gene symbols**.** Column 2 displays the H/L ratio calculated using MaxQuant. Columns 3-6 show the number times the protein was detected in each sample. Intensity represents the summed signals of all detection events for the protein.

3.4.2 FLAG IP with plasmid expression of 3xFLAG-NEP

The large amounts of other influenza proteins present in the viral I-DIRT immunopurifications confounded the identification of NEP specific interactors so a complementary proteomics dataset in the absence of other influenza components was generated. FLAG immunopurifications were performed following expression of 3xFLAG-NEP or a control (3xFLAG-GFP) in the same cell line used for viral I-DIRT. In this case the mixed buffer IP procedure described in 2.4.2.3 was used to minimize background interactions. Coomassie blue stained gels of the eluates for two replicate immunopurifications are shown in Figure 27. The major band in both control immunopurifications was 3xFLAG-GFP and it ran true to size at approximately 30kD. For the control IP, it appeared that more 3xFLAG-GFP was recovered in the second replicate (Figure 27B) while the inverse was true for the 3xFLAG-NEP IPs. An intense band in the first 3xFLAG-NEP IP (Figure 27A) ran at about 40kD and was likely actin (MW 42kD). While actin is a common IP contaminant, it is not clear why so much was present in this IP as all of the other eluates in Figure 27 have only a faint band at this MW including the second replicate of the 3xFLAG-NEP IP. The 3xFLAG-NEP band just above 14kD was visible in both of the target immunopurifications though it was recovered better in Figure 27B. Qualitatively the 3xFLAG-NEP immunopurifications contained numerous unique proteins compared to the control 3xFLAG-GFP IP. Proteins from the glycine eluates only were prepared for mass spectrometry analysis by both in solution (3.3.4) and in gel (2.3.9.1) digest procedures. Data acquisition and .RAW processing were performed as previously indicated. Replicates for 3xFLAG-NEP or 3xFLAG-GFP were combined in MaxQuant for analysis. Summary data for this proteomics dataset are presented in Table 5.

Western blots of the IP samples are shown in Figure 28. The concentration of 3xFLAG-GFP was much greater than that of 3xFLAG-NEP in their respective lysates. To compensate for the comparatively low expression level, more cells were used for the 3xFLAG-NEP IP. These blots also show that 3xFLAG-NEP was being substantially depleted from the lysates since the target was only faintly observed in the post bind samples (Figure 28A lane 6 and C lane 7).

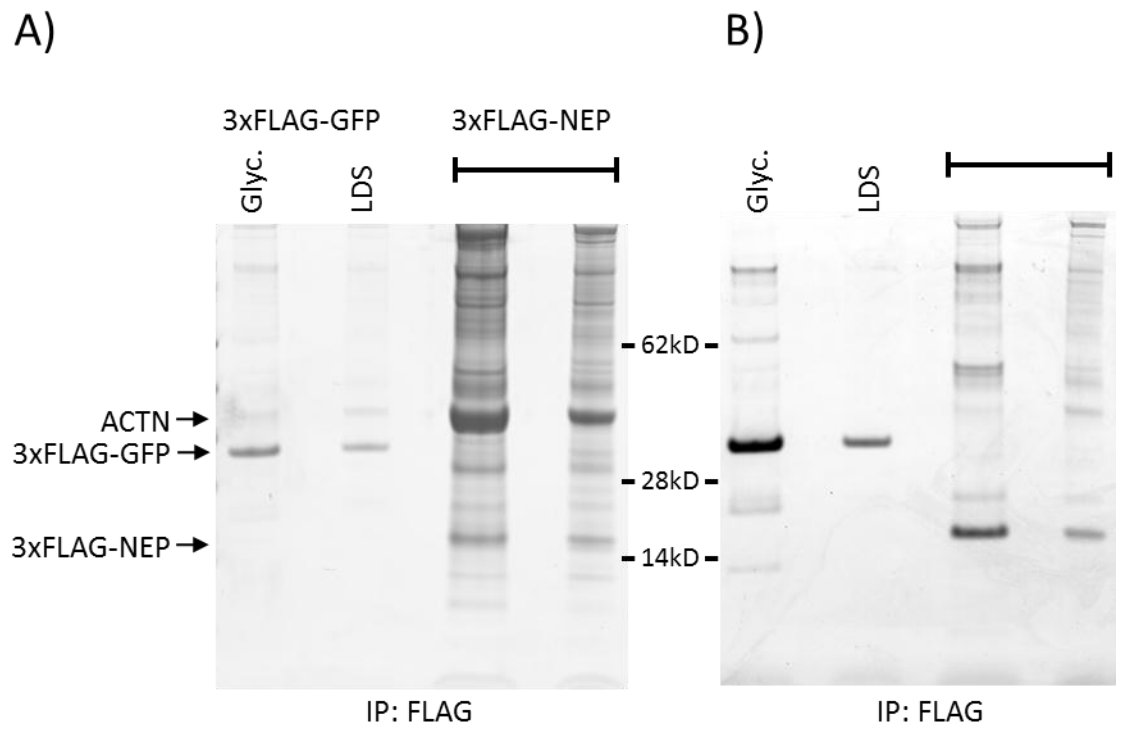


Figure 27. 3xFLAG plasmid expression immunopurifications
HEK293T cells in 15cm plates were transfected with pAc3xFLAG-GFP or pAc3xFLAG-NEP. Cells were washed and frozen 3 days post transfection. Cell lysates were immunopurified for FLAG in HB + 0.5% TX-100 followed by washing with PBS. Interacting proteins were eluted in 200mM glycine pH 3 with heating at 42°C for 10 minutes followed by LDS sample buffer. The same amount of beads were used for all samples; however, the 3xFLAG-NEP IP was performed with cells from three 15cm plates in 4.8mL buffer while the 3xFLAG-GFP IP used cells from one 15cm plate in 1.6mL buffer. Approximately 30% of the eluate was analyzed by SDS-PAGE with Coomassie blue staining for two replicate immunopurifications (A) and (B). The MW of 3xFLAG-NEP is 17.4kD and 3xFLAG-GFP is 30kD.

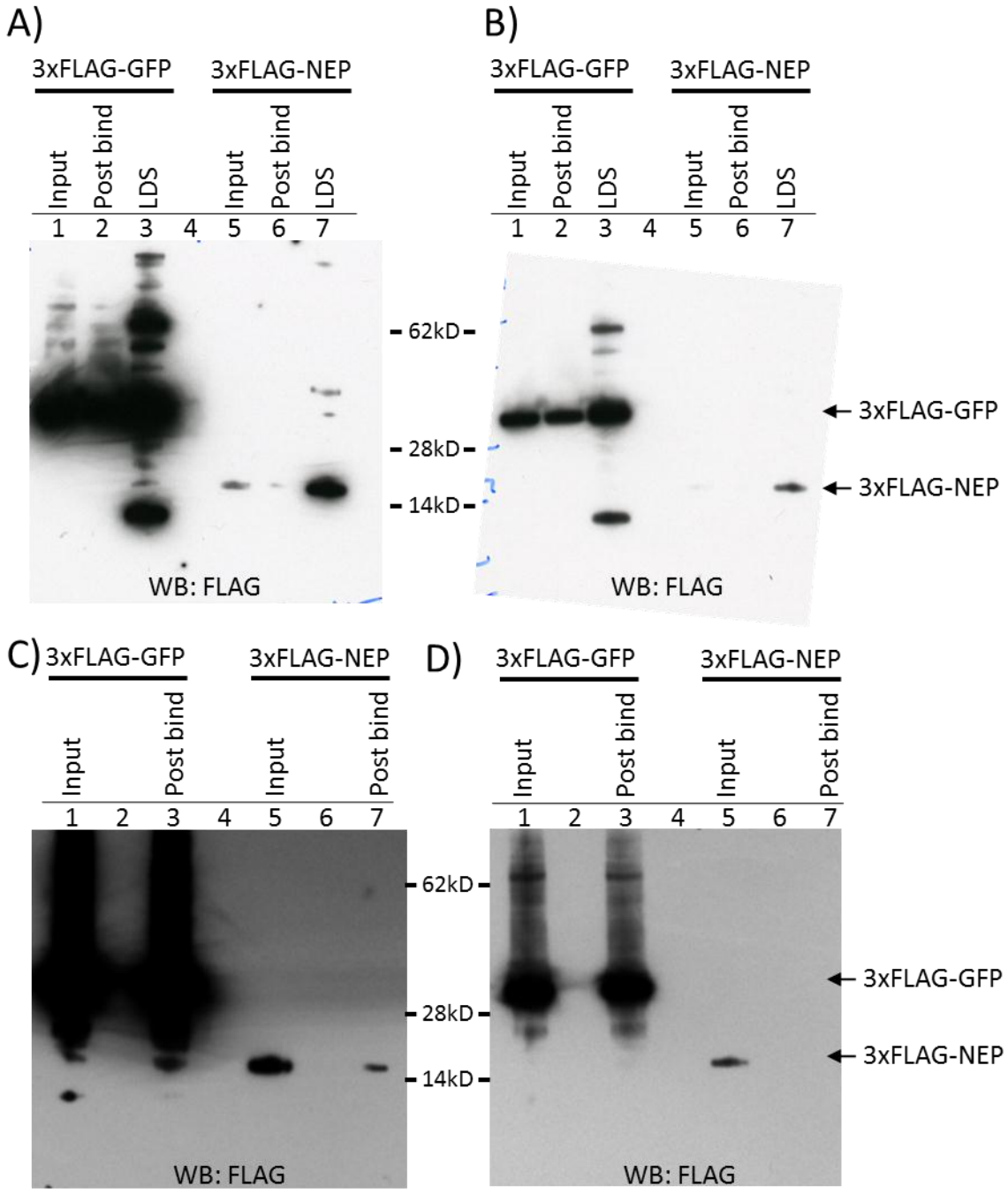


Figure 28. Western blot of FLAG immunopurifications from plasmid expression
Samples from the immunopurifications shown in Figure 27 were analyzed by western blotting for FLAG. (A) and (B) present long and short exposures from a western blot of input, post bind and LDS eluate samples from the IP shown in Figure 27 IP(A). (C) and (D) show long and short western blot exposures from input and post bind samples from the IP shown in Figure 27 IP(B).

Table 5. Summary of mass spectrometry data from immunopurifications of 3xFLAG-NEP or 3xFLAG-GFP from plasmid expression

Sample ID	MS	MS/MS Submitted	MS/MS Identified	MS/MS Identified [%]	Peptide Sequences Identified
IP(A)3xF-GFP	3671	13440	813	6	400
IP(A)3xF-GFP_F1-2	3910	10516	309	3	118
IP(B)3xF-GFP	4054	17826	567	3	147
IP(B)3xF-GFP_F1-3	5378	17596	299	2	71
IP(B)3xF-GFP_F7-9	4226	18473	215	1	46
IP(B)3xF-GFP_F4-7	5691	12380	161	1	48
IP(A)3xF-NEP	2815	13493	2663	20	1462
IP(A)3xF-NEP_F1-4	3206	9794	636	6	328
IP(A)3xF-NEP_F5-7,8	1936	14282	2992	21	1319
IP(A)3xF-NEP_F6	2096	13035	1447	11	497
IP(A)3xF-NEP_F9-11	2150	10831	1473	14	703
IP(A)3xF-NEP_F12-14	2359	9732	1740	18	788
IP(B)3xF-NEP	4033	16471	2073	13	589
IP(B)3xF-NEP_F1-3	5678	13131	79	1	32
IP(B)3xF-NEP_F4-6	4822	14696	334	2	115
IP(B)3xF-NEP_F7-9	4875	11404	104	1	49
Average	3806	13569	994	8	420
Standard dev.	1241	2750	922	7	440
Total	**60900**	**217100**	**15905**	**8**	**6712**

The sample ID corresponds to the labels in Figure 27. Column 2 indicates the total number of MS1 spectra. Column 3 shows the number of MS2 spectra submitted to the Andromeda search engine.

A total of 126 proteins were identified in the 3xFLAG-GFP immunopurifications and 634 in the 3xFLAG-NEP immunopurifications. In both datasets, FPRs were less than 4%. Filtering with the same criteria used for the I-DIRT experiment gave working sets of 13 and 98 proteins respectively. These lists are shown in Table 6 and are ranked by the summed intensity of all peptide spectra. This dataset had lower percentages of identified MS2 spectra especially for the control. The total intensities in the control were approximately 2 orders of magnitude lower than for the 3xFLAG-NEP dataset.

Table 6. Proteins identified in FLAG immunopurifications from plasmid expression

3xFLAG-NEP			
Gene symbol	IP(A)	IP(B)	Intensity
ACTG1	548	44	3.41E+09
VIM	252	122	1.06E+09
ACTA1	63	3	4.75E+08
TUBA1B	160	16	4.13E+08
AMOT	257	72	4.03E+08
SPTAN1	454	67	2.91E+08
TUBB4B	143	16	2.79E+08
CORO1C	80	5	2.03E+08
SPTBN1	333	46	2.01E+08
TUBB	54	6	1.61E+08
ATP5B	78	88	1.45E+08
3xFLAG-NEP*	23	89	1.34E+08
ATP5A1	42	55	9.43E+07
HIST1H2BL	16	8	4.38E+07
ACTB	15	4	4.33E+07
MYH10	151	4	4.23E+07
LIMA1	81	3	4.03E+07
FLNB	123	2	3.85E+07
NPM1	51	1	3.47E+07
UBB	25	1	3.36E+07
FLNA	96	2	2.98E+07
TRIM21	20	24	2.47E+07
HIST1H1C	17	10	2.40E+07
HIST1H2AJ	10	8	2.36E+07
HNRNPA2B1	31	1	2.35E+07
HIST1H4A	17	14	2.18E+07
HSPA8	39	12	1.60E+07
SLC25A5	20	4	1.53E+07
INA	27	6	1.53E+07
MYH9	85	1	1.50E+07
ABLIM1	32	1	1.32E+07
MPP5	37	4	1.17E+07
HNRNPC	17	2	1.16E+07
TJP1	43	5	1.15E+07
RBMX	14	1	1.12E+07
RPL7A	10	2	1.10E+07
EEF1A1	12	5	1.09E+07
CFL1	16	2	9.58E+06
POLDIP3	6	18	9.10E+06

3xFLAG-GFP			
Gene symbol	IP(A)	IP(B)	Intensity
IGK	2	11	1.40E+07
VIM	29	1	7.52E+06
POLDIP3	6	14	7.13E+06
UBB	3	2	1.90E+06
HIST1H2BL	4	1	1.30E+06
RPL13	3	3	1.29E+06
TRIM21	2	5	1.23E+06
HIST1H1C	5	3	1.06E+06
HIST1H2AJ	2	2	1.03E+06
RPL23	4	1	5.43E+05
GTSE1	1	2	4.69E+05
TMPPE	1	1	2.45E+05

SPTBN2	52	3	8.21E+06
MPDZ	36	2	6.94E+06
HSPA1A	30	4	6.81E+06
DYNLL1	12	1	6.36E+06
INADL	36	2	6.31E+06
SFPQ	19	1	6.03E+06
AKAP2	27	3	5.96E+06
PHGDH	12	1	5.84E+06
MPRIP	33	1	5.44E+06
H3F3A	5	3	5.13E+06
LIN7C	11	2	5.00E+06
NCL	20	1	4.63E+06
ATP5O	9	12	4.60E+06
TPM1	6	1	4.44E+06
SPTBN1	4	1	4.25E+06
MYL12A	13	1	3.76E+06
IMMT	17	1	3.68E+06
RPL6	9	1	3.55E+06
LMO7	25	3	3.48E+06
NEFM	12	8	3.14E+06
PLEC	28	6	3.10E+06
RPL7	6	1	2.90E+06
RPL13	11	5	2.81E+06
CHCHD3	10	3	2.62E+06
YAP1	11	2	2.60E+06
AMOTL1	18	4	2.47E+06
ATP5J	1	6	2.38E+06
ATP5F1	2	9	2.36E+06
RPL23	6	1	2.23E+06
CHMP4B	4	1	2.16E+06
SLC25A11	6	1	2.11E+06
RPL12	9	2	1.99E+06
SPTAN1	5	1	1.90E+06
ATP5H	1	6	1.80E+06
ALYREF	4	1	1.76E+06
RBM14	11	2	1.72E+06
DSP	16	1	1.64E+06
RPL23A	5	1	1.50E+06
SLC25A3	4	1	1.38E+06
MAP4	1	5	1.11E+06
HIST1H1E	1	1	9.21E+05
ATP5C1	1	4	8.56E+05
HSPA6	1	2	8.30E+05

HSP90AB1	5	1	8.00E+05
ATP5I		2	7.94E+05
RPL18	5	1	7.74E+05
CHTOP	3	3	7.65E+05
HIST2H2BE	1	1	7.07E+05
RPS8	7	1	6.89E+05
RPL11	4	1	5.78E+05
ERH	2	2	5.31E+05
RPLP2	3	1	5.04E+05
CKAP2	1	4	4.89E+05
RPL29	2	1	4.75E+05
RPL19	5	1	3.88E+05
IGF2BP1	5	1	3.06E+05
RPS11	1	1	2.29E+05

Proteins are identified by their official gene symbols. IP(A) and IP(B) columns display the number of times a given protein was detected in the sample. Intensity represents the summed spectral signals of all detection events for the protein.

3.4.3 Gene function classification with DAVID shows an enrichment of proteins involved in cellular processes important for the viral lifecycle

The protein identifications from the viral I-DIRT IP (Table 4) were uploaded to the DAVID bioinformatics platform. Using the default settings, the gene function classification tool returned 13 functionally enriched groups of proteins displayed in Table 7. The major functional or structural attribute of the grouping is given in italics along with the relative enrichment *p* value. In two cases, the defining feature of the grouping was a protein domain (either LIM or WD40). These functional groups were incorporated into a network analysis (3.4.4).

Table 7. Enriched functional groups in viral I-DIRT

Translation/ribosome p = 7.1E-67		*Actin-binding p = 5.8E-33*	*RNA Processing p = 1.5E-30*	*Chromatin Organization p = 3.1E-16*
RPL23A	RPS20	TPM3	RBM3	HIST2H2Ac
RPL13	RPL14	MPRIP	SF3B2	HIST1H2AB
RPS28	RPS7	ABLIM1	SYNCRIP	H3F3A
RPL8	RPL19	TMOD3	HNRNPU	HMGN2
RPS25	RPLP1	CFL2	SFRS6	H2AFZ
RPL12	RPS16	TPM4	HNRNPAB	HIST1H1C
RPS26	RPL5	*WD40 Domain p = 6.2E-14*	SFRS3	HIST1H2BL
RPS8	RPL38		HNRNPA2B1	HIST2H2BE
RPL18	RPS27		POLDIP3	H2AFY
RPL23	RPS11	DYNC1I2	HNRNPK	HIST1H1E
RPS24	RPL27	CORO1C	PTBP1	HIST2H3A
RPS23	RPL27A	CORO1B	RBM4, RBM14	HIST1H4A
RPL22	FAU	GNB2L1	HNRNPH3	*Molecular Chaperone p = 2.0E-11*
RPL6	RPS18	*Intermediate Filaments p = 1.2E-11*	NCL	
RPS12	RPL28		SFPQ	
RPS5	RPLP0		PABPC1	
RPS2	RPL26		IGF2BP1	
RPS10	*Microtubules p = 3.6E-12*		SNRNP70	HSP90AB1
RPL29		NEFM	PCBP2	ATAD3A
RPS21		VIM	SFRS7	TCP1
RPL31	TUBB	LMNA	HNRNPM	ACTBL2
RPL30	TUBB8	INA	SNRPD2	HSPA9
RPL11	TUBA1C	*Cytoskeleton p = 1.5E-10*	HNRNPD	hspa1b
RPL4	TCP1		RBMX	CCT3
RPL37A	TUBB2c		SNRPD3	HSPA5
RPS6	*LIM Domain p = 9.4E-11*		SFRS2	HSPA6
RPL35		SPTBN1	FUS	HSPA8
RPL17		SPTBN2	HNRNPH1	*Helicase p = 5.5E-10*
RPL7		SPTAN1	MATR3	
RPS17	ABLIM1	CAPZA1	SFRS1	
RPL3	PDLIM7	LIMA1	CPSF6	
RPL10A	LIMA1	*Tight junctions p = 2.4E-6*	NONO	ATAD3A
RPS13	CSRP2		ELAVL1	DDX3X
RPLP2	*Mitochondrial membrane p = 1.7E-7*		SNRPD1	HNRNPU
RPSA			HNRNPL	DDX17
RPL13A		LIN7C	SF3B1	DDX5
RPS4X		INADL	HNRNPA1	DHX9
RPS3		TJP1	YBX1	EIF4A1
RPS3A	CHCHD3	TJP2	Thoc4	G3BP1
RPL9	SLC25A5	CGN	HNRNPA0	DDX21
RPL7A	IMMT	AMOTL1	HNRNPC	
RPS19	SLC25A6	MPP5	HNRNPA3	
RPL24	SLC25A11	MPDZ	SNRPN	

3.4.4 Network Analysis in Cytoscape

The two IP datasets, one from viral I-DIRT and one from plasmid expression of 3xFLAG-NEP, shared 80 proteins which represent the NEP interactome (Table 8). The NEP interactome was queried against the CRAPome contaminant repository and the percent abundance value, displayed in Table 8, is the number of experiments in which the protein was observed divided by the total number of experiments in the database (411). Gorai et al. also used FLAG tag-based IP-MS to identify 37 host interactors of NEP and 15 of these proteins were present in our NEP interactome (denoted with * in Table 8). Venn diagrams in Figure 29 depict the number of overlapping proteins between the different datasets. The significance of the overlap was calculated using hypergeometric probability ($p < 2.6E-26$) from a background of 17,000 possible proteins (the draft human proteome contains 17,294 proteins (Kim et al. 2014) so the total was rounded down conservatively). The significance of the overlap was robust to the size of the background and only fell below 0.05 when the total possible protein pool was less than 290. This result demonstrated congruence with published data.

The two proteomics datasets were combined in Cytoscape to create a network model of the NEP interactome (Figure 30). Each node, other than the two central nodes which represent 3xFLAG-NEP from the plasmid (left) and viral IPs (right), represents a protein identified in immunopurifications from virally infected cells. The 80 proteins shared between the two datasets are colored yellow and represent the NEP interactome. Blue edges show interactions that have been previously reported between any influenza protein and the host factor curated from the literature (J. Chen, Huang, and Chen 2010; Demirov et al. 2012; Deng et al. 2006; Han et al. 2012; Jorba et al. 2008; Liu et al. 2009; Mahmoudian et al. 2009; Mayer et al. 2007; ONeill, Talon, and Palese 1998; Satterly et al. 2007; Tawaratsumida et al. 2014; Tripathi et al. 2013; Watanabe et al. 2006; Watanabe et al. 2014). Red edges represent interactions unique to this dataset. Grey edges are intra-host interactions derived from the STRING database. Edge width is proportional to the H/L ratio for the virally derived interactions with thicker edges indicating a higher ratio and thus a more stable interaction. Triangular nodes show those proteins that have been identified as functionally important for the influenza virus by at least one of 6 genome-wide screens (Brass et al. 2009; Hao et al. 2008; Karlas et al. 2010; König et al. 2010;

Shapira et al. 2009; Sui et al. 2009). The node size indicates the relative abundance of the protein in the CRAPome. Larger nodes are less prevalent in the CRAPome and are less likely to be background contaminants. Proteins are clustered based on the DAVID analysis (Table 7). Nodes from the two structural clusters (LIM and WD40 domain) were moved to the unclustered group to simplify the visualization and to focus on the functional enrichment in the dataset.

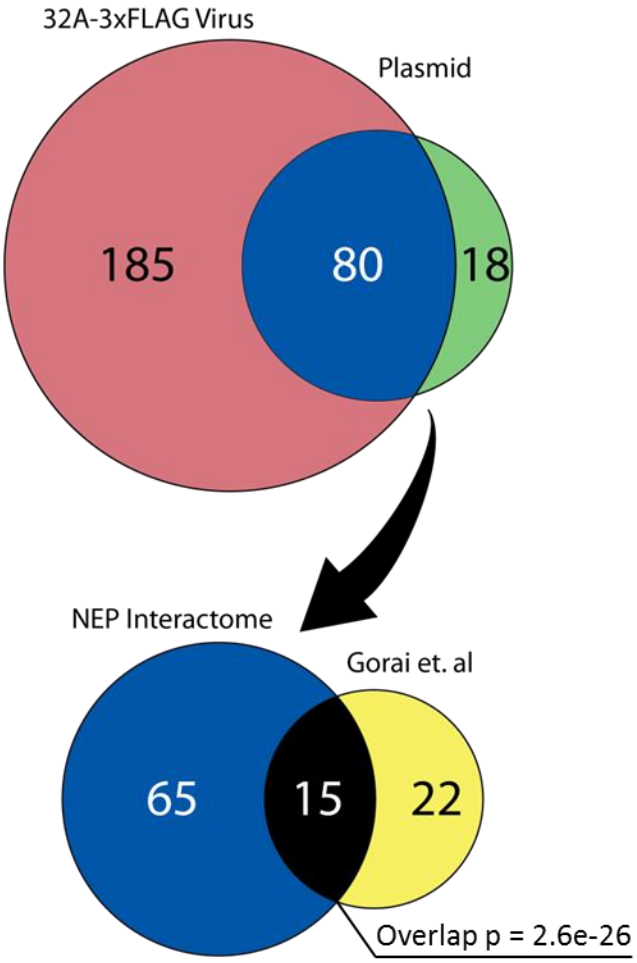


Figure 29. Venn diagram showing overlap of different NEP interaction datasets
The viral I-DIRT protein set is represented in pink and the set from 3xFLAG-NEP plasmid expression is green. The NEP interactome is shown in blue and interactors reported by Gorai et al. are represented by yellow. The intersection of these two is shown in black. The significance of this overlap is indicated.

The graph features 3 highly populated groupings (RNA processing, translation, and unclustered proteins) and 9 smaller clusters. The 97 proteins from the viral IP that did not get placed into any of the groups by the gene functional classification tool from DAVID are located in the unclustered group. The two highest scoring functional enrichment clusters are those related to translation and RNA processing, and many of these interactions have already been reported (as shown by the predominance of blue edges) indicating congruence with existing data. These two clusters were also dominated by blue nodes indicating little overlap with the proteins identified by plasmid expression of 3xFLAG-NEP and suggesting that many of these interactions are dependent on the presence of the other influenza proteins. There are also a number of triangular nodes representing pro-viral factors in these two clusters, especially in RNA processing, indicating that this functional group is highly important for the viral lifecycle.

After visually assessing the 9 smaller clusters in a similar fashion, one grouping in particular stands out. The tight junction cluster, enlarged in Figure 31, consists of mostly yellow nodes indicating the interactors were identified in 3xFLAG-NEP IPs in both the presence and absence of other influenza components. The red edges in the cluster indicate that these host factors have not previously been identified as influenza interactors. Additionally one of the tight junction components, AMOTL1, has been implicated as functionally important for viral replication by siRNA screening (Brass et al. 2009). We used the CRAPome database to assess the likelihood these 8 factors were contaminants. In Table 8, the NEP interactome proteins were ranked according to their representation in the database and the five proteins with lowest representation were all members of the tight junction cluster. A one-tailed t-test between the number experiments in which all proteins from the database appeared (excluding the eight tight junction proteins) and the number of experiments in which the 8 tight junction proteins were detected, yielded a *p* value of 0.023. This result showed that the tight junction cluster had a statistically significant low representation in the contaminant repository.

In summary, proteins with the tight junction functional annotation were significantly enriched in the viral IP-MS dataset and were also identified as NEP interactors in the absence of other influenza components. The tight junction components identified in this work have not

previously been reported as directly interacting with influenza proteins, though one has been identified as a pro-viral host factor by siRNA screening. The statistical analysis discussed above indicated that they were not likely to be IP contaminants. These results suggested that NEP could be mediating functionally important interactions with tight junction components during infection.

Table 8. NEP interactome: proteins shared between viral I-DIRT and plasmid expression proteomics datasets

Gene Symbol	CRAPome	Total I-DIRT	Total plasmid	H/L	Gene Symbol	CRAPome	Total I-DIRT	Total plasmid	H/L
AMOTL1	1%	11	22	1.1	RPL6*	42%	9	10	1.4
LIN7C	1%	10	13	0.9	RPL7*	43%	12	7	1.4
MPP5	1%	26	41	1.0	ALYREF	43%	20	5	0.8
INADL	1%	9	38	1.0	HNRNPC	44%	45	19	1.1
MPDZ	1%	17	38	1.1	RPL18*	46%	4	6	1.4
YAP1	4%	8	13	0.6	CFL1	48%	28	18	0.8
CHCHD3	4%	7	13	0.8	RPL7A	48%	15	12	1.5
AMOT	5%	780	329	1.2	RPL12*	49%	29	11	1.2
ABLIM1	7%	6	33	0.7	SFPQ	49%	17	20	0.8
AKAP2	8%	39	30	0.9	MYH9	49%	44	86	0.9
LMO7	8%	23	28	0.8	RPL11*	51%	8	5	1.2
MPRIP	9%	30	34	0.8	ATP5A1*	51%	10	97	3.3
TJP1	9%	76	48	0.8	RPL19*	51%	7	6	1.2
CKAP2	10%	7	5	0.8	ATP5B*	51%	27	166	3.1
IMMT	11%	42	18	0.7	RPLP2*	52%	30	4	1.3
SPTBN2	11%	70	55	0.8	RPL13	53%	24	16	1.2
RBM14	13%	11	13	0.9	SLC25A5	54%	33	24	1.2
SLC25A11	13%	6	7	1.1	RPL23A	55%	14	6	1.2
DYNLL1	16%	16	13	1.1	HIST1H4A	55%	107	31	1.4
LIMA1	18%	165	84	0.9	FLNA	55%	228	98	0.8
POLDIP3	19%	82	24	0.7	HIST1H2BL	55%	94	24	1.3
INA	20%	12	33	0.7	HIST2H2BE	55%	4	2	1.9
NEFM	21%	19	20	0.7	RPS8	55%	11	8	1.2
MAP4	21%	29	6	0.9	RPL23	59%	18	7	1.2
CORO1C	24%	67	85	0.8	NPM1	61%	125	52	0.7
PLEC	27%	8	34	0.8	VIM*	63%	711	374	0.8
TPM1	27%	7	7	0.7	NCL*	64%	79	21	0.8
IGF2BP1	28%	16	6	1.1	HNRNPA2B1	64%	48	32	0.8
DSP	31%	29	17	0.6	HSP90AB1	67%	23	6	1.0
TPM3	33%	180	83	0.9	HIST1H1E	73%	63	2	1.3
SPTBN1	35%	631	379	0.9	HIST1H1C*	73%	11	27	1.2
MYL12A	35%	6	14	1.0	EEF1A1	85%	52	17	2.2
ERH*	36%	13	4	1.0	ACTA1	86%	70	66	1.1
RPS11	37%	5	2	1.1	ACTG1	88%	200	592	1.1
PHGDH	37%	18	13	3.2	ACTB*	88%	1451	19	1.0
SPTAN1	39%	800	521	0.9	TUBB4B	91%	28	159	1.5
MYH10	40%	125	155	1.2	HSPA6	92%	9	3	0.7
RBMX	41%	25	15	0.9	TUBB	93%	223	60	1.7
RPL29	41%	14	3	1.3	HSPA1A	96%	134	34	0.7
FLNB	42%	158	125	0.8	HSPA8*	96%	109	51	0.8

Proteins are ranked by % representation in the CRAPome. The total detections from each dataset as well as the H/L ratio are presented in columns 3-5. Members (6 of 8) of the tight junction functional cluster are in bold. * denotes NEP interactors from Gorai et al. 2012.

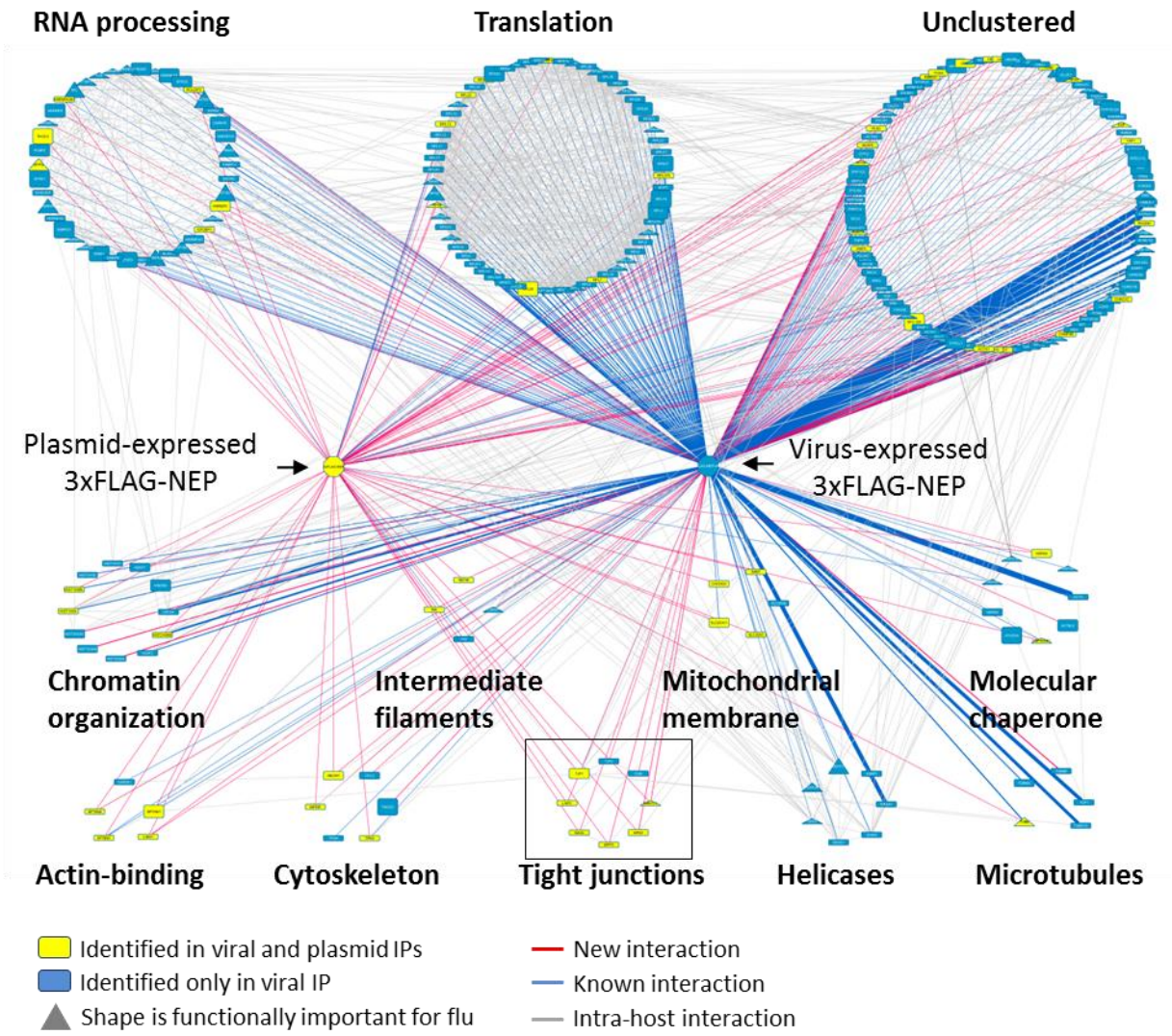


Figure 30. Network representation of IP-MS datasets for 3xFLAG-NEP virus and 3xFLAG-NEP plasmid immunopurifications

Nodes represent proteins identified in the viral I-DIRT IP. Yellow nodes are proteins also identified in the IP from plasmid expression of 3xFLAG-NEP and represent the NEP interactome. Node size is indicative of abundance in an IP contaminant database. Larger nodes are less represented in the database. Triangular nodes have been implicated as functionally important for influenza. Red edges represent new host-virus interactions between 3xFLAG-NEP from plasmid expression (left, central hexagonal node) or from viral expression (right, central hexagonal node) and human proteins. Blue edges represent published interactions. Edge width is proportional to the H/L peptide ratio (thicker edges indicate more stable interaction). Grey edges show intra-host interactions from the STRING database. Node clustering was derived from gene function classification in DAVID. The tight junction functional cluster (black box) is shown enlarged in Figure 31.

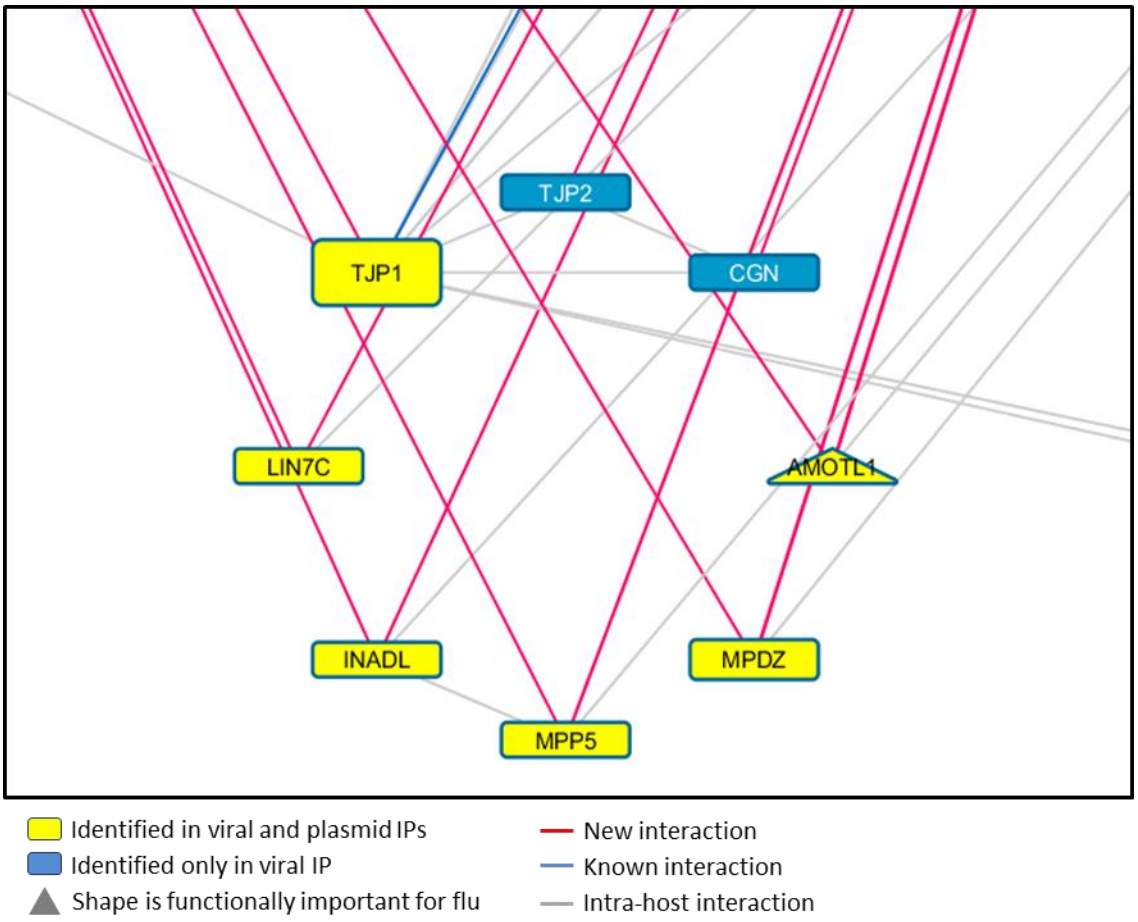


Figure 31. Expanded view of the tight junction functional cluster
Shown is an enlargement of the tight junction functional cluster inside the black box in Figure 30. Six of the eight nodes in the tight junction cluster are yellow indicating they were identified as NEP interactors during infection and also when NEP was expressed alone from a plasmid. The connecting edges are mostly red indicating these host proteins have not previously been identified as influenza interactors. AMOTL1 has been identified as pro-viral by genome-wide siRNA screening.

3.5 Discussion

The recombinant 32A-3xFLAG virus developed in Chapter 2 was employed in IP-MS experiments to discover new host factors interacting with NEP. We observed an enrichment of tight junction proteins in the proteomics data. Other groups have performed IP-MS experiments on NEP but have not reported an association with tight junction components so this result is unique to our approach. Reasons for this are probably due to the particular conditions utilized for the affinity purifications. We used a rapid, 30-minute process that was designed to capture transient interactors while past studies used extended 18-hour procedures (Gorai et al. 2012). Additionally, the buffers used for lysis, immunopurification, and washing were different between the studies and therefore it is not surprising that the set of identified interactors are also different (LaCava et al. 2015). Overall the mass spectrometry data were of comparable quality to published datasets as assessed by Andromeda peptide identification scores, percentage of tandem mass spectrometry identified, and protein FDR. I-DIRT provides a quantitative assessment of the stability of a bait-prey interaction in the form of a heavy to light peptide ratio. However, when we analyzed the H/L ratios of the proteins identified in the viral IP-MS experiment we did not find a convincing enrichment of specific and stable host interactors. A total of 19 proteins had H/L ratios exceeding the statistical threshold of 1 standard deviation from the mean. In comparison, 9 proteins had ratios less than 1 standard deviation from the mean and these represented false negatives. Since the heavy and light proteins were mixed in equal ratios prior to the IP, theoretically H/L ratios less than 1 should not have been observed. We therefore attributed the false negatives to error in the calculation of the ratios from the mass spectrometry data. If we assume the number of false positives and false negatives are equal, this equates to an FPR of 47% (9/19). In addition, many of the proteins with high H/L ratios are common IP contaminants.

It is likely that multiple factors are contributing to the lack of differentiation between stable interactors and background. One important consideration is that 3xFLAG-NEP was not the most abundant protein in the IP. SDS-PAGE analysis as well as mass spectrometry intensity data indicated that many host and influenza components were present in excess over NEP. As a

result, the signal from host factors directly interacting with 3xFLAG-NEP may be rather low. Another factor could be that under the IP conditions used, interactions with NEP may have been unstable or more dynamic. A more robust examination of alternative buffers, detergents, and salts may identify conditions better suited for the recovery of transient interactions.

There are a few experimental alterations that could improve detection of stable and dynamic interactors. One is in the mixing step. In these experiments, mixing occurred after lysis and this step should be done prior to lysis. The physio-chemical protein environment is drastically altered during cell breakage leading to non-specific protein-protein interactions. As a result contaminants may aggregate onto the target even before mixing. This could be one reason why we observed common contaminants with high H/L ratios. Another has to do with the immunopurification timeline. I-DIRT experiments must balance IP speed with the amount of target recovered and we used a 10 minute binding time. Affinity capture experiments using components interacting with sub-nanomolar K_Ds may enable shorter binding periods and increase the H/L ratios of transient interactors. Also performing the I-DIRT experiment with multiple capture times (i.e. 5, 10 and 20 minutes) may enable the detection of more weakly interacting proteins since we would expect their ratios to fall with time while ratios for stable and background interactors should remain relatively unchanged. Lastly, purification conditions could be altered to improve the detection of specifically interacting proteins. For the I-DIRT IP, we wanted to maximize the recovery of interacting proteins and therefore opted to wash with HB. However, previous results indicated that washing with PBS lowered background (Figure 18) and this may be why more actin was observed in the I-DIRT IPs (Figure 25). Although the I-DIRT methodology does not use protein abundance to identify contaminants, it is still subject to signal-to-noise constraints and having a high overall background will impact the ability to detect enriched H/L ratios. Competitive elution with the 3xFLAG peptide is another way we could have increased the signal from NEP specific interactors. We did not test this elution condition because 3xFLAG-NEP was successfully liberated from M2 with low pH. Nonetheless, a 3xFLAG peptide elution may have been preferable to the sequential high salt, low pH elutions that were used.

Although the quantitative aspect of the I-DIRT IP did not yield a compelling list of NEP-specific interactors, the dataset contains a large number of host-influenza interactions derived from an active infection. In order to determine the subset of host factors interacting directly with NEP, we performed an IP-MS experiment without the other influenza proteins present by expressing 3xFLAG-NEP from a plasmid. We used 3xFLAG-GFP to control for background and identified 80 host proteins in the 3xFLAG-NEP dataset that were also identified in the viral IP-MS experiment. These 80 proteins represent the NEP interactome. Other researchers recently identified host interactors of NEP, also using FLAG IP-MS (Gorai et al. 2012), and 15 of the 37 proteins in that dataset were also present in our NEP interactome: a statistically significant overlap. This indicates that our NEP interactome is representative of published results.

Ideally large high throughput datasets will accurately represent the underlying biological processes upon which their collection was based. In our case, we performed IP-MS experiments that isolated an influenza protein from infected cells and identified many interacting host and viral components. Therefore we expected that proteins from cellular processes functionally important for the viral lifecycle would be enriched in our dataset. The gene functional classification tool from DAVID identified 13 groups of functionally related proteins that were statistically enriched in our viral IP-MS protein list, and the processes associated with most of these clusters have in fact been previously identified as important for influenza replication.

The polymerase complex and vRNPs have previously been shown to associate with ribosomal and spliceosomal components (Jorba et al. 2008; Mayer et al. 2007; Watanabe et al. 2014), and the two most significantly enriched functional annotation groups were those related to translation and RNA processing. The large number of triangular nodes in these clusters highlights the central role these processes play in viral replication (Karlas et al. 2010; T. Watanabe, Watanabe, and Kawaoka 2010). The presence of chromatin organizing proteins was explained by the fact that vRNPs tightly associate with cellular chromatin (Takizawa et al. 2006) to gain access to host export factors (Chase et al. 2011), and that influenza proteins associate with histones (Garcia-Robles et al. 2005; Zhirnov and Klenk 1997). Four different clusters related to cellular structural elements reflecting known associations with cytoskeletal

components (Avalos, Yu, and Nayak 1997; Zhang and Lamb 1996). Structural elements are important for viral protein translation (Arcangeletti, Pinardi, and Missorini 1997), and assembly and budding of new virions (König et al. 2010; Simpson-Holley et al. 2002). NEP associates with mitochondrial adenosine triphosphate (ATP) synthase components (ATP5A/B), so the proteins in the mitochondrial membrane cluster may have represented bridged interactions. Nodes and connecting edges in the helicase and molecular chaperone clusters were nearly all blue indicating that these were mostly unique to the viral IP and previously observed. Helicases in particular have been shown to enhance activity of the viral polymerase complex (Bortz et al. 2011), again demonstrating congruence between processes import for viral replication and functional enrichment in the dataset. Two of the clusters represented protein domain enrichments (LIM and WD40). Both domains mediate protein-protein interactions and are involved in many different cellular processes (Bach 2000; Stirnimann et al. 2010). These clusters had only 7 unique members (LIMA1 was also grouped in the cytoskeletal cluster), and since we focused on functional enrichment these clusters were not included in the network visualization.

The final cluster represented an enrichment of proteins associated with cellular tight junctions. Six of the eight proteins in this group were identified in both IP-MS datasets and had not been previously known to associate with influenza before. This raised the possibility that NEP may be physically associating with tight junction components during infection. We used the CRAPome to assess the likely hood that proteins from the tight junction cluster were contaminants and found that of all proteins in the NEP interactome these had the lowest representation in the database. The abundance of the tight junction proteins was also statistically low in the database suggesting that the functional enrichment observed was not due to background association. Tight junctions are important cellular structures that establish cellular polarity and physically isolate the apical and basolateral membranes and are targeted by numerous viral and bacterial pathogens. Based on these results, we hypothesized that NEP may be modulating tight junctions during the viral lifecycle and the ensuing chapter presents results from experiments testing whether or not NEP affects tight junctions.

Chapter 4: NEP antagonizes tight junction function

4.1 Summary

Network analysis of the NEP proteomics data revealed an enrichment of tight junction components in the NEP interactome. We assessed the effect of NEP on tight functioning by measuring transepithelial electrical resistance and macromolecular diffusion across polarized cell monolayers. These assays indicate that NEP does impair the barrier function of tight junctions. We also observed mislocalization of the tight junction protein occludin in the presence of NEP by immunofluorescence. Tight junctions isolate the basolateral cell surface from the external environment and enable the establishment of cellular polarity. These results implicate NEP in the modulation of an important cellular system. How these findings relate to the viral lifecycle is discussed in Chapter 5.

4.2 Introduction

Specialized intercellular junctions were first observed in the 1950's when the lens of the recently invented electron microscope was turned to the study of epithelial tissues (Fawcett 1958; Porter 1956).These early images implicated the junctional structure in cell-cell adhesion (Hama 1960) and epithelial permeability (Peachey and Rasmussen 1961). Work by Marilyn Farquhar and George Palade in the early 1960's demonstrated the tripartite nature of the junctional complex and established that the barrier function was mediated by the most apical structure—the *zonula occludens* or tight junction (Farquhar and Palade 1963). A short time later images of tight junctions reformed following invasion by *Salmonella typhimurium* hinted at the dynamic character of the structure (Takeuchi 1967), a key property that was subsequently confirmed (Galli et al. 1976; Martinez-Palomo et al. 1980; Staehelin 1973) and is still under

study today (Capaldo et al. 2014; Sasaki et al. 2003). In 1972, Eberhard Fromter and Jared Diamond showed that tight junctions regulate ion flux (Fromter and Diamond 1972) and Diamond subsequently proposed that they serve as a selective gate to the paracellular flux of ions and solutes. Around this same time, studies on the compositions and properties of the apical and basolateral surfaces of epithelial cells revealed distinct differences (Fujita et al. 1973; Lewis et al. 1975; Lewis and Diamond 1976) and this led Diamond to propose an additional "fence" function for tight junctions in the maintenance of cellular polarity (Diamond 1977).

The physiochemical functions of tight junctions, gate and fence, have been established for many years, but we now know that these structures are also involved in regulating diverse processes inside the cell, including cell proliferation (Balda, Garrett, and Matter 2003), differentiation (Balda and Matter 2000), transcription (Huerta et al. 2007), and immune responses (Kumar, Liu, and Rice 2012). These regulatory properties are mediated by a host of adaptor proteins that associate with the cytoplasmic junctional plaque thereby linking tight junctions to cellular signaling networks (Balda and Matter 2009; González-Mariscal et al. 2014; Schneeberger and Lynch 2004). The specific localization of these adaptor proteins, which include transcription factors, is also important for their regulatory functions because some cycle between the apical junction and the nucleus (Balda 2003; Matter and Balda 2007). The diagram in Figure 32 shows the basic composition of the tight junction complex. The barrier-forming transmembrane proteins seal the intracellular space and are coordinated through interactions with scaffolding and adaptor proteins that make up the cytoplasmic plaque.

The pathophysiology of tight junctions has been studied hand-in-hand with their natural functioning, and pathogens have been implicated in the modulation these important cellular structures since shortly after their discovery (Takeuchi 1967). Polarized epithelial cells line the surfaces of our bodies that come into contact with our environment and serve as a first line of defense against pathogenic invaders. Strategies utilized by both bacteria and viruses to overcome the epithelial barrier have been examined in some detail (Guttman and Finlay 2009; Zihni, Balda, and Matter 2014) and include direct interactions with tight junction components (Sonoda et al. 1999) and indirect disruption by pathogenic effector proteins (Boyle, Brown, and

Finlay 2006; Simonovic et al. 2001; Viswanathan et al. 2004). Several viruses directly bind to tight junction components for entry into host cells, including hepatitis C virus (Evans et al. 2007), reovirus (Barton et al. 2001), coxsackie virus (Bergelson et al. 1997), and adenovirus (Bergelson et al. 1997). SARS-Coronavirus E protein binds the adaptor protein MPP5 leading to altered tight junctions and epithelial morphology (Teoh et al. 2010). The adenovirus E4-ORF1 protein interacts with several tight junction adaptors through a PDZ domain and these interactions affect cell polarity and barrier integrity (Latorre et al. 2005).

Given the diverse cellular processes influenced by tight junctions and the central role these structures play in the maintenance of cell polarity, it is not surprising pathogens have evolved to target them. The influenza virus is one of the most common causes of respiratory disease in humans and infections can lead to acute respiratory distress syndrome (ARDS) and death (Taubenberger and Morens 2008). Severe and fatal cases are characterized by pulmonary edema from damage to the epithelial barrier and systemic viremia (de Jong et al. 2006). These pathological outcomes have been primarily attributed to pneumonia from secondary bacterial infections and hypercytokinemia driven by high viral loads (Brundage and Shanks 2008; de Jong et al. 2006; Metersky et al. 2012; Z. Wang et al. 2014). However, recent work implicates the influenza NS1 protein in the disruption of tight junctions through direct interactions with tight junction adaptor proteins, Dlg1 and Scribble. The binding is conferred by a PDZ motif (ESEV) located on the C-terminus of NS1 proteins from highly virulent H5N1 strains (de Jong et al. 2006). Golebiewski et al. found that this association led to mislocalization of tight junction proteins Lin7C, ZO-1 and occludin (Golebiewski et al. 2011). These results suggest that the pathophysiology of severe influenza infections may be driven, in part, by the disruption of the epithelial barrier by NS1. Based on the proteomics work presented in Chapter 3, we hypothesized that influenza NEP interacts with tight junction components during infection and this chapter presents data showing that NEP affects tight junction functioning in polarized cell monolayers.

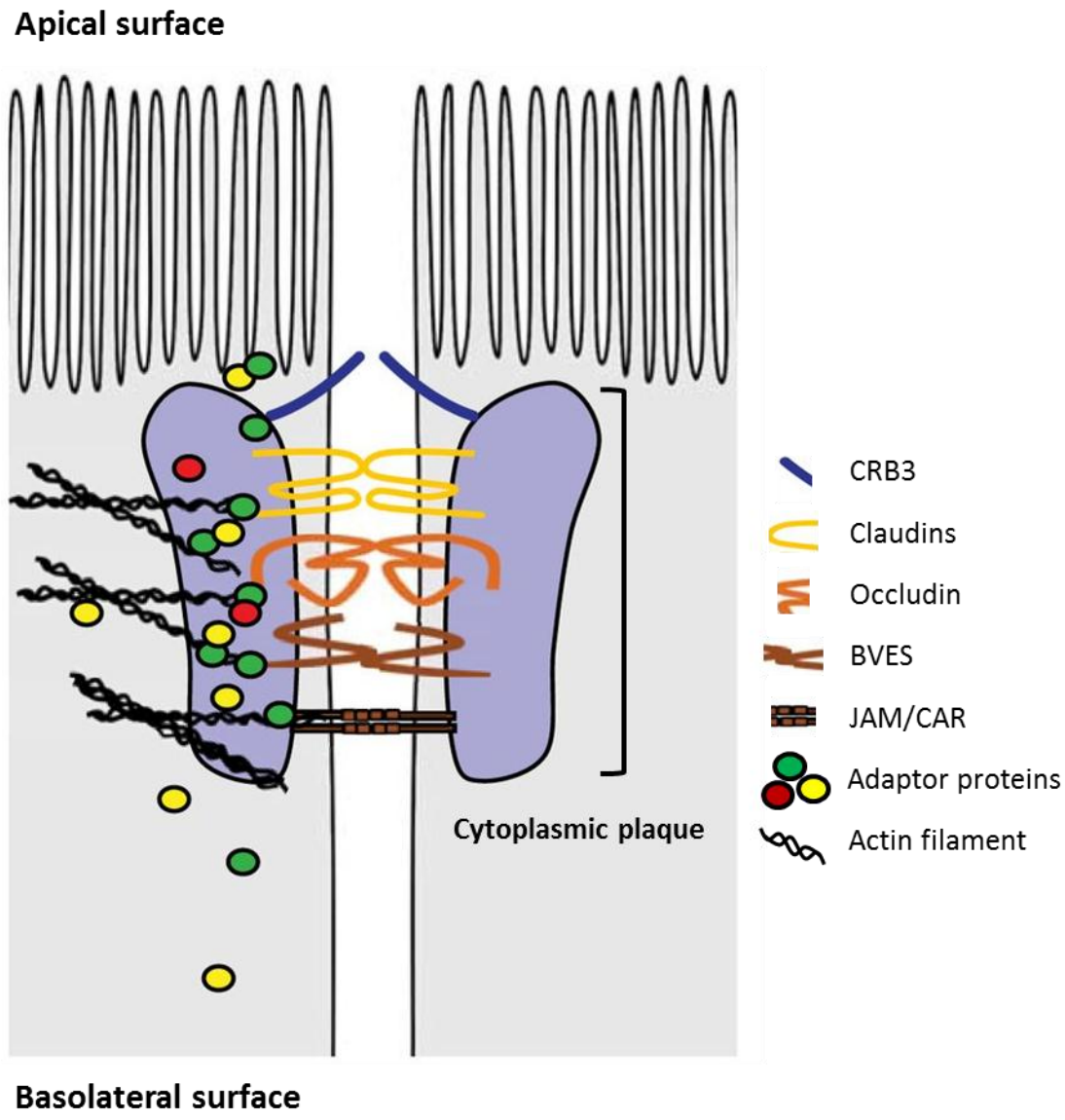


Figure 32. Model of the tight junction complex
Tight junctions are composed of transmembrane proteins that bridge the space between two adjacent cells. CRB3 regulates polarization and apical differentiation. Claudins, occludin and BVES mediate paracellular permeability while JAM and CAR are essential for junction integrity. The organization and functioning of the transmembrane proteins are regulated through interactions with a network of cytoplasmic proteins that localize to the cytoplasmic plaque. Formation of the cytoplasmic plaque is mediated by scaffolding proteins that link the transmembrane proteins and associated factors with filamentous actin. Figure adapted from Zihni et al. 2014.

4.3 Methods

4.3.1 Lentiviral Production

Lentivirus was produced using standard procedures. First, 75% confluent HEK293T cells in 6 well plates were transfected using 5µL Mirus LT1 reagent in 200µL OptiMEM with the following plasmids: 1µg vector encoding a tetracycline responsive (Tet-on) promoter upstream of either NEP or GFP, 0.75µg psPAX2 encoding gag and pol, and 0.2µg of MD2.G encoding lentiviral envelope protein. At 24 hours post-transfection, media was replaced with 2mL cDMEM. At 48 hours post-transfection, supernatant containing lentivirus was cleared by centrifugation and 1.5mL was added to 75% confluent MDCKII or MDCKI cells. Transduced cells were grown for two days and then expanded into 10cm plates with either 2 or 4µg/mL puromycin respectively for selection. Transduced, selected MDCKII or MDCKI cells were identified by the transgene prefix (ie. NEP-MDCKII or GFP-MDCKI).

4.3.2 Fluorescein isothiocyanate (FITC)-inulin Diffusion Assay

After transduction and puromycin selection, cells (150K) were seeded on 12mm, polyester transwell inserts with 0.4µm pores (VWR, #29442-078) with or without doxycycline (dox), (0.5µg/mL). Cells were grown for varied times (detailed in results) prior to addition of 2mg/mL FITC-inulin (Sigma, F3272-1G) dissolved in growth media. Fluorescence readings were performed on a SpectraMax plate reader.

4.3.3 Transepithelial electrical resistance (TEER) measurements

TEER measurements were performed on MDCKII or MDCKI cells grown in transwells as described in 4.3.2 using a Millipore Millicell-ERS resistance meter.

4.3.4 Immunofluorescence

MDCKI cells were seeded on glass coverslips and grown for 48 hours in the presence or absence of doxycycline (0.5µg/mL). Cells were fixed with 3.7% formaldehyde for 15 minutes then permeabilized with 0.1% Triton X-100 for 40 minutes at RT. Cells were blocked with 5%

BSA, and 0.02% Triton X-100 at 4°C for 4 hours. Alexa-488 conjugated anti-ZO-1 and Alexa-568 conjugated anti-occludin antibodies (Life Technologies, 339188 and 331594) were added at 1:5000 and 1:4000 dilutions in 0.02% Triton X-100 and 0.1% BSA for 16 hours. Cells were washed 3x in PBS and then mounted on glass slides with SlowFade® Gold Antifade Mountant with DAPI (Life Technologies, S36939). Cells were imaged using a Delta Vision deconvolution microscope with an Olympus PlanApo N 60x objective.

4.4 Results

4.4.1 Transepithelial barrier function in MDCKII cells is impaired by expression of NEP but not GFP

First the time period for the establishment of barrier integrity in MDCKII cell monolayers was assessed. Parental MDCKII cells were plated in transwells and FITC-inulin diffusion was measured at 24, 48, and 72 hours after seeding at two different densities 100,000 and 200,000 (Figure 33A). On day 1, FITC-inulin readily diffuses across the transwell but by day 2, the monolayers had established junctional integrity. The levels of diffusion were comparable between days 2 and 3 and between the different seeding densities. The transepithelial electrical resistance was also measured in transduced but uninduced MDCKII cells (Figure 33B). The resistance rose from day 1 to day 2, but plateaued thereafter until declining on day 5. This is consistent with the observations in Figure 33A where a barrier formed between the first and second days after seeding. Overall the TEER values in these MDCKII cells were low given that the baseline resistance for a transwell without any cells (blank) was already greater than 100Ω.

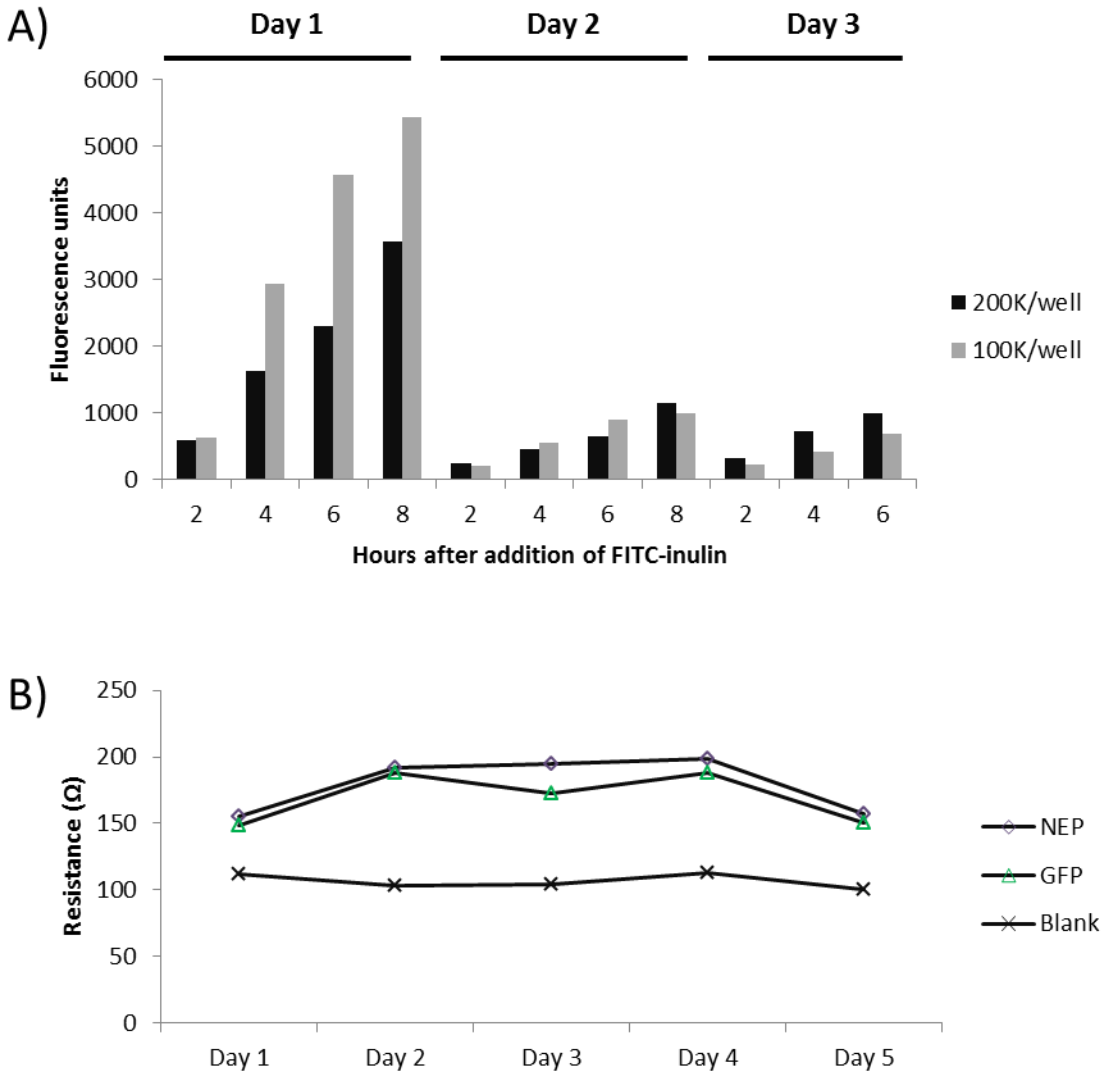


Figure 33. Timeline for the establishment of tight junction integrity in MDCKII cells
(A) MDCKII cells were plated at 100K and 200K/per transwell. At 24 hours post seeding FITC-inulin was added to two transwell cultures (one from the 100K seeding and one from the 200K). The well solution was sampled at 2 hour intervals for 8 hours on days 1 and 2 and every 2 hours for 6 hours on day 3. (B) TEER measurements were performed on uninduced cells or cells expressing GFP or NEP-MDCKII on days1-5 post seeding.

To measure the effect of NEP on the FITC-inulin diffusion rate, cells were seeded in transwells and induced to express NEP or GFP (6 replicates each) with 0.5% doxycycline (dox). FITC-inulin was added to the cultures after 3 days and the well solution was sampled at 4, 8, and 24 hours (Figure 34A). NEP induction resulted in a roughly 2 fold increase in the amount of FITC-inulin diffusing through the cell monolayer compared to GFP expression. Cell lysates from 6 different NEP-MDCKII cultures (from 6 different transductions) or GFP-MDCKII cells were blotted for NEP after doxycycline induction (Figure 34B).

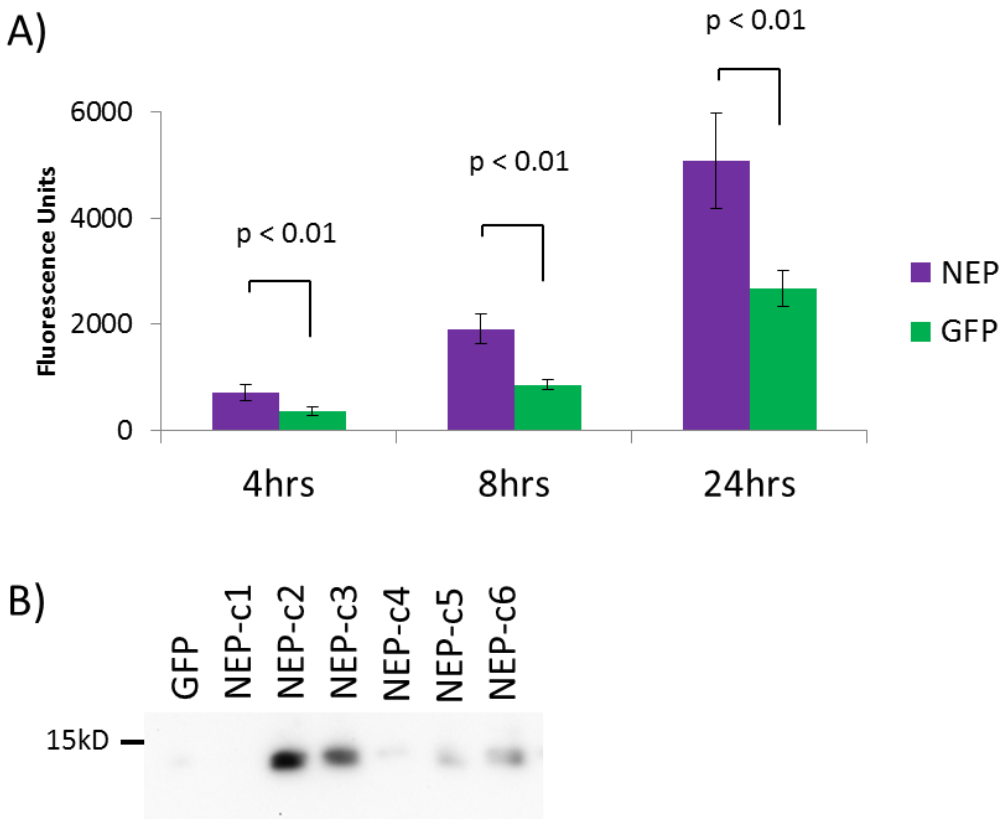


Figure 34. Induction of NEP expression increases FITC-inulin diffusion versus GFP expression in MDCKII cells.
(A) GFP and NEP-MDCKII cells (100K) were plated in transwells. At 3 days post seeding FITC-inulin was added to all cultures and the well solution was sampled at 4, 8, and 24 hours. Values are averages of 6 biological replicates. Error bars represent standard deviations. The *p* values were calculated with Student's t-test. (B) shows a western blot for NEP from GFP or NEP-MDCKII cells after induction with doxycycline. Samples from 1 GFP clone and 6 different NEP clones are shown. Clone c5 was used in Figure 34A.

4.4.2 Tight junction function in high resistance MDCKI cells is impaired by expression of NEP but not GFP

NEP-MDCKI or GFP-MDCKI cells were used to test whether NEP has an effect on ion permeability in addition to the observed effect on macromolecular diffusion. The low inherent electrical resistance of MDCKII cells (Figure 33B) precluded use of TEER assays in that system. Again baseline measurements were first established in the parental cell line (Figure 35). In contrast to MDCKII cells the MDCKI monolayers continued to build resistance until reaching a peak of approximately 8,000Ω 4 days after seeding (Figure 35A). Macromolecular diffusion was also much more attenuated in MDCKI monolayers as shown in Figure 35B. Next, the effect of NEP expression on TEER and FITCI-inulin diffusion in MDCKI cells was tested (Figure 36). Five days after cells were with or without doxycycline to induce expression of NEP or GFP, TEER was measured. A significant decrease in resistance was observed when NEP-MDCKI cells were induced to express NEP, whereas expression of a control protein had no effect (Figure 36A). Increased paracellular flux of FITC-inulin was also observed when NEP was induced, but not when untransduced parental MDCKI cells were treated with doxycycline (Figure 36B).

Since the previous experiments were all carried out with early induction (cells were seeded with doxycycline to induce expression before tight junctions formed), the effects of late expression of NEP were also examined. NEP, GFP or parental MDCKI cells were allowed to establish tight junctions for 5 days then FITC-inulin was added to all transwells. Twenty-four hours later doxycycline was added to 2 of 4 transwells from each group. Fluorescence signal in the well solution was examined 24 hours after induction with doxycycline. The fold change in the signal between treated and untreated groups is shown in Figure 36C. There was a significant difference between the NEP and GFP-MDCKI fold changes at 24 hours by Student's t-test while the difference between the NEP and parental MDCKI cells was just below significance ($p = 0.054$). TEER was measured in this experiment as well but differences were not significant.

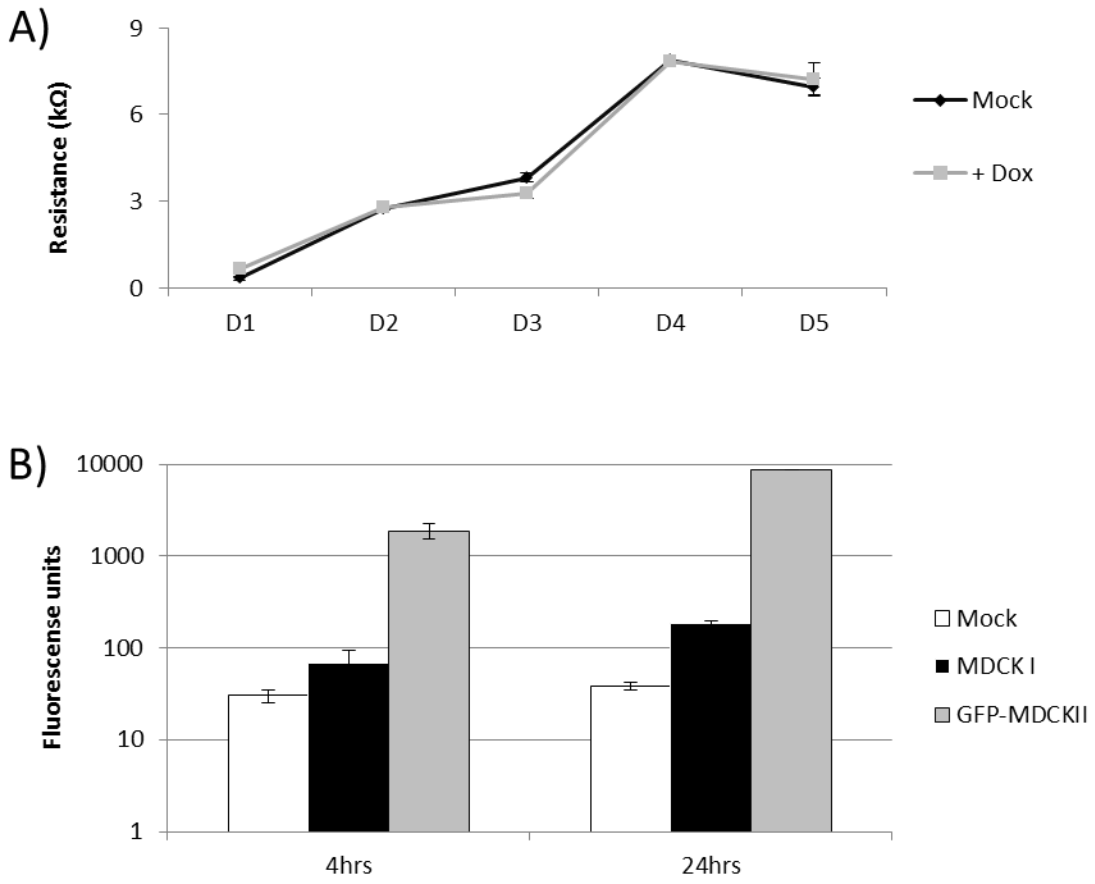


Figure 35. MDCKI cells establish cell monolayers with high resistance and low permeability

(A) TEER measurements were performed on MDCKI cells treated or not (mock) with doxycycline (0.5µg/mL). (B) At day 4 post seeding, FITCI-inulin was added to MDCKI cells or uninduced GFP-MDCKII cells and fluorescence was measured in the well solution at 4 hours and 24 hours. Mock represents background signal in culture medium. All values are the means of 2 biological replicates. Error bars represent standard deviations.

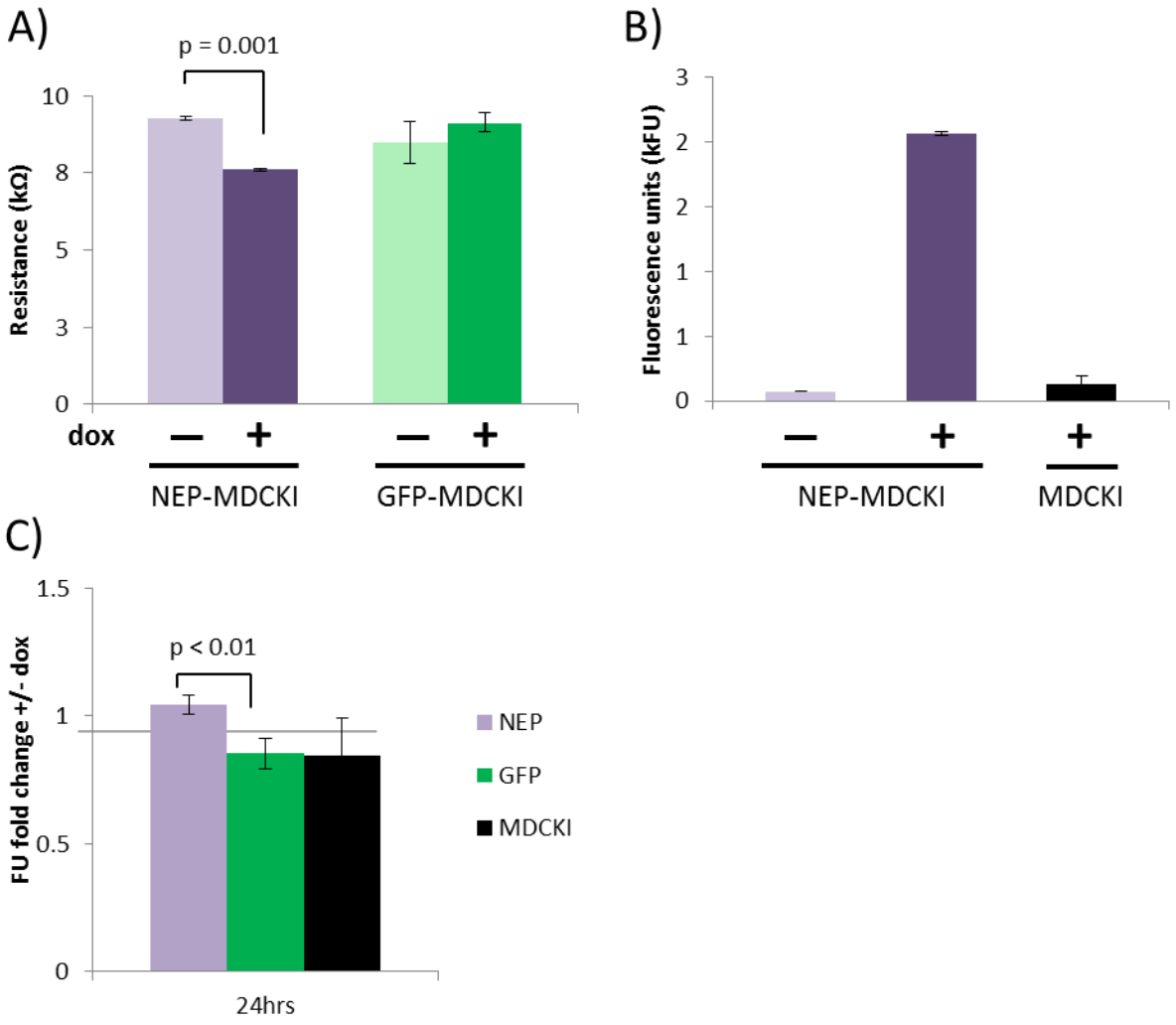


Figure 36. Induction of NEP expression increases FITC-inulin diffusion and decreases TEER in high resistance MDCKI cells

(A) TEER measurement at day 5 post-seeding for NEP or GFP-MDCKI cells +/- dox. (B) MDCKI cells were seeded in transwells with (NEP and untransduced MDCKI) or without (NEP only) dox. FITC-inulin was added to the transwells on day 5 and well solution was assayed for fluorescence at 48 hours. In (C) FITC-inulin was added to all transwells (NEP, GFP and untransduced MDCKI cells) at 5 days post seeding. Dox was added the following day to half the cultures and fluorescence signal in the well solutions of treated (n=2) and untreated (n=2) cultures was measured 24 hours after induction. Data from (A), (B) and (C) were from independent experiments. All values are the means from two biological replicates. Error bars indicate standard deviations.

4.4.3 Tight junction protein occludin is mislocalized when NEP is induced

We used immunofluorescence microscopy to examine the cellular localization of tight junction proteins ZO-1 and occludin. GFP-MDCKI or NEP-MDCKI cells were grown with or without doxycycline, then fixed and stained (Figure 37). Normal localization along the cell boundaries was observed for both ZO-1 and occludin in the absence of doxycycline (Figure 37A). The staining patterns remained normal when the MDCKI cells were induced to express GFP (Figure 37B) but occludin puncta on the cellular interior were detected when NEP was induced.

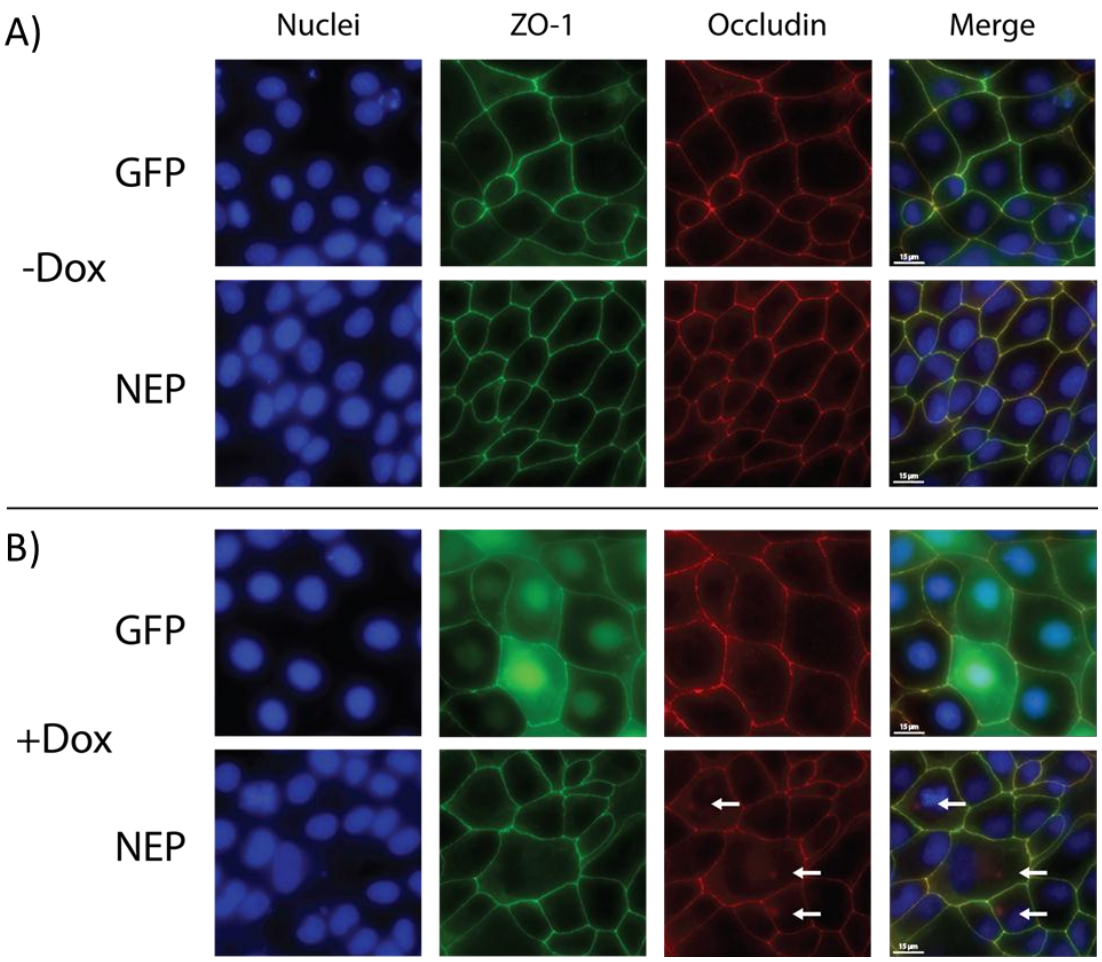
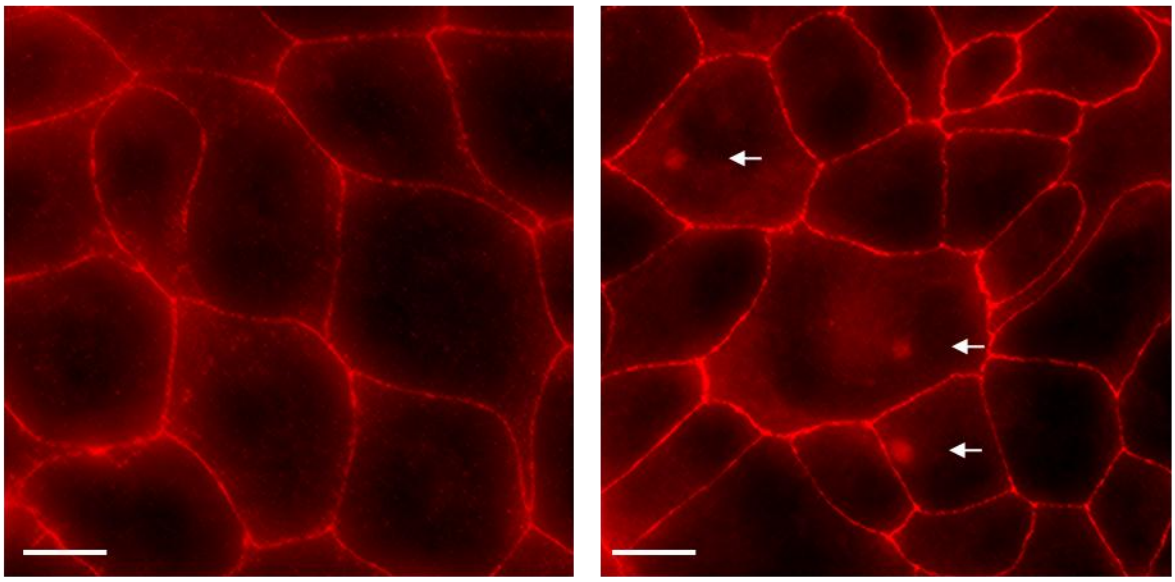


Figure 37. Tight junction protein occludin is mislocalized in the presence of NEP expression

GFP or NEP-MDCKII were grown for 4 days on glass coverslips +/-dox then fixed and stained for ZO-1, occludin, or DNA. Top section (A) shows staining under growth conditions without doxycycline, and ZO-1 and occludin show typical localization. In (B) the localization of ZO-1 and occludin were unchanged by GFP expression. However when NEP was induced, occludin was mislocalized to large round structures in the cellular interior (white arrows). White scale bars in the merged images represent 15μM. Panels showing occludin staining are enlarged in Figure 38.

A)



B)

Percentage of cells with occludin puncta

	1	2	3	4	5	Ave	Total cells
NEP-MDCKI	16%	13%	14%	16%	13%	14%	152
GFP-MDCKI	0%	0%	6%	0%	7%	2%	107

$p < 0.001$

Figure 38. High contrast enlargement of occludin staining in induced GFP or NEP-MDCKI cells

Higher contrast images from Figure 37B showing occludin localization in induced GFP or NEP-MDCKI cells were enlarged to highlight occludin puncta. Occludin puncta in 5 separate image panels from induced GFP or NEP-MDCKI cells were quantified. Values reported in (B) are the percentage of cells in each panel containing occludin puncta. The *p* value was calculated using Student's t-test. White scale bars represent 15μM.

4.5 Discussion

This chapter presents results from experiments that tested for an effect of NEP on tight junction function. Tight junctions form a dynamic barrier between epithelial cells that line the respiratory and gastrointestinal tracts and influenza infects each of these systems in different hosts (humans and birds respectively). The integrity of the barrier can be assessed by measuring the resistance to ion flow and also the flux of macromolecules using cultured cells that form polarized monolayers. Madine-Darby canine kidney (MDCK) cells are one of the most commonly used cell lines for studying tight junctions and characterization of their epithelial properties spans 4 decades beginning in the 1970's and continuing today (Cereijido et al. 1978; Lu, Stewart, and Wilson 2015; Misfeldt, Hamamoto, and Pitelka 1976). Two human colon carcinoma cell lines, Caco-2 (Hidalgo, Raub, and Borchardt 1989) and T84 (Dharmsathaphorn et al. 1984) are also widely used, but we opted for the MDCK line because it has been established that influenza replicates and buds efficiently in these cells (Tobita et al. 1975). Interestingly there are two different MDCK cell lines that differ markedly in their barrier properties and both were used in this work (Stevenson et al. 1988). MDCKI cell monolayers exhibit much higher transepithelial electrical resistance than do those of MDCKII cells due to the differential expression of claudin proteins (Furuse et al. 2001).

The functional integrity of tight junctions is typically measured in two ways: resistance to ion flow (TEER) and paracellular flux of macromolecules. Measurements are conducted in transwells which suspend a semipermeable membrane in growth media to facilitate polarization. This situation mimics *in vivo* conditions by allowing both the apical and basolateral surfaces to contact the culture solution. Since the apical face is physically isolated from the basolateral face by the structure of the transwell, ions or macromolecules must travel through the monolayer to reach the other side. In this manner the resistance of the monolayer can be measured by placing electrodes in the apical and basolateral culture solutions, and paracellular flux can be monitored by sampling the basolateral solution after adding inert tracer macromolecules to the apical side.

We first established a timeline for barrier formation in MDCKII cells and found that by the third day after seeding, the monolayers had formed tight junctions that impede the paracellular flux of inulin (a plant-derived fructofuranose polysaccharide that is not metabolized by humans (André et al. 1996) and is widely used to study tight junctions) coupled to a fluorescent molecule (FITC) (Figure 33A). The inulin used in our experiments had a molecular weight range of 2,000 to 5,000 daltons. In order to test for an effect of NEP on tight junctions we transduced MDCKII cells with lentivirus to produce cell lines with doxycycline-inducible expression of either GFP or NEP. We observed increased flux of FITC-inulin across the NEP-MDCKII monolayers compared to the control GFP-MDCKII monolayers when the cells were cultured with doxycycline indicating that expression of NEP decreased barrier integrity (Figure 34).

We also wanted to compare ion flow measurements between monolayers of cells expressing NEP or GFP but the MDCKII cells do not establish barriers with high resistance (Figure 33B) so we used lentivirus to generate transgenic MDCKI cell lines. MDCKI cell monolayers reached maximal resistance by the fourth day after seeding in transwells, and the junction integrity was not affected by the presence of doxycycline (Figure 35A). The resistance levels of the MDCKI cells have been reported to be 30-60 times greater than levels in MDCKII cells (Stevenson et al. 1988) and our observations fall squarely in this range (Figure 33B versus Figure 35A). We also observed significantly lower rates of FITC-inulin flux in the MDCKI cell line compared to the MDCKII cell line (Figure 35B). We measured the TEER of NEP and GFP-MDCKI cell monolayers grown for 5 days both with and without doxycycline and observed a significant decrease in resistance when NEP was expressed but not when GFP was expressed (Figure 36A). We also observed a significant increase in the paracellular flux of FITC-inulin when the NEP-MDCKI cells were cultured with doxycycline but not when the parental MDCKI cells were grown with doxycycline (Figure 36B). These results indicate that the specific induction of NEP expression led to decreased monolayer resistance and increased permeability. We also assessed whether late expression of NEP after tight junctions had already formed would affect the barrier properties. We observed a significant increase in FITC-inulin flux through the NEP-MDCKI cells with late addition of doxycycline compared to GFP-MDCKI cells but not compared to the parental MDCKI cells (though the result was just outside the significance threshold of p =

0.050). TEER was also measured after late addition of doxycycline but no significant difference was detected. This might due to the lower sensitivity of the TEER assay which utilizes a single, instantaneous measurement, compared to the diffusion assay which measures the accumulation of signal over time.

It is important to note that TEER and FITC-inulin flux measure distinct properties of tight junctions that can be independently regulated. Balda et al. observed seemingly paradoxical results after a mutant version of occludin was expressed in MDCKII cells—increased flux and increased TEER (Balda et al. 1996). Hasegawa et al. observed the same phenomenon after Rho activation in MDCKII cells (Hasegawa et al. 1999). Subsequent studies have shown altered TEER but unchanged paracellular flux when claudin levels are varied in MDCK cells (Amasheh et al. 2002; Van Itallie, Rahner, and Anderson 2001). Furthermore TEER is an instantaneous measurement while paracellular flux measures the accumulation of tracer molecules overtime. As a result, the magnitude of the differences in the measurements can be quite large as we observed in Figure 36A versus B. Importantly, the differences in TEER and flux in this work are comparable to published data. Golebiewski et al. observed a decrease in TEER of approximately 20% between control and experimental conditions (Figure 39A) while we measured a decrease of 18% when NEP was expressed in MDCKI cells Figure 36A. Similarly for FITC-inulin flux, Golebiewski et al. reported an increase of approximately 150% between control and experimental samples (Figure 39B) and we observed an increase of 100% (Figure 34).

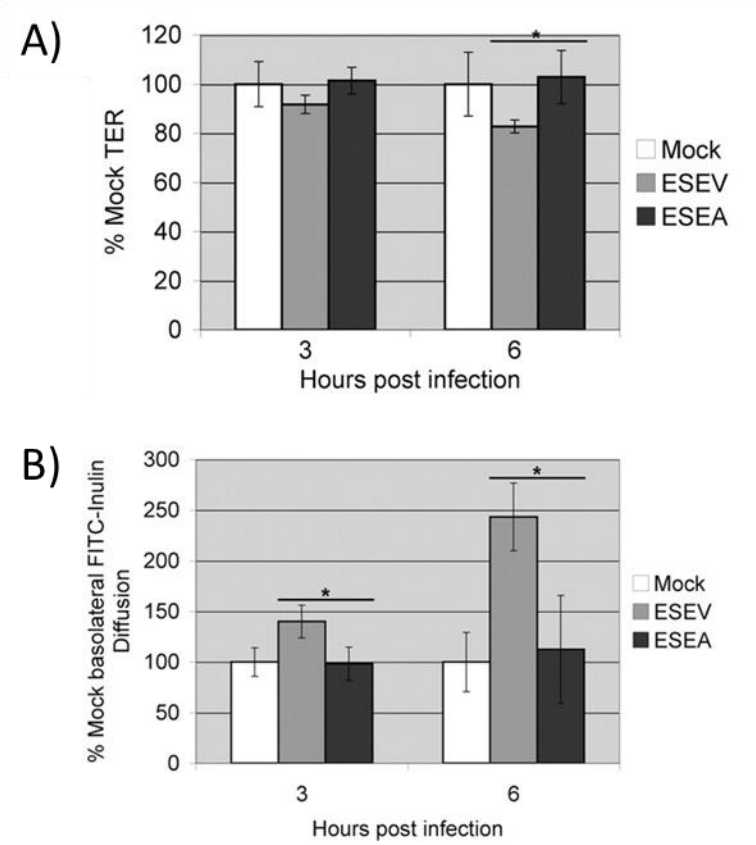


Figure 39. TEER and FITC-inulin diffusion in infected monolayers from Golebiewski et al. 2011

In (A) Golebiewski et al. observed an approximately 20% decrease in resistance in monolayers of MDCK cells infected with an influenza strain expressing NS1 variant ESEV compared to controls (Mock infected and strain ESEA) while flux increased approximately 150% between the control and experimental conditions (B).

After it was established that expression of NEP had an effect on tight junction function, we analyzed the localization of tight junction proteins ZO-1 and occludin by immunofluorescence. ZO-1 is a scaffolding protein essential for tight junction formation (Stevenson et al. 1986), and occludin is a transmembrane protein that influences barrier properties (Furuse et al. 1993). ZO-1 links occludin to the actin cytoskeleton (Fanning et al. 1998) and both proteins are used to visually assess tight junction integrity by immunofluorescence microscopy (Golebiewski et al. 2011). When we stained NEP or GFP-MDCKI cells for ZO-1 we did not observe any difference between cells grown with or without

doxycycline (Figure 37). However, we did observe large diffuse occludin puncta in some of the NEP-MDCKI cells grown with doxycycline. The panels showing occludin localization in Figure 37B were enlarged with increased contrast for better visualization in Figure 38A. We compared the percentage of NEP-MDCKI cells with occludin puncta to the control cells and found a significant difference (Figure 38B). The overall rate of puncta observed was somewhat low at 14% of MDCKI cells expressing NEP and this may be due to heterogeneous expression levels in the cell line. Variable levels of expression were observed for the transgenic NEP and GFP-MDCK cell lines both by western blot (Figure 34B) and visually by fluorescence microscopy (data not shown). After induction, the fluorescence of GFP-MDCK cultures was patchy with some cells producing only low levels of GFP while others produced high amounts. It seems likely that the expression of NEP followed a similar pattern. However, we were not able to successfully visualize NEP by immunofluorescence so this remains unconfirmed.

In sum, the results from our TEER and paracellular flux assays indicate that NEP altered the barrier properties of tight junctions and that the magnitudes of the observed effects were comparable to those in published reports. We also observed differential localization of the tight junction protein occludin between MDCKI cells expressing NEP and GFP and it is possible that this mislocalization led to altered barrier properties. Further discussion of these results in the context of the viral lifecycle as well as interesting directions for future studies are presented in Chapter 5.

Chapter 5: Perspectives

5.1 Lessons from the study of host interactors of NEP

The study of host-pathogen interactions is central to the development of new therapeutic strategies. This principle is embodied in the title of a recent publication "Influenza Virus-Host Interactome Screen as a Platform for Anitviral Drug Development" (Watanabe et al. 2014). In that work, researchers performed IP-MS with FLAG-tagged influenza proteins expressed from plasmids to identify host interactors. Functional assays were subsequently used to determine which of the interacting host factors were functionally important for viral replication, and pharmaceutical targeting of several host factors led to reduced viral replication. However there are some limitations to the study. Importantly, the interactions were not interrogated during an active infection. The cellular state is drastically altered during infection (Huang et al. 2001; Jovanovic et al. 2015), and it is likely that some interactions will only occur during active viral replication. Ideally, chemotherapies would target interactions unique to infected cells to minimize effects on normal functioning.

Studying host-pathogen interactions during active infection is a challenging endeavor. There have been numerous studies, including many referenced in this work, that have identified interactions between pathogens and their hosts, but only a handful of these studies have been conducted during an active infection. Reasons for this range from the additional precautions that must be taken when experimenting with infectious agents to the difficulties associated with genetically modifying viral genomes. Researchers who used IP-MS to study host interactions with the influenza polymerase complex stated that "[a]n optimal experimental design would be to rescue viruses in which affinity purification tags have been introduced into the polymerase subunits. However, such genetic modifications led repetitively to lethal virus phenotypes (data not shown)" (Jorba et al. 2008). Instead, they relied on plasmid-based expression of tagged polymerase proteins—a suboptimal, non-infectious context.

We overcame several hurdles in the development of viruses that express tagged NEP during infection. Early attempts to tag NEP on the NS segment met with failure, and led us to place the entire coding sequence of tagged NEP on the NA segment. This enabled the rescue of viruses encoding tagged NEP and represented a significant milestone because it was not known at the outset whether a tag would be functionally compatible with NEP's essential roles in the viral lifecycle. After establishing a viable tagging strategy, we had to re-engineer different aspects of the recombinant virus for IP-MS studies. Our final construct, the 32A-3xFLAG virus, featured an N-terminal 3xFLAG-tag on NEP and a triple-tandem repetition of the 2A self-cleaving peptide sequence (2.4.2.4).

Obtaining consistent results from IP-MS was also challenging. There are many steps in an IP-MS experiment, from growth and infection of cells through data acquisition and analysis that can each inject variability into the final output. We used several different strategies to control for variability but ultimately the quality of the IP-MS results will be a direct reflection of the consistency of the entire procedure. Thus optimization work prior to execution of the IP-MS experiment is highly important. We noticed greater consistency (albeit higher background) between SBP affinity purifications compared to FLAG and this was probably because the streptavidin-Dynabeads are commercially prepared while the anti-FLAG beads were prepared prior to each IP. Coupling anti-FLAG M2 to the PrG-Dynabeads is a multi-step process subject to batch variations (both the antibody and beads) and experimental variation in the procedure. Simplifying the overall procedure would help minimize variability. It is also highly important to minimize the background in the IP and to maximize the signal from the biologically relevant interactors. One way to accomplish this is using competitive elution, and the signal from NEP specific interactors might have been increased by eluting with FLAG peptide as opposed to high-salt and low pH.

Once new host interactors have been identified by IP-MS, it can be difficult to prioritize candidates for additional studies. To address this issue we used I-DIRT to rank identified host interactors by their H/L peptide ratios (a quantitative measure of the stability of their interaction with NEP) and we conducted follow-up studies on an interaction with EEF1A1.

Despite the fact that EEF1A1 was among the 16 host factors with high H/L ratios we were not able to conclusively show an interaction specifically with NEP using co-IP (data not shown). It remains possible that NEP and EEF1A1 do interact *in vivo*. In our case, integrating published data and a functional enrichment analysis proved helpful in identifying avenues for further study.

In general we found it difficult to express NEP in mammalian cells. Transfection of HEK293T cells with plasmid encoding native NEP did not result in detectable levels of expression (data not shown). Placing N-terminal tags on NEP such as 3xFLAG improved production but expression levels remained attenuated compared to 3xFLAG-GFP (Figure 28). Based on our experience with plasmid expression in HEK293T cells, we used lentivirus for NEP expression in MDCK cells, and again observed low expression levels compared to GFP (data not shown). In retrospect, inclusion of the N-terminal 3xFLAG tag might have increased expression levels and facilitated detection of NEP by immunofluorescence microscopy.

5.2 Potential roles for tight junctions in the influenza lifecycle

In wild birds influenza is a disease of the gastrointestinal tract and in humans the virus causes respiratory disease. In both cases it must invade, replicate in, and bud from polarized epithelial cells. Therefore it is probable that the virus utilizes cellular processes and features common to these cell types. All polarized epithelial cells maintain distinct apical and basolateral surfaces through the formation of tight junctions between neighboring cells. The resulting barrier effectively isolates the basolateral membrane from the external milieu. Numerous studies have demonstrated that this barrier is compromised as result of infection by both bacterial and viral pathogens that enter and exit through epithelial cells (discussed in 4.2). In the case of influenza, severe infections are characterized by disruption of the respiratory epithelium that results in fluid accumulation in the lungs and viral dissemination to other tissues. The body of work presented in this thesis, as well as recent studies by other researchers, suggest that the integrity of the epithelial barrier may be directly modulated by the influenza virus through interactions with tight junction components.

Why would altered tight junctions benefit the virus? Golebewski et al. (2011) proposed that the disruption of tight junctions and cell polarity are likely to enhance "viral replication, dissemination in the host, or spread to other hosts", and it has been known for many years that tight junctions are not essential for budding (Rodriguez-Boulan, Paskiet, and Sabatini 1983). Tight junction disruption may provide access to lipids or other biomolecules sequestered in the basolateral membrane. It could also alter the curvature or other physical characteristics of the cell to favor budding. However these arguments are not entirely satisfying because during a natural infection, influenza components localize specifically to the apical surface in preparation for budding and the ensuing release of new virions into the environment enables dissemination to a new host thereby perpetuating the viral lifecycle. In this light, the maintenance of polarity would seem to favor the assembly and budding process. It is also important to consider that the tight junction complex is involved in the regulation of a number of cellular processes in addition to barrier maintenance (discussed in 4.2). Influenza may, in fact, be targeting one of these other pathways and the observed effects on tight junction integrity may actually be a secondary outcome of this modulation.

The tight junction complex also serves as a location beacon for the trafficking of cellular components to and from the apical plasma membrane. Numerous proteins involved in vesicular trafficking localize to tight junctions including members of the Rab family of guanosine triphosphatases (GTPases), exocyst proteins and SNARE (Soluble N-ethylmaleimide-sensitive factor Associated REceptor) components (Grindstaff et al. 1998; Köhler, Louvard, and Zahraoui 2004; Naydenov et al. 2012). These proteins help to sort, target, and release cargos that arrive and depart from the apical-basolateral interface (Köhler and Zahraoui 2005). Some of these cargos are components of the junctions themselves including the transmembrane proteins. Occludin, claudins, and cadherin are subject to Rab-regulated turnover through endocytosis and recycling (Fletcher et al. 2014; Linford et al. 2012; Lu et al. 2014). Interestingly this process has also been linked to the Crumbs complex which localizes to tight junctions and is a key regulator of polarity and cell shape (Bulgakova and Knust 2009). The 3 core cytoplasmic proteins of the complex (MPP5, INADL, and LIN7C) were identified in our work as NEP interactors, and Wang et al. found that silencing of MPP5 lead to misregulated trafficking of E-cadherin to the cell

periphery and its accumulation in puncta (Wang, Chen, and Margolis 2007). We observed a similar phenotype with respect to occludin in MDCKI cells expressing NEP (Figure 38).

Additionally it has been established through independent studies that both Crb (the transmembrane component of the Crumbs complex) and influenza vRNPs are trafficked to the apical plasma membrane through endosomal recycling pathways in a Rab11 dependent manner (Eisfeld et al. 2011; Momose et al. 2011; Roeth et al. 2009). These studies revealed that vRNPs newly exported from the nucleus concentrate at the microtubule organizing center (MTOC) prior to transport to the apical membrane on Rab11 positive recycling endosomes. However Eisfeld et al. noted that Rab11 did not co-localize with vRNPs at the plasma membrane suggesting that they are delivered to staging areas before incorporation into budding virions (Eisfeld et al. 2011). They proposed that the cytoplasmic trafficking of vRNPs consists of at least four steps: (i) aggregation at the MTOC, (ii) association with Rab11 for transport, (iii) accumulation near the apical plasma membrane, and (iv) selection/sorting for budding. However details regarding the hosts factors and processes involved in steps (iii) and (iv) are still unknown. Based on our data, we hypothesize that NEP interacts with the Crumbs complex proteins to promote the accumulation of vRNPs at the tight junction complex and/or the subsequent targeting of vRNPs to bud sites on the plasma membrane. Figure 40 illustrates how our findings could enhance the understanding of processes involved in the final steps of the viral lifecycle.

Another host factor in the NEP interactome, angiomotin (AMOT), associates with both tight junctions and Rab11 positive endosomes (Bratt et al. 2005; Heller et al. 2010). AMOT was one of the host proteins identified in this work as specific to the SBP-NEP and 3xFLAG-NEP affinity purifications by Coomassie blue staining (Figure 19 and Figure 23), and it was also identified in the NEP interactome (Table 8). However it was not grouped into the tight junction functional cluster, likely representing a false negative result of the enrichment analysis. AMOT interacts preferentially with curved, cholesterol rich membranes and promotes the tubulation of liposomes (Heller et al. 2010). Intriguingly, AMOT plays a key role in release of HIV-1 virions

from cell membranes (Mercenne et al. 2015), thus NEP could be interacting with AMOT to promote influenza budding (Figure 40).

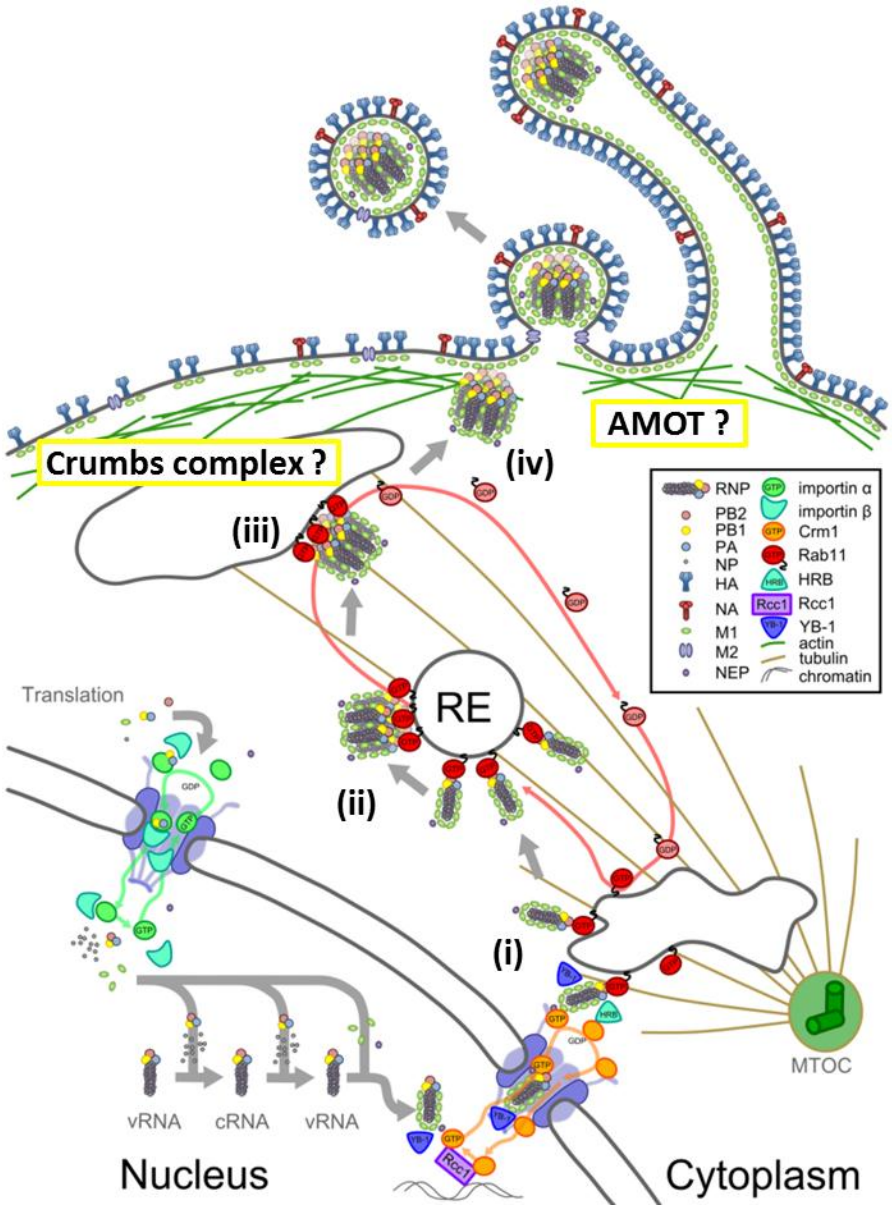


Figure 40. Diagram showing the post entry steps in the influenza virus lifecycle
Replication of the viral genome as well as production of vmRNAs takes place in the nucleus. After translation, some viral proteins are imported into the nucleus to form new vRNPs (discussed in 1.1.3). Following nuclear export, vRNPs aggregate on RAB11 positive recycling endosomes (i) for transport to the apical membrane (ii). Rab11 dissociates from vRNPs near the cell surface (iii) and it is not known how they reach the plasma membrane (iv). The processes and factors involved in budding are also not fully understood. It is possible interactions between NEP and the Crumbs polarity complex enable the assembly and targeting of vRNPs to bud sites on the plasma membrane and interactions with AMOT may facilitate bud formation and release (yellow boxes). Figure adapted from Eisfeld, Neumann, and Kawaoka 2014.

5.3 Future directions

Additional studies exploring the role of NEP, AMOT and the Crumbs complex in viral assembly and budding are warranted. The host factors and processes involved in these important steps in the viral lifecycle are not fully understood. Segment specific signals seem to promote the co-localization of some vRNPs in the cytoplasm prior to packaging into progeny virions as a supramolecular complex (Chou et al. 2013; Fujii et al. 2003; Gerber et al. 2014; Lakdawala et al. 2014; Muramoto et al. 2006). However it is not known where the final assembly of the vRNPs takes place or how they reach the plasma membrane from recycling endosomes (Eisfeld, Neumann, and Kawaoka 2014). Host factors involved in influenza budding also remain largely unknown (Gorai et al. 2012). Inter-vRNP interactions and the host factors that facilitate assembly and budding bear directly on the ability of two different strains to reassort—a process that can lead to influenza pandemics (discussed in 1.1.1). A comprehensive understanding of these aspects of the viral lifecycle could lead to new therapeutic targets and better pandemic preparedness.

The results presented in this thesis suggest several avenues for follow up studies. Targeted co-immunoprecipitation (co-IP) experiments to identify which components of the tight junction complex are interacting directly with NEP would be a logical starting point. Finer mapping of the interactions and phenotypes to specific residues or motifs on NEP would inform model construction and facilitate sequence-based comparisons across influenza strains. In a complementary fashion, mapping the interaction sites on the host factors would help inform follow on functional studies and mechanistic hypotheses about the role of the interaction in the viral lifecycle. Immunofluorescence experiments showing the localization of NEP with respect to the Crumbs complex and AMOT during infection would provide additional information about the relevance of these interactions to specific stages of the viral lifecycle. In parallel, functional studies should also be performed to better understand the role of the tight junction complex during influenza replication. Presently it is not clear if disrupted tight junctions are pro-viral, and establishing whether or not barrier integrity has an impact on specific stages of the life cycle would benefit the entire field. Additional functional studies on the tight junction

components specifically identified in this work and ideally validated through co-IP, should also be performed to assess their relevance to viral replication.

As a pathogen that enters, replicates in, and buds from epithelial cells, it is logical that influenza has evolved to interact with one of the defining features of these cells—the tight junction complex. A growing body of evidence implicates tight junctions in the regulation of a range of cellular processes and many other bacterial and viral pathogens interact with tight junction components. Future studies based on the work presented in this thesis and discussed in this section would help define the role tight junction proteins play in the influenza lifecycle potentially leading to new therapeutic strategies and new insights into fundamental biological processes.

As a starting point for host-pathogen interaction studies on influenza NEP, we engineered affinity tag sequences into the viral genome. We used the recombinant virus in proteomics experiments to identify host factors that interact with NEP during infection. Functional enrichment analysis of the proteomics data showed an overrepresentation of tight junction proteins and we hypothesized that NEP alters tight junction properties. We tested for an effect of NEP expression on tight junction functioning in an *in vitro* cell culture model and showed that NEP disrupts tight junctions. In the end, these findings suggest a new role for NEP during viral replication as a tight junction modulator.

Biliography

Aggarwal, Shilpa, Stephen Dewhurst, Toru Takimoto, and Baek Kim. 2011. “Biochemical Impact of the Host Adaptation-Associated PB2 E627K Mutation on the Temperature-Dependent RNA Synthesis Kinetics of Influenza A Virus Polymerase Complex.” *Journal of Biological Chemistry* 286 (40): 34504–13.

Akarsu, Hatice, Wilhelm P Burmeister, Carlo Petosa, Isabelle Petit, Christoph W Müller, Rob W H Ruigrok, and Florence Baudin. 2003. “Crystal Structure of the M1 Protein-Binding Domain of the Influenza A Virus Nuclear Export Protein (NEP/NS2).” *The EMBO journal* 22(18): 4646–55.

Akerström, B, T Brodin, K Reis, and L Björck. 1985. “Protein G: A Powerful Tool for Binding and Detection of Monoclonal and Polyclonal Antibodies.” *Journal of immunology (Baltimore, Md. : 1950)* 135(4): 2589–92.

Alber, Frank, Svetlana Dokudovskaya, Liesbeth M Veenhoff, Wenzhu Zhang, Julia Kipper, Damien Devos, Adisetyantari Suprapto, Orit Karni-Schmidt, Rosemary Williams, Brian T Chait, Andrej Sali, and Michael P Rout. 2007. “The Molecular Architecture of the Nuclear Pore Complex.” *Nature* 450(7170): 695–701.

Alexander, Dennis J. 2007. “An Overview of the Epidemiology of Avian Influenza.” *Vaccine* 25(30 SPEC. ISS.): 5637–44.

Allison, Andrew B, Jennifer R Ballard, Robert B Tesh, Justin D Brown, Mark G Ruder, M Kevin Keel, Brandon A Munk, Randall M Mickley, Samantha E J Gibbs, Amelia P A Travassos da Rosa, Julie C Ellis, Hon S Ip, Valerie I Shearn-Bochsler, Matthew B Rogers, Elodie Ghedin, Edward C Holmes, Colin R Parrish, and Chris Dwyer. 2015. “Cyclic Avian Mass Mortality in the Northeastern United States Is Associated with a Novel Orthomyxovirus.” *Journal of Virology* 89 (2): 1389–1403.

Alvord, Gregory, Jean Roayaei, Robert Stephens, Michael W Baseler, H Clifford Lane, and Richard A Lempicki. 2007. “The DAVID Gene Functional Classification Tool: A Novel Biological Module-Centric Algorithm to Functionally Analyze Large Gene Lists.” *Genome Biol* 8(9): 183.

Amasheh, Salah, Noga Meiri, Alfred H Gitter, Torsten Schöneberg, Joachim Mankertz, Jörg D Schulzke, and Michael Fromm. 2002. “Claudin-2 Expression Induces Cation-Selective Channels in Tight Junctions of Epithelial Cells.” *Journal of Cell Science* 115(24): 4969–76.

André, I, J L Putaux, H Chanzy, F R Taravel, J W Timmermans, and D de Wit. 1996. “Single Crystals of Inulin.” *International Journal of Biological Macromolecules* 18(3): 195–204.

Arcangeletti, M C, F Pinardi, and S Missorini. 1997. “Modification of Cytoskeleton and Prosome Networks in Relation to Protein Synthesis in Influenza A Virus-Infected LLC-MK2 Cells.” *Virus Research* 51: 19–34.

Avalos, R T, Z Yu, and D P Nayak. 1997. “Association of Influenza Virus NP and M1 Proteins with Cellular Cytoskeletal Elements in Influenza Virus-Infected Cells.” *Journal of virology* 71(4): 2947–58.

Bach, Ingolf. 2000. “The LIM Domain: Regulation by Association.” *Mechanisms of Development* 91(1–2): 5–17.

Baicu, Simona C, and Michael J Taylor. 2002. “Acid–base Buffering in Organ Preservation Solutions as a Function of Temperature: New Parameters for Comparing Buffer Capacity and Efficiency.” *Cryobiology* 45(1): 33–48.

Balda, M. 2003. “Epithelial Cell Adhesion and the Regulation of Gene Expression.” *Trends in Cell Biology* 13(6): 310–18.

Balda, M S, J A Whitney, C Flores, S González, M Cereijido, and K Matter. 1996. “Functional Dissociation of Paracellular Permeability and Transepithelial Electrical Resistance and Disruption of the Apical-Basolateral Intramembrane Diffusion Barrier by Expression of a Mutant Tight Junction Membrane Protein.” *The Journal of Cell Biology* 134 (4): 1031–49.

Balda, Maria S, Michelle D Garrett, and Karl Matter. 2003. “The ZO-1–associated Y-Box Factor ZONAB Regulates Epithelial Cell Proliferation and Cell Density.” *The Journal of Cell Biology* 160 (3): 423–32.

Balda, Maria S, and Karl Matter. 2000. “The Tight Junction Protein ZO-1 and an Interacting Transcription Factor Regulate ErbB-2 Expression.” *The EMBO Journal* 19(9): 2024–33.

Balda, Maria S, and Karl Matter. 2009. “Tight Junctions and the Regulation of Gene Expression.” *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1788(4): 761–67.

Banks, J, E S Speidel, E Moore, L Plowright, A Piccirillo, I Capua, P Cordioli, A Fioretti, and D J Alexander. 2001. “Changes in the Haemagglutinin and the Neuraminidase Genes prior to the Emergence of Highly Pathogenic H7N1 Avian Influenza Viruses in Italy.” *Archives of Virology* 146(5): 963–73.

Bantscheff, Marcus, Simone Lemeer, Mikhail M. Savitski, and Bernhard Kuster. 2012. “Quantitative Mass Spectrometry in Proteomics: Critical Review Update from 2007 to the Present.” *Analytical and Bioanalytical Chemistry* 404(4): 939–65.

Barton, Erik S, J Craig Forrest, Jodi L Connolly, James D Chappell, Yuan Liu, Frederick J Schnell, Asma Nusrat, Charles A Parkos, and Terence S Dermody. 2001. “Junction Adhesion Molecule Is a Receptor for Reovirus.” *Cell* 104(3): 441–51.

Battle, A., Z. Khan, S. H. Wang, A. Mitrano, M. J. Ford, Jonathan K Pritchard, and Yoav Gilad. 2015. “Impact of Regulatory Variation from RNA to Protein.” *Science* 347(6222): 664–67.

Beare, A.S, G.C Schild, and J.W Craig. 1975. “TRIALS IN MAN WITH LIVE RECOMBINANTS MADE FROM A/PR/8/34 (H0 N1) AND WILD H3 N2 INFLUENZA VIRUSES.” *The Lancet* 306(7938): 729–32.

Bergelson, Jeffrey M, Jennifer A Cunningham, Gustavo Droguett, Evelyn A Kurt-Jones, Anita Krithivas, Jeong S Hong, Marshall S Horwitz, Richard L Crowell, and Robert W Finberg. 1997. “Isolation of a Common Receptor for Coxsackie B Viruses and Adenoviruses 2 and 5.” *Science* 275(5304): 1320–23.

Biegeleisen Jr., Joseph Z, Malcolm S Mitchell, Betsy B Marcus, Dwane L Rhoden, and Richard W Blumberg. 1966. “Immunofluorescence Techniques for Demonstrating Bacterial Pathogens Associated with Cerebrospinal Meningitis. I. Clinical Evaluation of Conjugates on Smears Prepared Directly from Cerebrospinal Fluid Sediments.” *The Journal of Laboratory and Clinical Medicine* 65(6): 976–89.

Biggar, R J, M Melbye, P Ebbesen, S Alexander, J O Nielsen, P Sarin, and V Faber. 1985. “Variation in Human T Lymphotropic Virus III (HTLV-III) Antibodies in Homosexual Men: Decline before Onset of Illness Related to Acquired Immune Deficiency Syndrome (AIDS).” *British medical journal (Clinical research ed.)* 291(6501): 997–98.

Bortz, Eric, Liset Westera, Jad Maamary, and J Steel. 2011. “Host-and Strain-Specific Regulation of Influenza Virus Polymerase Activity by Interacting Cellular Proteins.” *MBio*.

Boyle, Erin C, Nat F Brown, and B Brett Finlay. 2006. “Salmonella Enterica Serovar Typhimurium Effectors SopB, SopE, SopE2 and SipA Disrupt Tight Junction Structure and Function.” *Cellular microbiology* 8(12): 1946–57.

Brass, Abraham L, I C Huang, Y Benita, S P John, M N Krishnan, E M Feeley, B J Ryan, J L Weyer, L van der Weyden, E Fikrig, D J Adams, R J Xavier, M Farzan, and S J Elledge. 2009. “The IFITM Proteins Mediate Cellular Resistance to Influenza A H1N1 Virus, West Nile Virus, and Dengue Virus.” *Cell* 139(7): 1243–54.

Bratt, Anders, Olivier Birot, Indranil Sinha, Niina Veitonmäki, Karin Aase, Mira Ernkvist, and Lars Holmgren. 2005. “Angiomotin Regulates Endothelial Cell-Cell Junctions and Cell Motility.” *Journal of Biological Chemistry* 280(41): 34859–69.

Le Breton, Marc, Laurène Meyniel-Schicklin, Alexandre Deloire, Bruno Coutard, Bruno Canard, Xavier de Lamballerie, Patrice Andre, Chantal Rabourdin-Combe, Vincent Lotteau, and Nathalie Davoust. 2011. “Flavivirus NS3 and NS5 Proteins Interaction Network: A High-Throughput Yeast Two-Hybrid Screen.” *BMC Microbiology* 11(1): 234.

Brizzard, Bill L, R G Chubet, and D L Vizard. 1994. “Immunoaffinity Purification of FLAG Epitope-Tagged Bacterial Alkaline Phosphatase Using a Novel Monoclonal Antibody and Peptide Elution.” *Biotechniques* 16(4): 730–35.

Brown, Joseph N, Robert E Palermo, Carole R Baskin, Marina Gritsenko, Patrick J Sabourin, James P Long, Carol L Sabourin, Helle Bielefeldt-Ohmann, Adolfo García-Sastre, Randy Albrecht, Terrence M Tumpey, Jon M Jacobs, Richard D Smith, and Michael G Katze. 2010. “Macaque Proteome Response to Highly Pathogenic Avian Influenza and 1918 Reassortant Influenza Virus Infections.” *Journal of Virology* 84 (22): 12058–68.

Brundage, John F, and G Dennis Shanks. 2008. “Deaths from Bacterial Pneumonia during 1918–19 Influenza Pandemic.” *Emerging infectious diseases* 14(8): 1193.

Brunner, Erich, Christian H Ahrens, Sonali Mohanty, Hansruedi Baetschmann, Sandra Loevenich, Frank Potthast, Eric W Deutsch, Christian Panse, Ulrik de Lichtenberg, Oliver Rinner, Hookeun Lee, Patrick G A Pedrioli, Johan Malmstrom, Katja Koehler, Sabine Schrimpf, Jeroen Krijgsveld, Floyd Kregenow, Albert J R Heck, Ernst Hafen, Ralph Schlapbach, and Ruedi Aebersold. 2007. “A High-Quality Catalog of the Drosophila Melanogaster Proteome.” *Nat Biotech* 25(5): 576–83.

Brunotte, L., J. Flies, H. Bolte, P. Reuther, F. Vreede, and M. Schwemmle. 2014. “The Nuclear Export Protein of H5N1 Influenza A Viruses Recruits Matrix 1 (M1) Protein to the Viral Ribonucleoprotein to Mediate Nuclear Export.” *Journal of Biological Chemistry* 289(29): 20067–77.

Bui, M, E G Wills, a Helenius, and G R Whittaker. 2000. “Role of the Influenza Virus M1 Protein in Nuclear Export of Viral Ribonucleoproteins.” *Journal of virology* 74(4): 1781–86.

Bulgakova, Natalia A, and Elisabeth Knust. 2009. “The Crumbs Complex: From Epithelial-Cell Polarity to Retinal Degeneration.” *Journal of Cell Science* 122(15): 2587–96.

Bullido, Rosario, P Gomez-Puertas, and MJ Saiz. 2001. “Influenza A Virus NEP (NS2 Protein) Downregulates RNA Synthesis of Model Template RNAs.” *Journal of* 75(10): 4912–17.

Cann, D C, and Margaret E Willox. 1965. “Analysis of Multimolecular Enzymes as an Aid to the Identification of Certain Rapidly Growing Mycobacteria, Using Starch Gel Electrophoresis.” *Journal of Applied Bacteriology* 28(1): 165–73.

Capaldo, Christopher T, Attila E Farkas, Roland S Hilgarth, Susanne M Krug, Mattie F Wolf, Jeremy K Benedik, Michael Fromm, Michael Koval, Charles Parkos, and Asma Nusrat. 2014. “Proinflammatory Cytokine-Induced Tight Junction Remodeling through Dynamic Self-Assembly of Claudins.” *Molecular biology of the cell* 25(18): 2710–19.

Caron, Alexandre, Nicolas Gaidet, Michel de Garine-Wichatitsky, Serge Morand, and Elissa Z Cameron. 2009. “Evolutionary Biology, Community Ecology and Avian Influenza Research.” *Infection, Genetics and Evolution* 9(2): 298–303.

Carpp, Lindsay N., Richard S. Rogers, Robert L. Moritz, and John D. Aitchison. 2014. “Quantitative Proteomic Analysis of Host-Virus Interactions Reveals a Role for Golgi Brefeldin A Resistance Factor 1 (GBF1) in Dengue Infection.” *Molecular & Cellular Proteomics* 13(11): 2836–54.

Cereijido, M, E S Robbins, W J Dolan, C A Rotunno, and D D Sabatini. 1978. “Polarized Monolayers Formed by Epithelial Cells on a Permeable and Translucent Support.” *The Journal of cell biology* 77(3): 853–80.

Chaemchuen, Suwittra, Suang Rungpragayphan, Yong Poovorawan, and Kanitha Patarakul. 2011. “Identification of Candidate Host Proteins That Interact with LipL32, the Major Outer Membrane Protein of Pathogenic Leptospira, by Random Phage Display Peptide Library.” *Veterinary Microbiology* 153(1-2): 178–85.

Chase, Geoffrey P, Marie-Anne Rameix-Welti, Aurelija Zvirbliene, Gintautas Zvirblis, Veronika Götz, Thorsten Wolff, Nadia Naffakh, and Martin Schwemmle. 2011. “Influenza Virus Ribonucleoprotein Complexes Gain Preferential Access to Cellular Export Machinery through Chromatin Targeting.” *PLoS pathogens* 7(9): e1002187.

de Chassey, B, V Navratil, L Tafforeau, M S Hiet, a Aublin-Gex, S Agaugué, G Meiffren, F Pradezynski, B F Faria, T Chantier, M Le Breton, J Pellet, N Davoust, P E Mangeot, a Chaboud, F Penin, Y Jacob, P O Vidalain, M Vidal, P André, C Rabourdin-Combe, and V Lotteau. 2008. “Hepatitis C Virus Infection Protein Network.” *Molecular systems biology* 4: 230.

de Chassey, Benoît, Anne Aublin-Gex, Alessia Ruggieri, Laurène Meyniel-Schicklin, Fabrine Pradezynski, Nathalie Davoust, Thibault Chantier, Lionel Tafforeau, Philippe-Emmanuel Mangeot, Claire Ciancia, Laure Perrin-Cocon, Ralf Bartenschlager, Patrice André, and Vincent Lotteau. 2013. “The Interactomes of Influenza Virus NS1 and NS2 Proteins Identify New Host Factors and Provide Insights for ADAR1 Playing a Supportive Role in Virus Replication.” *PLoS pathogens* 9(7): e1003440.

Chen, Benjamin J, George P Leser, David Jackson, and Robert a Lamb. 2008. “The Influenza Virus M2 Protein Cytoplasmic Tail Interacts with the M1 Protein and Influences Virus Assembly at the Site of Virus Budding.” *Journal of virology* 82(20): 10059–70.

Chen, Jingjing, Shengping Huang, and Ze Chen. 2010. “Human Cellular Protein Nucleoporin hNup98 Interacts with Influenza A Virus NS2/nuclear Export Protein and Overexpression of Its GLFG Repeat Domain Can Inhibit Virus Propagation.” *The Journal of general virology* 91(Pt 10): 2474–84.

Choi, Hyungwon, Brett Larsen, Zhen-Yuan Lin, Ashton Breitkreutz, Dattatreya Mellacheruvu, Damian Fermin, Zhaohui S Qin, Mike Tyers, Anne-Claude Gingras, and Alexey I Nesvizhskii. 2011. “SAINT: Probabilistic Scoring of Affinity Purification-Mass Spectrometry Data.” *Nature methods* 8(1): 70–73.

Chou, Yi-ying, Nicholas S Heaton, Qinshan Gao, Peter Palese, Robert Singer, and Timothée Lionnet. 2013. “Colocalization of Different Influenza Viral RNA Segments in the Cytoplasm before Viral Budding as Shown by Single-Molecule Sensitivity FISH Analysis.” *PLoS Pathog* 9(5): e1003358.

Chua, Mark a., Sonja Schmid, Jasmine T. Perez, Ryan a. Langlois, and Benjamin R. tenOever. 2013. “Influenza A Virus Utilizes Suboptimal Splicing to Coordinate the Timing of Infection.” *Cell Reports* 3(1): 23–29.

Chubet, Richard G, and Bill L Brizzard. 1996. “Vectors for Expression and Secretion of FLAG Epitope-Tagged Proteins in Mammalian Cells.” *Biotechniques* 20(1): 136–41.

Coons, Albert H, H J Creech, R N Jones, and E Berliner. 1942. “The Demonstration of Pneumococcal Antigen in Tissues by the Use of Fluorescent Antibody.” *J. Immunol* 45(3): 159–70.

Coons, Albert H, and Melvin H Kaplan. 1950. “Localization of Antigen in Tissue Cells II. Improvements in a Method for the Detection of Antigen by Means of Fluorescent Antibody.” *The Journal of experimental medicine* 91(1): 1–13.

Cox, Jürgen, and Matthias Mann. 2008. “MaxQuant Enables High Peptide Identification Rates, Individualized P.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification.” *Nature biotechnology* 26(12): 1367–72.

Cox, Jürgen, Ivan Matic, Maximiliane Hilger, Nagarjuna Nagaraj, Matthias Selbach, Jesper V Olsen, and Matthias Mann. 2009. “A Practical Guide to the MaxQuant Computational Platform for SILAC-Based Quantitative Proteomics.” *Nature protocols* 4(5): 698–705.

Cravatt, Benjamin F, Gabriel M Simon, and John R Yates. 2007. “The Biological Impact of Mass-Spectrometry-Based Proteomics.” *Nature* 450(7172): 991–1000.

Demirov, Dimiter, Gülsah Gabriel, Carola Schneider, Heinrich Hohenberg, and Stephan Ludwig. 2012. “Interaction of Influenza A Virus Matrix Protein with RACK1 Is Required for Virus Release.” *Cellular microbiology* 14(5): 774–89.

Deng, Tao, Othmar G Engelhardt, Benjamin Thomas, Alexandre V Akoulitchev, George G Brownlee, and Ervin Fodor. 2006. “Role of Ran Binding Protein 5 in Nuclear Import and Assembly of the Influenza Virus RNA Polymerase Complex.” *Journal of virology* 80(24): 11911–19.

Dennis, Glynn, Brad T Sherman, Douglas A Hosack, Jun Yang, Wei Gao, H C Lane, and Richard A Lempicki. 2003. “DAVID: Database for Annotation, Visualization, and Integrated Discovery.” *Genome Biology* 4(9): R60.

Desselberger, U, K Nakajima, P Alfino, F S Pedersen, W A Haseltine, C Hannoun, and P Palese. 1978. “Biochemical Evidence That ‘New’ Influenza Virus Strains in Nature May Arise by Recombination (reassortment).” *Proceedings of the National Academy of Sciences* 75 (7): 3341–45.

Dharmsathaphorn, Kiertisin, J A McRoberts, K G Mandel, L D Tisdale, and H Masui. 1984. “A Human Colonic Tumor Cell Line That Maintains Vectorial Electrolyte Transport.” *American Journal of Physiology-Gastrointestinal and Liver Physiology* 246(2): G204–8.

Diamond, J M. 1977. “Twenty-First Bowditch Lecture. The Epithelial Junction: Bridge, Gate, and Fence.” *The Physiologist* 20(1): 10.

Donnelly, M L, G Luke, a Mehrotra, X Li, L E Hughes, D Gani, and M D Ryan. 2001. “Analysis of the Aphthovirus 2A/2B Polyprotein ‘Cleavage’ Mechanism Indicates Not a Proteolytic Reaction, but a Novel Translational Effect: A Putative Ribosomal ‘Skip’.” *The Journal of general virology* 82(Pt 5): 1013–25.

Ducatez, M.F., R.G. Webster, and R.J. Webby. 2008. “Animal Influenza Epidemiology.” *Vaccine* 26: D67–69.

Dyer, Matthew D., T. M. Murali, and Bruno W. Sobral. 2008. “The Landscape of Human Proteins Interacting with Viruses and Other Pathogens.” *PLoS Pathogens* 4(2).

Einhauer, A, and A. Jungbauer. 2001a. “Affinity of the Monoclonal Antibody M1 Directed against the FLAG Peptide.” *Journal of Chromatography A* 921(1): 25–30.

Einhauer, A, and A. Jungbauer. 2001b. “The FLAG Peptide, a Versatile Fusion Tag for the Purification of Recombinant Proteins.” *Journal of Biochemical and Biophysical Methods* 49(1-3): 455–65.

Eisfeld, Amie J, Eiryo Kawakami, Tokiko Watanabe, Gabriele Neumann, and Yoshihiro Kawaoka. 2011. “RAB11A Is Essential for Transport of the Influenza Virus Genome to the Plasma Membrane.” *Journal of virology* 85(13): 6117–26.

Eisfeld, Amie J, Gabriele Neumann, and Yoshihiro Kawaoka. 2014. “At the Centre: Influenza A Virus Ribonucleoproteins.” *Nature reviews. Microbiology* 13(1): 28–41.

Ellis, Trevor M, R Barry Bousfield, Lucy A Bissett, Kitman C Dyrting, Geraldine S M Luk, S T Tsim, Katharine Sturm-ramirez, Robert G Webster, Yi Guan, and J S Malik Peiris. 2004. “Investigation of Outbreaks of Highly Pathogenic H5N1 Avian Influenza in Waterfowl and Wild Birds in Hong Kong in Late 2002.” *Avian Pathology* 33(5): 492–505.

Elton, Debra, M. Simpson-Holley, Kate Archer, L. Medcalf, Roger Hallam, J. McCauley, and Paul Digard. 2001. “Interaction of the Influenza Virus Nucleoprotein with the Cellular CRM1-Mediated Nuclear Export Pathway.” *Journal of virology* 75(1): 408.

Evans, M J, and D W Kingsbury. 1969. “Separation of Newcastle Disease Virus Proteins by Polyacrylamide Gel Electrophoresis.” *Virology* 37(4): 597–604.

Evans, Matthew J, Thomas von Hahn, Donna M Tscherne, Andrew J Syder, Maryline Panis, Benno Wölk, Theodora Hatziioannou, Jane A McKeating, Paul D Bieniasz, and Charles M Rice. 2007. “Claudin-1 Is a Hepatitis C Virus Co-Receptor Required for a Late Step in Entry.” *Nature* 446(7137): 801–5.

Fabrini, Raffaele, Anastasia De Luca, Lorenzo Stella, Giampiero Mei, Barbara Orioni, Sarah Ciccone, Giorgio Federici, Mario Lo Bello, and Giorgio Ricci. 2009. “Monomer−Dimer Equilibrium in Glutathione Transferases: A Critical Re-Examination.” *Biochemistry* 48(43): 10473–82.

Fang, Jianmin, Jing-Jing Qian, Saili Yi, Thomas C Harding, Guang Huan Tu, Melinda VanRoey, and Karin Jooss. 2005. “Stable Antibody Expression at Therapeutic Levels Using the 2A Peptide.” *Nature biotechnology* 23(5): 584–90.

Fanning, Alan S, Brian J Jameson, Lynne A Jesaitis, and James Melvin Anderson. 1998. “The Tight Junction Protein ZO-1 Establishes a Link between the Transmembrane Protein Occludin and the Actin Cytoskeleton.” *Journal of Biological Chemistry* 273(45): 29745–53.

Farquhar, Marilyn G, and George E Palade. 1963. “JUNCTIONAL COMPLEXES IN VARIOUS EPITHELIA.” *The Journal of Cell Biology* 17 (2): 375–412.

Fawcett, D W. 1958. “Structural Specializations of the Cell Surface.” *Frontiers in Cytology,(SL Palay, editor), New Haven, Yale University Press* 19.

Fenn, J B, M Mann, C K Meng, S F Wong, and C M Whitehouse. 1989. “Electrospray Ionization for Mass Spectrometry of Large Biomolecules.” *Science* 246 (4926): 64–71.

Fields, S, and O Song. 1989. “A Novel Genetic System to Detect Protein-Protein Interactions.” *Nature* 340(6230): 245–46.

Fletcher, Sarah J, Mudassar Iqbal, Sara Jabbari, Dov Stekel, and Joshua Z Rappoport. 2014. “Analysis of Occludin Trafficking, Demonstrating Continuous Endocytosis, Degradation, Recycling and Biosynthetic Secretory Trafficking.” *PLoS ONE* 9(11): e111176.

Fodor, Ervin, Louise Devenish, Othmar G Engelhardt, Peter Palese, George G Brownlee, and Adolfo García-Sastre. 1999. “Rescue of Influenza A Virus from Recombinant DNA.” *Journal of virology* 73(11): 9679–82.

Folin, Otto, and Vintila Ciocalteau. 1927. “Tyrosine and Tryptophane in Proteins.” *Journal of Biological Chemistry* 73(2): 627–48.

Fowler, R C, D W Coble, N C Kramer, and T M Brown. 1963. “Starch Gel Electrophoresis of a Fraction of Certain of the Pleuropneumonia-Like Group of Microorganisms.” *Journal of bacteriology* 86(Table 1): 1145–51.

Franceschini, Andrea, Damian Szklarczyk, Sune Frankild, Michael Kuhn, Milan Simonovic, Alexander Roth, Jianyi Lin, Pablo Minguez, Peer Bork, Christian von Mering, and Lars J Jensen. 2013. “STRING v9.1: Protein-Protein Interaction Networks, with Increased Coverage and Integration.” *Nucleic Acids Research* 41 (D1): D808–15.

Franzosa, Eric a, and Yu Xia. 2011. “Structural Principles within the Human-Virus Protein-Protein Interaction Network.” *Proceedings of the National Academy of Sciences of the United States of America* 108(26): 10538–43.

Fromter, E, and J Diamond. 1972. “Route of Passive Ion Permeation in Epithelia.” *Nature new biol* 235(9): 13.

Fujii, Ken, Yutaka Fujii, Takeshi Noda, Yukiko Muramoto, Tokiko Watanabe, Ayato Takada, Hideo Goto, Taisuke Horimoto, and Yoshihiro Kawaoka. 2005. “Importance of Both the Coding and the Segment-Specific Noncoding Regions of the Influenza A Virus NS Segment for Its Efficient Incorporation into Virions.” *Journal of virology* 79(6): 3766–74.

Fujii, Yutaka, Hideo Goto, Tokiko Watanabe, Tetsuya Yoshida, and Yoshihiro Kawaoka. 2003. “Selective Incorporation of Influenza Virus RNA Segments into Virions.” *Proceedings of the National Academy of Sciences of the United States of America* 100(4): 2002–7.

Fujita, Michiya, Koichi Kawai, Shigetaka Asano, and Makoto Nakao. 1973. “Protein Components of Two Different Regions of an Intestinal Epithelial Cell Membrane: Regional Singularities.” *Biochimica et Biophysica Acta (BBA) - Biomembranes* 307(1): 141–51.

Furuse, Mikio, Kyoko Furuse, Hiroyuki Sasaki, and Shoichiro Tsukita. 2001. “Conversion of Zonulae Occludentes from Tight to Leaky Strand Type by Introducing Claudin-2 into Madin-Darby Canine Kidney I Cells.” *The Journal of Cell Biology* 153 (2): 263–72.

Furuse, Mikio, Tetsuaki Hirase, Masahiko Itoh, Akira Nagafuchi, Shigenobu Yonemura, and Sachiko Tsukita. 1993. “Occludin: A Novel Integral Membrane Protein Localizing at Tight Junctions.” *The Journal of cell biology* 123(6): 1777–88.

Galli, P, A Brenna, P De Camilli, and J Meldolesi. 1976. “Extracellular Calcium and the Organization of Tight Junctions in Pancreatic Acinar Cells.” *Experimental Cell Research* 99(1): 178–83.

Gao, Rongbao, Bin Cao, Yunwen Hu, Zijian Feng, Dayan Wang, Wanfu Hu, Jian Chen, Zhijun Jie, Haibo Qiu, Ke Xu, Xuewei Xu, Hongzhou Lu, Wenfei Zhu, Zhancheng Gao, Nijuan Xiang, Yinzhong Shen, Zebao He, Yong Gu, Zhiyong Zhang, Yi Yang, Xiang Zhao, Lei Zhou, Xiaodan Li, Shumei Zou, Ye Zhang, Xiyan Li, Lei Yang, Junfeng Guo, Jie Dong, Qun Li, Libo Dong, Yun Zhu, Tian Bai, Shiwen Wang, Pei Hao, Weizhong Yang, Yanping Zhang, Jun Han, Hongjie Yu, Dexin Li, George F Gao, Guizhen Wu, Yu Wang, Zhenghong Yuan, and Yuelong Shu. 2013. “Human Infection with a Novel Avian-Origin Influenza A (H7N9) Virus.” *New England Journal of Medicine* 368(20): 1888–97.

Gao, Shengyan, Shanshan Wang, Shuai Cao, Lei Sun, Jing Li, Yuhai Bi, George F Gao, and Wenjun Liu. 2014. “Characteristics of Nucleocytoplasmic Transport of H1N1 Influenza A Virus Nuclear Export Protein.” *Journal of virology* 88(13): 7455–63.

Gao, Shijuan, Jiaoxiang Wu, Ran-yi Liu, Jiandong Li, Liping Song, Yan Teng, Chunjie Sheng, Dong Liu, Chen Yao, Huiming Chen, Wei Jiang, Shuai Chen, and Wenlin Huang. 2015. “Interaction of NS2 with AIMP2 Facilitates

the Switch from Ubiquitination to SUMOylation of M1 in Influenza A Virus-Infected." *Journal of Virology* 89(1): 300–311.

Garcia-Robles, Inmaculada, Hatice Akarsu, Christoph W Müller, Rob W H Ruigrok, and Florence Baudin. 2005. "Interaction of Influenza Virus Proteins with Nucleosomes." *Virology* 332(1): 329–36.

Gatherer, Derek. 2009. "The 2009 H1N1 Influenza Outbreak in Its Historical Context." *Journal of Clinical Virology* 45(3): 174–78.

Gavin, Anne-Claude, Patrick Aloy, Paola Grandi, Roland Krause, Markus Boesche, Martina Marzioch, Christina Rau, Lars Juhl Jensen, Sonja Bastuck, Birgit Dümpelfeld, Angela Edelmann, Marie-Anne Heurtier, Verena Hoffman, Christian Hoefert, Karin Klein, Manuela Hudak, Anne-Marie Michon, Malgorzata Schelder, Markus Schirle, Marita Remor, Tatjana Rudi, Sean Hooper, Andreas Bauer, Tewis Bouwmeester, Georg Casari, Gerard Drewes, Gitte Neubauer, Jens M Rick, Bernhard Kuster, Peer Bork, Robert B Russell, and Giulio Superti-Furga. 2006. "Proteome Survey Reveals Modularity of the Yeast Cell Machinery." *Nature* 440(7084): 631–36.

Geiger, Tamar, Anja Wehner, Christoph Schaab, Juergen Cox, and Matthias Mann. 2012. "Comparative Proteomic Analysis of Eleven Common Cell Lines Reveals Ubiquitous but Varying Expression of Most Proteins." *Molecular & Cellular Proteomics* 11 (3).

Geli, V, D Baty, and C Lazdunski. 1988. "Use of a Foreign Epitope as a 'Tag' for the Localization of Minor Proteins within a Cell: The Case of the Immunity Protein to Colicin A." *Proceedings of the National Academy of Sciences of the United States of America* 85(February): 689–93.

Gerber, Marie, Catherine Isel, Vincent Moules, and Roland Marquet. 2014. "Selective Packaging of the Influenza A Genome and Consequences for Genetic Reassortment." *Trends in Microbiology* 22(8): 446–55.

Gilchrist, Mark, Vesteinn Thorsson, Bin Li, Alistair G Rust, Martin Korb, Jared C Roach, Kathleen Kennedy, Tsonwin Hai, Hamid Bolouri, and Alan Aderem. 2006. "Systems Biology Approaches Identify ATF3 as a Negative Regulator of Toll-like Receptor 4." *Nature* 441(7090): 173–78.

Gog, Julia R, Emmanuel Dos Santos Afonso, Rosa M Dalton, India Leclercq, Laurence Tiley, Debra Elton, Johann C von Kirchbach, Nadia Naffakh, Nicolas Escriou, and Paul Digard. 2007. "Codon Conservation in the Influenza A Virus Genome Defines RNA Packaging Signals." *Nucleic acids research* 35(6): 1897–1907.

Goldberg, Debra S, and Frederick P Roth. 2003. "Assessing Experimentally Derived Interactions in a Small World." *Proceedings of the National Academy of Sciences* 100 (8): 4372–76.

Golebiewski, L., H. Liu, R. T. Javier, and a. P. Rice. 2011. "The Avian Influenza Virus NS1 ESEV PDZ Binding Motif Associates with Dlg1 and Scribble To Disrupt Cellular Tight Junctions." *Journal of Virology* 85(20): 10639–48.

González-Mariscal, Lorenza, Alaide Domínguez-Calderón, Arturo Raya-Sandino, José Mario Ortega-Olvera, Orlando Vargas-Sierra, and Gabriela Martínez-Revollar. 2014. "Tight Junctions and the Regulation of Gene Expression." *Seminars in Cell & Developmental Biology* 36: 213–23.

Good, Norman E, G Douglas Winget, Wilhelmina Winter, Thomas N Connolly, Seikichi Izawa, and Raizada M M Singh. 1966. "Hydrogen Ion Buffers for Biological Research*." *Biochemistry* 5(2): 467–77.

Gorai, Takeo, Hideo Goto, Takeshi Noda, Tokiko Watanabe, Hiroko Kozuka-Hata, Masaaki Oyama, Ryo Takano, Gabriele Neumann, Shinji Watanabe, and Yoshihiro Kawaoka. 2012. "F1Fo-ATPase, F-Type Proton-Translocating ATPase, at the Plasma Membrane Is Critical for Efficient Influenza Virus Budding." *Proceedings of the National Academy of Sciences of the United States of America* 109(12): 4615–20.

Grindstaff, Kent K, Charles Yeaman, Niroshana Anandasabapathy, Shu-Chan Hsu, Enrique Rodriguez-Boulan, Richard H Scheller, and W.James Nelson. 1998. "Sec6/8 Complex Is Recruited to Cell–Cell Contacts and Specifies Transport Vesicle Delivery to the Basal-Lateral Membrane in Epithelial Cells." *Cell* 93(5): 731–40.

Gulig, P. a., and E. J. Hansen. 1985. "Coprecipitation of Lipopolysaccharide and the 39,000-Molecular-Weight Major Outer Membrane Protein of Haemophilus Influenzae Type B by Lipopolysaccharide-Directed Monoclonal Antibody." *Infection and Immunity* 49(3): 819–27.

Guttman, Julian a., and B. Brett Finlay. 2009. “Tight Junctions as Targets of Infectious Agents.” *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1788(4): 832–41.

Gygi, S P, B Rist, S a Gerber, F Turecek, M H Gelb, and R Aebersold. 1999. “Quantitative Analysis of Complex Protein Mixtures Using Isotope-Coded Affinity Tags.” *Nature biotechnology* 17(10): 994–99.

Hama, Kiyoshi. 1960. “The Fine Structure of the Desmosomes in Frog Mesothelium.” *The Journal of biophysical and biochemical cytology* 7(3): 575–77.

Han, Xueqing, Zhihui Li, Hongjun Chen, Huiyu Wang, Lin Mei, Shaoqiang Wu, Tianyi Zhang, Bohua Liu, and Xiangmei Lin. 2012. “Influenza Virus A/Beijing/501/2009 (H1N1) NS1 Interacts with β-Tubulin and Induces Disruption of the Microtubule Network and Apoptosis on A549 Cells.”

Hao, Linhui, Akira Sakurai, Tokiko Watanabe, Ericka Sorensen, Chairul a Nidom, Michael a Newton, Paul Ahlquist, and Yoshihiro Kawaoka. 2008. “Drosophila RNAi Screen Identifies Host Genes Important for Influenza Virus Replication.” *Nature* 454(7206): 890–93.

Harris, John O, and R M Kline. 1956. “Electrophoresis of Motile Bacteria.” *Journal of bacteriology* 72(4): 530.

Hasegawa, Hiroshi, Hirotada Fujita, Hironori Katoh, Junko Aoki, Kazuhiro Nakamura, Atsushi Ichikawa, and Manabu Negishi. 1999. “Opposite Regulation of Transepithelial Electrical Resistance and Paracellular Permeability by Rho in Madin-Darby Canine Kidney Cells.” *Journal of Biological Chemistry* 274(30): 20982–88.

Hatta, Masato, Peng Gao, Peter Halfmann, and Yoshihiro Kawaoka. 2001. “Molecular Basis for High Virulence of Hong Kong H5N1 Influenza A Viruses.” *Science* 293 (5536): 1840–42.

Heller, Brigitte, Emmanuel Adu-Gyamfi, Whitney Smith-Kinnaman, Cliff Babbey, Mohsin Vora, Yi Xue, Robert Bittman, Robert V. Stahelin, and Clark D. Wells. 2010. “Amot Recognizes a Juxtanuclear Endocytic Recycling Compartment via a Novel Lipid Binding Domain.” *Journal of Biological Chemistry* 285(16): 12308–20.

Hernan, Ron, Ken Heuermann, and Bill Brizzard. 2000. “Multiple Epitope Tagging of Expressed Proteins for Enhanced Detection.” *BioTechniques* 28(4): 789–93.

Hidalgo, Ismael J, Thomas J Raub, and Ronald T Borchardt. 1989. “Characterization of the Human Colon Carcinoma Cell Line (Caco-2) as a Model System for Intestinal Epithelial Permeability.” *Gastroenterology* (96): 736–49.

Hodgkin, A, and A Huxley. 1952. “A Quantitative Description of Membrane Current and Its Application to Conduction and Excitation in Nerve.” *Bulletin of Mathematical Biology* 52(1-2): 25–71.

Hoffmann, E, G Neumann, Y Kawaoka, G Hobom, and R G Webster. 2000. “A DNA Transfection System for Generation of Influenza A Virus from Eight Plasmids.” *Proceedings of the National Academy of Sciences of the United States of America* 97(11): 6108–13.

Hopp, Thomas P, Kathryn S Prickett, Virginia L Price, Randell T Libby, Carl J March, Douglas Pat Cerretti, David L Urdal, and Paul J Conlon. 1988. “A Short Polypeptide Marker Sequence Useful for Recombinant Protein Identification and Purification.” *Biotechnology* 6(10): 1204–10.

Hosack, Douglas A, Glynn Dennis Jr, Brad T Sherman, H Clifford Lane, and Richard A Lempicki. 2003. “Identifying Biological Themes within Lists of Genes with EASE.” *Genome Biol* 4(10): R70.

Houel, Stephane, Robert Abernathy, Kutralanathan Renganathan, Karen Meyer-Arendt, Natalie G Ahn, and William M Old. 2010. “Quantifying the Impact of Chimera MS/MS Spectra on Peptide Identification in Large Scale Proteomics Studies.” *Journal of proteome research* 9(8): 4152–60.

Huang, Da Wei, Brad T Sherman, and Richard A Lempicki. 2009a. “Bioinformatics Enrichment Tools: Paths toward the Comprehensive Functional Analysis of Large Gene Lists.” *Nucleic Acids Research* 37 (1): 1–13.

Huang, Da Wei, Brad T Sherman, and Richard a Lempicki. 2009b. “Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources.” *Nature protocols* 4(1): 44–57.

Huang, Qian, Dongyu Liu, Paul Majewski, Leah C Schulte, Joshua M Korn, Richard A Young, Eric S Lander, and Nir Hacohen. 2001. “The Plasticity of Dendritic Cell Responses to Pathogens and Their Components.” *Science* 294 (5543): 870–75.

Huang, S., J. Chen, Q. Chen, H. Wang, Y. Yao, J. Chen, and Z. Chen. 2012. “A Second CRM1-Dependent Nuclear Export Signal in the Influenza A Virus NS2 Protein (NEP) Contributes to the Nuclear Export of Viral Ribonucleoproteins.” *Journal of Virology* 87(2): 767–78.

Huang, Shengping, Jingjing Chen, Huadong Wang, Bing Sun, Hanzhong Wang, Zhiping Zhang, Xianen Zhang, and Ze Chen. 2009. “Influenza A Virus Matrix Protein 1 Interacts with hTFIIIC102-S, a Short Isoform of the Polypeptide 3 Subunit of Human General Transcription Factor IIIC.” *Archives of virology* 154(7): 1101–10.

Huerta, Miriam, Rodrigo Muñoz, Rocío Tapia, Ernesto Soto-Reyes, Leticia Ramírez, Félix Recillas-Targa, Lorenza González-Mariscal, and Esther López-Bayghen. 2007. “Cyclin D1 Is Transcriptionally Down-Regulated by ZO-2 via an E Box and the Transcription Factor c-Myc.” *Molecular Biology of the Cell* 18 (12): 4826–36.

Van Itallie, Christina, Christoph Rahner, and James Melvin Anderson. 2001. “Regulated Expression of Claudin-4 Decreases Paracellular Conductance through a Selective Decrease in Sodium Permeability.” *The Journal of Clinical Investigation* 107(10): 1319–27.

Ito, T. 2000. “Interspecies Transmission and Receptor Recognition of Influenza A Viruses.” *Microbiol.Immunol.* 44(0385-5600 (Print)): 423–30.

Jäger, Stefanie, Peter Cimermancic, Natali Gulbahce, Jeffrey R. Johnson, Kathryn E. McGovern, Starlynn C. Clarke, Michael Shales, Gaelle Mercenne, Lars Pache, Kathy Li, Hilda Hernandez, Gwendolyn M. Jang, Shoshannah L. Roth, Eyal Akiva, John Marlett, Melanie Stephens, Ivan D’Orso, Jason Fernandes, Marie Fahey, Cathal Mahon, Anthony J. O’Donoghue, Aleksandar Todorovic, John H. Morris, David a. Maltby, Tom Alber, Gerard Cagney, Frederic D. Bushman, John a. Young, Sumit K. Chanda, Wesley I. Sundquist, Tanja Kortemme, Ryan D. Hernandez, Charles S. Craik, Alma Burlingame, Andrej Sali, Alan D. Frankel, and Nevan J. Krogan. 2011. “Global Landscape of HIV–human Protein Complexes.” *Nature*: 5–10.

Ji, Xinhua, Pinghui Zhang, Richard N Armstrong, and Gary L Gilliland. 1992. “The Three-Dimensional Structure of a Glutathione S-Transferase from the Mu Gene Class. Structural Analysis of the Binary Complex of Isoenzyme 3-3 and Glutathione at 2.2-.ANG. Resolution.” *Biochemistry* 31(42): 10169–84.

Jin, Hong, George P Leser, Jie Zhang, and Robert A Lamb. 1997. “Influenza Virus Hemagglutinin and Neuraminidase Cytoplasmic Tails Control Particle Shape.” *The EMBO Journal* 16(6): 1236–47.

Johnson, Niall P A S, and Juergen Mueller. 2002. “Updating the Accounts: Global Mortality of the 1918-1920 ‘Spanish’ Influenza Pandemic.” *Bulletin of the history of medicine* 76(1): 105–15.

de Jong, Menno D, Cameron P Simmons, Tran Tan Thanh, Vo Minh Hien, Gavin J D Smith, Tran Nguyen Bich Chau, Dang Minh Hoang, Nguyen Van Vinh Chau, Truong Huu Khanh, Vo Cong Dong, Phan Tu Qui, Bach Van Cam, Do Quang Ha, Yi Guan, J S Malik Peiris, Nguyen Tran Chinh, Tran Tinh Hien, and Jeremy Farrar. 2006. “Fatal Outcome of Human Influenza A (H5N1) Is Associated with High Viral Load and Hypercytokinemia.” *Nat Med* 12(10): 1203–7.

Jorba, Núria, Silvia Juarez, Eva Torreira, Pablo Gastaminza, Noelia Zamarreño, Juan Pablo Albar, and Juan Ortín. 2008. “Analysis of the Interaction of Influenza Virus Polymerase Complex with Human Cell Factors.” *Proteomics* 8(10): 2077–88.

Jovanovic, M., M. S. Rooney, P. Mertins, D. Przybylski, N. Chevrier, R. Satija, E. H. Rodriguez, A. P. Fields, S. Schwartz, R. Raychowdhury, M. R. Mumbach, T. Eisenhaure, M. Rabani, D. Gennert, D. Lu, T. Delorey, J. S. Weissman, S. A. Carr, N. Hacohen, and A. Regev. 2015. “Dynamic Profiling of the Protein Life Cycle in Response to Pathogens.” *Science* 347(6226): 1259038–1259038.

Juozapaitis, Mindaugas, Étori Aguiar Moreira, Ignacio Mena, Sebastian Giese, David Riegger, Anne Pohlmann, Dirk Höper, Gert Zimmer, Martin Beer, Adolfo García-Sastre, and Martin Schwemmle. 2014. “An Infectious Bat-Derived Chimeric Influenza Virus Harbouring the Entry Machinery of an Influenza A Virus.” *Nat Commun* 5.

Karas, M, and F Hillenkamp. 1988. “Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10,000 Daltons.” *Analytical chemistry* 60(20): 2299–2301.

Karlas, Alexander, Nikolaus Machuy, Yujin Shin, Klaus-Peter Pleissner, Anita Artarini, Dagmar Heuer, Daniel Becker, Hany Khalil, Lesley a Ogilvie, Simone Hess, André P Mäurer, Elke Müller, Thorsten Wolff, Thomas Rudel, and Thomas F Meyer. 2010. "Genome-Wide RNAi Screen Identifies Human Host Factors Crucial for Influenza Virus Replication." *Nature* 463(7282): 818–22.

Karlsson, Robert, Anne Michaelsson, and Lars Mattsson. 1991. "Kinetic Analysis of Monoclonal Antibody-Antigen Interactions with a New Biosensor Based Analytical System." *Journal of Immunological Methods* 145(1): 229–40.

Kashyap, Arun K, John Steel, Ahmet F Oner, Michael a Dillon, Ryann E Swale, Katherine M Wall, Kimberly J Perry, Aleksandr Faynboym, Mahmut Ilhan, Michael Horowitz, Lawrence Horowitz, Peter Palese, Ramesh R Bhatt, and Richard a Lerner. 2008. "Combinatorial Antibody Libraries from Survivors of the Turkish H5N1 Avian Influenza Outbreak Reveal Virus Neutralization Strategies." *Proceedings of the National Academy of Sciences of the United States of America* 105(16): 5986–91.

Keefe, a D, D S Wilson, B Seelig, and J W Szostak. 2001. "One-Step Purification of Recombinant Proteins Using a Nanomolar-Affinity Streptavidin-Binding Peptide, the SBP-Tag." *Protein expression and purification* 23(3): 440–46.

Khadka, S., a. D. Vangeloff, C. Zhang, P. Siddavatam, N. S. Heaton, L. Wang, R. Sengupta, S. Sahasrabudhe, G. Randall, M. Gribskov, R. J. Kuhn, R. Perera, and D. J. LaCount. 2011. "A Physical Interaction Network of Dengue Virus and Human Proteins." *Molecular & Cellular Proteomics* 10(12): M111.012187–M111.012187.

Kim, Jin Hee, Sang-Rok Lee, Li-Hua Li, Hye-Jeong Park, Jeong-Hoh Park, Kwang Youl Lee, Myeong-Kyu Kim, Boo Ahn Shin, and Seok-Yong Choi. 2011. "High Cleavage Efficiency of a 2A Peptide Derived from Porcine Teschovirus-1 in Human Cell Lines, Zebrafish and Mice." *PloS one* 6(4): e18556.

Kim, Min-Sik, Sneha M Pinto, Derese Getnet, Raja Sekhar Nirujogi, Srikanth S Manda, Raghothama Chaerkady, Anil K Madugundu, Dhanashree S Kelkar, Ruth Isserlin, Shobhit Jain, Joji K Thomas, Babylakshmi Muthusamy, Pamela Leal-Rojas, Praveen Kumar, Nandini A Sahasrabuddhe, Lavanya Balakrishnan, Jayshree Advani, Bijesh George, Santosh Renuse, Lakshmi Dhevi N Selvan, Arun H Patil, Vishalakshi Nanjappa, Aneesha Radhakrishnan, Samarjeet Prasad, Tejaswini Subbannayya, Rajesh Raju, Manish Kumar, Sreelakshmi K Sreenivasamurthy, Arivusudar Marimuthu, Gajanan J Sathe, Sandip Chavan, Keshava K Datta, Yashwanth Subbannayya, Apeksha Sahu, Soujanya D Yelamanchi, Savita Jayaram, Pavithra Rajagopalan, Jyoti Sharma, Krishna R Murthy, Nazia Syed, Renu Goel, Aafaque A Khan, Sartaj Ahmad, Gourav Dey, Keshav Mudgal, Aditi Chatterjee, Tai-Chung Huang, Jun Zhong, Xinyan Wu, Patrick G Shaw, Donald Freed, Muhammad S Zahari, Kanchan K Mukherjee, Subramanian Shankar, Anita Mahadevan, Henry Lam, Christopher J Mitchell, Susarla Krishna Shankar, Parthasarathy Satishchandra, John T Schroeder, Ravi Sirdeshmukh, Anirban Maitra, Steven D Leach, Charles G Drake, Marc K Halushka, T S Keshava Prasad, Ralph H Hruban, Candace L Kerr, Gary D Bader, Christine A Iacobuzio-Donahue, Harsha Gowda, and Akhilesh Pandey. 2014. "A Draft Map of the Human Proteome." *Nature* 509(7502): 575–81.

Köhler, Katja, Daniel Louvard, and Ahmed Zahraoui. 2004. "Rab13 Regulates PKA Signaling during Tight Junction Assembly." *The Journal of Cell Biology* 165 (2): 175–80.

Köhler, Katja, and Ahmed Zahraoui. 2005. "Tight Junction: A Co-ordinator of Cell Signalling and Membrane Trafficking." *Biology of the Cell* 97(8): 659–65.

König, Renate, Silke Stertz, Yingyao Zhou, Atsushi Inoue, H-Heinrich Hoffmann, Suchita Bhattacharyya, Judith G Alamares, Donna M Tscherne, Mila B Ortigoza, Yuhong Liang, Qinshan Gao, Shane E Andrews, Sourav Bandyopadhyay, Paul De Jesus, Buu P Tu, Lars Pache, Crystal Shih, Anthony Orth, Ghislain Bonamy, Loren Miraglia, Trey Ideker, Adolfo García-Sastre, John A T Young, Peter Palese, Megan L Shaw, and Sumit K Chanda. 2010. "Human Host Factors Required for Influenza Virus Replication." *Nature* 463(7282): 813–17.

Krogan, Nevan J, Gerard Cagney, Haiyuan Yu, Gouqing Zhong, Xinghua Guo, Alexandr Ignatchenko, Joyce Li, Shuye Pu, Nira Datta, Aaron P Tikuisis, Thanuja Punna, José M Peregrín-Alvarez, Michael Shales, Xin Zhang, Michael Davey, Mark D Robinson, Alberto Paccanaro, James E Bray, Anthony Sheung, Bryan Beattie, Dawn P Richards, Veronica Canadien, Atanas Lalev, Frank Mena, Peter Wong, Andrei Starostine, Myra M Canete, James Vlasblom, Samuel Wu, Chris Orsi, Sean R Collins, Shamanta Chandran, Robin Haw, Jennifer J Rilstone, Kiran Gandi, Natalie J Thompson, Gabe Musso, Peter St Onge, Shaun Ghanny, Mandy H Y Lam, Gareth Butland, Amin M Altaf-Ul, Shigehiko Kanaya, Ali Shilatifard, Erin O'Shea, Jonathan S Weissman, C James Ingles, Timothy R Hughes, John Parkinson, Mark Gerstein, Shoshana J Wodak, Andrew Emili, and Jack F Greenblatt. 2006. "Global Landscape of Protein Complexes in the Yeast Saccharomyces Cerevisiae." *Nature* 440(7084): 637–43.

Kumar, Manish, Hongbing Liu, and Andrew P Rice. 2012. "Regulation of Interferon-Beta by MAGI-1 and Its Interaction with Influenza A Virus NS1 Protein with ESEV PBM." *PloS one* 7(7): e41251.

LaCava, John, Kelly R Molloy, Martin S Taylor, Michal Domanski, Brian T Chait, and Michael P Rout. 2015. "Affinity Proteomics to Study Endogenous Protein Complexes: Pointers, Pitfalls, Preferences and Perspectives." *BioTechniques* 58(3): 103–19.

LaCount, Douglas J, Marissa Vignali, Rakesh Chettier, Amit Phansalkar, Russell Bell, Jay R Hesselberth, Lori W Schoenfeld, Irene Ota, Sudhir Sahasrabudhe, Cornelia Kurschner, Stanley Fields, and Robert E Hughes. 2005. "A Protein Interaction Network of the Malaria Parasite Plasmodium Falciparum." *Nature* 438(7064): 103–7.

Lakdawala, Seema S, Yicong Wu, Peter Wawrzusin, Juraj Kabat, Andrew J Broadbent, Elaine W Lamirande, Ervin Fodor, Nihal Altan-Bonnet, Hari Shroff, and Kanta Subbarao. 2014. "Influenza A Virus Assembly Intermediates Fuse in the Cytoplasm." *PLoS Pathog* 10(3): e1003971.

Lane, D P, V Simanis, R Bartsch, J Yewdell, J Gannon, and S Mole. 1985. "Cellular Targets for SV40 Large T-Antigen." *Proceedings of the Royal Society of London B: Biological Sciences* 226(1242): 25–42.

Latorre, Isabel J, Michael H Roh, Kristopher K Frese, Robert S Weiss, Ben Margolis, and Ronald T Javier. 2005. "Viral Oncoprotein-Induced Mislocalization of Select PDZ Proteins Disrupts Tight Junctions and Causes Polarity Defects in Epithelial Cells." *Journal of Cell Science* 118(18): 4283–93.

Lauterbach, Sonja B, Roberto Lanzillotti, and Theresa L Coetzer. 2003. "Libraries To Identify Host Parasite Interactions." *Malaria Journal* 4: 1–4.

Lee, Jun Han, Sung-Hak Kim, Philippe Noriel Q Pascua, Min-Suk Song, Yun Hee Baek, Xun Jin, Joong-Kook Choi, Chul-Joong Kim, Hyunggee Kim, and Young Ki Choi. 2010. "Direct Interaction of Cellular hnRNP-F and NS1 of Influenza A Virus Accelerates Viral Replication by Modulation of Viral Transcriptional Activity and Host Gene Expression." *Virology* 397(1): 89–99.

Lewis, B A, A Elkin, R H Michell, and R Coleman. 1975. "Basolateral Plasma Membranes of Intestinal Epithelial Cells. Identification by Lactoperoxidase-Catalysed Iodination and Isolation after Density Perturbation with Digitonin." *Biochemical Journal* 152(1): 71–84.

Lewis, Simon A, and Jared M Diamond. 1976. "Na+ Transport by Rabbit Urinary Bladder, a Tight Epithelium." *The Journal of membrane biology* 28(1): 1–40.

Li, Jingyi Jessica, and Mark D. Biggin. 2015. "Statistics Requantitates the Central Dogma." *Science* 347(6226): 1066–67.

Liao, Qing, Norman H L Chiu, Chang Shen, Yi Chen, and Paul Vouros. 2007. "Investigation of Enzymatic Behavior of Benzonase/alkaline Phosphatase in the Digestion of Oligonucleotides and DNA by ESI-LC/MS." *Analytical chemistry* 79(5): 1907–17.

Linford, Andrea, Shin-ichiro Yoshimura, Ricardo Nunes Bastos, Lars Langemeyer, Andreas Gerondopoulos, Daniel J. Rigden, and Francis A. Barr. 2012. "Rab14 and Its Exchange Factor FAM116 Link Endocytic Recycling and Adherens Junction Stability in Migrating Cells." *Developmental Cell* 22(5): 952–66.

Liu, J, H Xiao, F Lei, Q Zhu, K Qin, X.-w. Zhang, X.-l. Zhang, D Zhao, G Wang, Y Feng, J Ma, W Liu, J Wang, and G F Gao. 2005. "Highly Pathogenic H5N1 Influenza Virus Infection in Migratory Birds." *Science* 309 (5738): 1206.

Liu, Xiaoling, Lei Sun, Maorong Yu, Zengfu Wang, Chongfeng Xu, Qinghua Xue, Ke Zhang, Xin Ye, Yoshihiro Kitamura, and Wenjun Liu. 2009. “Cyclophilin A Interacts with Influenza A Virus M1 Protein and Impairs the Early Stage of the Viral Replication.” *Cellular microbiology* 11(February): 730–41.

Lu, Ruifeng, Debra L Johnson, Lorraine Stewart, Kelsey Waite, David Elliott, and Jean M Wilson. 2014. “Rab14 Regulation of Claudin-2 Trafficking Modulates Epithelial Permeability and Lumen Morphogenesis.” *Molecular biology of the cell* 25(11): 1744–54.

Lu, Ruifeng, Lorraine Stewart, and Jean M Wilson. 2015. “Scaffolding Protein GOPC Regulates Tight Junction Structure.” *Cell and tissue research* 360(2): 321–32.

Lund, Barbara. 1965. “A Comparison by the Use of Gel Electrophoresis of Soluble Protein Components and Esterase Enzymes of Some Group D Streptococci.” *Journal of general microbiology* 40(3): 413–19.

Luytjes, Willem, Mark Krystal, Masayoshi Enami, Jeffrey D. Parvin, and Peter Palese. 1989. “Amplification, Expression, and Packaging of a Foreign Gene by Influenza Virus.” *Cell* 59(6): 1107–13.

Mahmoudian, Shohreh, Sabrina Auerochs, Monika Gröne, and Manfred Marschall. 2009. “Influenza A Virus Proteins PB1 and NS1 Are Subject to Functionally Important Phosphorylation by Protein Kinase C.” *The Journal of general virology* 90(Pt 6): 1392–97.

Mänz, Benjamin, Linda Brunotte, Peter Reuther, and Martin Schwemmle. 2012. “Adaptive Mutations in NEP Compensate for Defective H5N1 RNA Replication in Cultured Human Cells.” *Nature communications* 3(May): 802.

Martinez-Palomo, A, I Meza, G Beaty, and M Cereijido. 1980. “Experimental Modulation of Occluding Junctions in a Cultured Transporting Epithelium.” *The Journal of Cell Biology* 87 (3): 736–45.

Massin, Pascale, Sylvie van der Werf, and Nadia Naffakh. 2001. “Residue 627 of PB2 Is a Determinant of Cold Sensitivity in RNA Replication of Avian Influenza Viruses.” *Journal of Virology* 75 (11): 5398–5404.

Matter, Karl, and Maria Susana Balda. 2007. “Epithelial Tight Junctions, Gene Expression and Nucleo-Junctional Interplay.” *Journal of Cell Science* 120(9): 1505–11.

Mayer, Daniel, Kaaweh Molawi, Luis Martı, Alexander Ghanem, Stefan Thomas, Sacha Baginsky, Jonas Grossmann, Adolfo Garcı, and Martin Schwemmle. 2007. “Identification of Cellular Interaction Partners of the Influenza Virus Ribonucleoprotein Complex and Polymerase Complex Using Proteomic-Based Approaches Research Articles.” *Journal of Proteome Research*: 672–82.

McChesney, Evan W, and Wm Kirk Swann Jr. 1937. “The Identification of the Amino Acids: P-Toluenesulfonyl1 Chloride as a Reagent.” *Journal of the American Chemical Society* 59(6): 1116–18.

McFadden, E R, B M Pichurko, H F Bowman, E Ingenito, S Burns, N Dowling, and J Solway. 1985. “Thermal Mapping of the Airways in Humans.” *Journal of Applied Physiology* 58(2): 564–70.

Mellacheruvu, Dattatreya, Zachary Wright, Amber L Couzens, Jean-Philippe Lambert, Nicole A St-Denis, Tuo Li, Yana V Miteva, Simon Hauri, Mihaela E Sardiu, and Teck Yew Low. 2013. “The CRAPome: A Contaminant Repository for Affinity Purification-Mass Spectrometry Data.” *Nature methods* 10(8): 730–36.

Mercenne, Gaelle, Steven L Alam, Jun Arii, Matthew S Lalonde, and Wesley I Sundquist. 2015. “Angiomotin Functions in HIV-1 Assembly and Budding.” *eLife* 4(January): 1–23.

Metersky, Mark L, Robert G Masterton, Hartmut Lode, Thomas M File Jr, and Timothy Babinchak. 2012. “Epidemiology, Microbiology, and Treatment Considerations for Bacterial Pneumonia Complicating Influenza.” *International Journal of Infectious Diseases* 16(5): e321–31.

Misfeldt, Dayton Stanley, Susan T Hamamoto, and Dorothy R Pitelka. 1976. “Transepithelial Transport in Cell Culture.” *Proceedings of the National Academy of Sciences* 73(4): 1212–16.

Mitzner, David, Sabine Eva Dudek, Nicole Studtrucker, Darisuren Anhlan, Igor Mazur, Josef Wissing, Lothar Jänsch, Ludmilla Wixler, Karsten Bruns, Alok Sharma, Victor Wray, Peter Henklein, Stephan Ludwig, and Ulrich Schubert. 2009. “Phosphorylation of the Influenza A Virus Protein PB1-F2 by PKC Is Crucial for Apoptosis Promoting Functions in Monocytes.” *Cellular microbiology* 11(10): 1502–16.

Momose, Fumitaka, Tetsuya Sekimoto, Takashi Ohkura, Shuichi Jo, Atsushi Kawaguchi, Kyosuke Nagata, and Yuko Morikawa. 2011. “Apical Transport of Influenza A Virus Ribonucleoprotein Requires Rab11-Positive Recycling Endosome.” *PLoS ONE* 6(6): e21123.

Muramoto, Yukiko, Ayato Takada, Ken Fujii, Takeshi Noda, Kiyoko Iwatsuki-Horimoto, Shinji Watanabe, Taisuke Horimoto, Hiroshi Kida, and Yoshihiro Kawaoka. 2006. “Hierarchy among Viral RNA (vRNA) Segments in Their Role in vRNA Incorporation into Influenza A Virions.” *Journal of virology* 80(5): 2318–25.

Nayak, D P, G W Kelley, G A Young, and N R Underdahl. 1964. “Progressive Descending Influenza Infection in Mice Determined by Immunofluorescence.” *Experimental Biology and Medicine* 116 (1): 200–206.

Nayak, Debi P, Eric Ka-Wai Hui, and Subrata Barman. 2004. “Assembly and Budding of Influenza Virus.” *Virus research* 106(2): 147–65.

Naydenov, Nayden G, Bryan Brown, Gianni Harris, Michael R Dohn, Victor M Morales, Somesh Baranwal, Albert B Reynolds, and Andrei I Ivanov. 2012. “A Membrane Fusion Protein αSNAP Is a Novel Regulator of Epithelial Apical Junctions.” *PLoS ONE* 7(4): e34320.

Nelson, Martha I, Cécile Viboud, Lone Simonsen, Ryan T Bennett, Sara B Griesemer, Kirsten St. George, Jill Taylor, David J Spiro, Naomi A Sengamalay, Elodie Ghedin, Jeffery K Taubenberger, and Edward C Holmes. 2008. “Multiple Reassortment Events in the Evolutionary History of H1N1 Influenza A Virus Since 1918.” *PLoS Pathog* 4(2): e1000012.

Nesvizhskii, Alexey I. 2012. “Computational and Informatics Strategies for Identification of Specific Protein Interaction Partners in Affinity Purification Mass Spectrometry Experiments.” *Proteomics* 12(10): 1639–55.

Neumann, G, M T Hughes, and Y Kawaoka. 2000. “Influenza A Virus NS2 Protein Mediates vRNP Nuclear Export through NES-Independent Interaction with hCRM1.” *The EMBO journal* 19(24): 6751–58.

Neumann, G, T Watanabe, H Ito, S Watanabe, H Goto, P Gao, M Hughes, D R Perez, R Donis, E Hoffmann, G Hobom, and Y Kawaoka. 1999. “Generation of Influenza A Viruses Entirely from Cloned cDNAs.” *Proceedings of the National Academy of Sciences of the United States of America* 96(16): 9345–50.

Nguyen-Van-Tam, J. 2003. “The Epidemiology and Clinical Impact of Pandemic Influenza.” *Vaccine* 21(16): 1762–68.

Oda, Y, K Huang, F R Cross, D Cowburn, and B T Chait. 1999. “Accurate Quantitation of Protein Expression and Site-Specific Phosphorylation.” *Proceedings of the National Academy of Sciences of the United States of America* 96(12): 6591–96.

ONeill, R E, J Talon, and P Palese. 1998. “The Influenza Virus NEP (NS2 Protein) Mediates the Nuclear Export of Viral Ribonucleoproteins.” *The EMBO journal* 17(1): 288–96.

Ong, S.-E. 2002. “Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics.” *Molecular & Cellular Proteomics* 1(5): 376–86.

Palese, P, and M L Shaw. 2007. “Fields Virology.” *Orthomyxoviridae: The Viruses and Their Replication, 5th edn, Philadelphia, PA: Lippincott Williams & Wilkins, Wolters Kluwer Business*: 1647–89.

Paša-Tolić, Ljiljana, Pamela K Jensen, Gordon A Anderson, Mary S Lipton, Kim K Peden, Suzana Martinović, Nikola Tolić, James E Bruce, and Richard D Smith. 1999. “High Throughput Proteome-Wide Precision Measurements of Protein Expression Using Mass Spectrometry.” *Journal of the American Chemical Society* 121(34): 7949–50.

Peachey, Lee D, and Howard Rasmussen. 1961. “STRUCTURE OF THE TOAD’S URINARY BLADDER AS RELATED TO ITS PHYSIOLOGY.” *The Journal of Biophysical and Biochemical Cytology* 10(4): 529–53.

Pena, Lindomar, Troy Sutton, Ashok Chockalingam, Sachin Kumar, Matthew Angel, Hongxia Shao, Hongjun Chen, Weizhong Li, and Daniel R Perez. 2013. “Influenza Viruses with Rearranged Genomes as Live-Attenuated Vaccines.” *Journal of Virology* 87 (9): 5118–27.

Percy, N, W S Barclay, a García-Sastre, and P Palese. 1994. “Expression of a Foreign Protein by Influenza A Virus.” *Journal of virology* 68(7): 4486–92.

Perwitasari, Olivia, Scott Johnson, Xiuzhen Yan, Elizabeth Howerth, Sharon Shacham, Yosef Landesman, Erkan Baloglu, Dilara McCauley, Sharon Tamir, S Mark Tompkins, and Ralph A Tripp. 2014. "Verdinexor, a Novel Selective Inhibitor of Nuclear Export, Reduces Influenza A Virus Replication In Vitro and In Vivo." *Journal of Virology* 88 (17): 10228–43.

Peter, Isabelle S, and Eric H Davidson. 2011. "A Gene Regulatory Network Controlling the Embryonic Specification of Endoderm." *Nature* 474(7353): 635–39.

Porter, Keith R. 1956. "Observations on the Fine Structure of Animal Epidermis." In *Proceedings of the International Congress on Electron Microscopy. Royal Microscopical Society, London*,.

Presti, Rachel M, Guoyan Zhao, Wandy L Beatty, Kathie A Mihindukulasuriya, Amelia P A Travassos da Rosa, Vsevolod L Popov, Robert B Tesh, Herbert W Virgin, and David Wang. 2009. "Quaranfil, Johnston Atoll, and Lake Chad Viruses Are Novel Members of the Family Orthomyxoviridae ." *Journal of Virology* 83 (22): 11599–606.

Prickett, Kathryn S, David C Amberg, and Thomas P Hopp. 1989. "A Calcium-Dependent Antibody for Identification and Purification of Recombinant Proteins." *Biotechniques* 7(6): 580–89.

Rain, J C, L Selig, H De Reuse, V Battaglia, C Reverdy, S Simon, G Lenzen, F Petel, J Wojcik, V Schächter, Y Chemama, a Labigne, and P Legrain. 2001. "The Protein-Protein Interaction Map of Helicobacter Pylori." *Nature* 409(6817): 211–15.

Raman, C S, Ronald Jemmerson, Barry T Nall, and Michael J Allen. 1992. "Diffusion-Limited Rates for Monoclonal Antibody Binding to Cytochrome c." *Biochemistry* 31(42): 10370–79.

Ray, Sandipan, Karthik S Kamath, Rajneesh Srivastava, Dinesh Raghu, Kishore Gollapalli, Rekha Jain, Shipra V Gupta, Sayantan Ray, Santosh Taur, Snigdha Dhali, Nithya Gogtay, Urmila Thatte, Rapole Srikanth, Swati Patankar, and Sanjeeva Srivastava. 2012. "Serum Proteome Analysis of Vivax Malaria: An Insight into the Disease Pathogenesis and Host Immune Response." *Journal of Proteomics* 75(10): 3063–80.

Reinhardt, J, and T Wolff. 2000. "The Influenza A Virus M1 Protein Interacts with the Cellular Receptor of Activated C Kinase (RACK) 1 and Can Be Phosphorylated by Protein Kinase C." *Veterinary microbiology* 74(1-2): 87–100.

Reuther, Peter, Sebastian Giese, Veronika Götz, Normann Kilb, Benjamin Mänz, Linda Brunotte, and Martin Schwemmle. 2014. "Adaptive Mutations in the Nuclear Export Protein of Human-Derived H5N1 Strains Facilitate a Polymerase Activity-Enhancing Conformation." *Journal of virology* 88(1): 263–71.

Reuther, Peter, Kristina Göpfert, Alexandra H. Dudek, Monika Heiner, Susanne Herold, and Martin Schwemmle. 2015. "Generation of a Variety of Stable Influenza A Reporter Viruses by Genetic Engineering of the NS Gene Segment." *Scientific Reports* 5(January 2015): 11346.

Robb, Nicole C, Matt Smith, Frank T Vreede, and Ervin Fodor. 2009. "NS2/NEP Protein Regulates Transcription and Replication of the Influenza Virus RNA Genome." *The Journal of general virology* 90(Pt 6): 1398–1407.

Rodriguez-Boulan, Enrique, Kevin T Paskiet, and David D Sabatini. 1983. "Assembly of Enveloped Viruses in Madin-Darby Canine Kidney Cells: Polarized Budding from Single Attached Cells and from Clusters of Cells in Suspension." *The Journal of cell biology* 96(3): 866–74.

Roeth, Jeremiah F, Jessica K Sawyer, Daniel A Wilner, and Mark Peifer. 2009. "Rab11 Helps Maintain Apical Crumbs and Adherens Junctions in the Drosophila Embryonic Ectoderm." *PLoS ONE* 4(10): e7634.

Roy, Koushik, George M Hilliard, David J Hamilton, Jiwen Luo, Marguerite M Ostmann, and James M Fleckenstein. 2009. "Enterotoxigenic Escherichia Coli EtpA Mediates Adhesion between Flagella and Host Cells." *Nature* 457(7229): 594–98.

Ryan, Martin D, A M King, and G Paul Thomas. 1991. "Cleavage of Foot-and-Mouth Disease Virus Polyprotein Is Mediated by Residues Located within a 19 Amino Acid Sequence." *J Gen Virol* 72(Pt 11): 2727–32.

Saleem, Ramsey A, Barbara Knoblach, Fred D Mast, Jennifer J Smith, John Boyle, C Melissa Dobson, Rose Long-O'Donnell, Richard A Rachubinski, and John D Aitchison. 2008. "Genome-Wide Analysis of Signaling Networks Regulating Fatty Acid–induced Gene Expression and Organelle Biogenesis." *The Journal of Cell Biology* 181 (2): 281–92.

Sasaki, Hiroyuki, Chiyuki Matsui, Kyoko Furuse, Yuko Mimori-Kiyosue, Mikio Furuse, and Shoichiro Tsukita. 2003. "Dynamic Behavior of Paired Claudin Strands within Apposing Plasma Membranes." *Proceedings of the National Academy of Sciences* 100 (7): 3971–76.

Satterly, Neal, Pei-Ling Tsai, Jan van Deursen, Daniel R Nussenzveig, Yaming Wang, Paula a Faria, Agata Levay, David E Levy, and Beatriz M a Fontoura. 2007. "Influenza Virus Targets the mRNA Export Machinery and the Nuclear Pore Complex." *Proceedings of the National Academy of Sciences of the United States of America* 104(6): 1853–58.

Schmidt, Peter M, Lindsay G Sparrow, Rebecca M Attwood, Xiaowen Xiao, Tim E Adams, and Jennifer L McKimm-Breschkin. 2012. "Taking down the FLAG! How Insect Cell Expression Challenges an Established Tag-System." *PloS one* 7(6): e37779.

Schneeberger, Eveline E, and Robert D Lynch. 2004. "The Tight Junction: A Multifunctional Complex." *American Journal of Physiology - Cell Physiology* 286(6): C1213–28.

Schrimpf, Sabine P, Manuel Weiss, Lukas Reiter, Christian H Ahrens, Marko Jovanovic, Johan Malmström, Erich Brunner, Sonali Mohanty, Martin J Lercher, Peter E Hunziker, Ruedi Aebersold, Christian von Mering, and Michael O Hengartner. 2009. "Comparative Functional Analysis of the Caenorhabditis Elegans and Drosophila Melanogaster Proteomes." *PLoS biology* 7(3): e48.

Selbach, Matthias, Florian Ernst Paul, Sabine Brandt, Patrick Guye, Oliver Daumke, Steffen Backert, Christoph Dehio, and Matthias Mann. 2009. "Host Cell Interactome of Tyrosine-Phosphorylated Bacterial Proteins." *Cell Host and Microbe* 5(4): 397–403.

Shannon, Paul, Andrew Markiel, Owen Ozier, Nitin S Baliga, Jonathan T Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. 2003. "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks." *Genome research* 13(11): 2498–2504.

Shapira, Sagi D, Irit Gat-Viks, Bennett O V Shum, Amelie Dricot, Marciela M de Grace, Liguo Wu, Piyush B Gupta, Tong Hao, Serena J Silver, David E Root, David E Hill, Aviv Regev, and Nir Hacohen. 2009. "A Physical and Regulatory Map of Host-Influenza Interactions Reveals Pathways in H1N1 Infection." *Cell* 139(7): 1255–67.

Shi, Yi, Ying Wu, Wei Zhang, Jianxun Qi, and George F Gao. 2014. "Enabling the 'Host Jump': Structural Determinants of Receptor-Binding Specificity in Influenza A Viruses." *Nature reviews. Microbiology* 12(12): 822–31.

Shimizu, Teppei, Naoki Takizawa, Ken Watanabe, Kyosuke Nagata, and Nobuyuki Kobayashi. 2010. "Crucial Role of the Influenza Virus NS2 (NEP) C-Terminal Domain in M1 Binding and Nuclear Export of vRNP." *FEBS letters* 2: 1–6.

Shui, Wenqing, Sarah a Gilmore, Leslie Sheu, Jun Liu, Jay D Keasling, and Carolyn R Bertozzi. 2009. "Quantitative Proteomic Profiling of Host - Pathogen Interactions : The Macrophage Response to Mycobacterium Tuberculosis Lipids Research Articles." : 282–89.

Simonovic, Ivana, Monique Arpin, Athanasia Koutsouris, Holly J Falk-Krzesinski, and Gail Hecht. 2001. "Enteropathogenic Escherichia Coli Activates Ezrin, Which Participates in Disruption of Tight Junction Barrier Function." *Infection and Immunity* 69 (9): 5679–88.

Simpson-Holley, Martha, Darren Ellis, Dawn Fisher, Debra Elton, John McCauley, and Paul Digard. 2002. "A Functional Link between the Actin Cytoskeleton and Lipid Rafts during Budding of Filamentous Influenza Virions." *Virology* 301(2): 212–25.

Sjöbring, U, L Björck, and W Kastern. 1991. "Streptococcal Protein G. Gene Structure and Protein Binding Properties." *Journal of Biological Chemistry* 266 (1): 399–405.

Smith, Gavin J D, Dhanasekaran Vijaykrishna, Justin Bahl, Samantha J Lycett, Michael Worobey, Oliver G Pybus, Siu Kit Ma, Chung Lam Cheung, Jayna Raghwani, Samir Bhatt, J S Malik Peiris, Yi Guan, and Andrew Rambaut. 2009. "Origins and Evolutionary Genomics of the 2009 Swine-Origin H1N1 Influenza A Epidemic." *Nature* 459(7250): 1122–25.

Smith, George P, and Valery a Petrenko. 1997. "Phage Display." *Chemical Reviews* 97(96): 391–410.

Smithies, O. 1955. "Zone Electrophoresis in Starch Gels: Group Variations in the Serum Proteins of Normal Human Adults." *The Biochemical journal* 61(4): 629–41.

Song, Min-Suk, Philippe Noriel Q Pascua, Jun Han Lee, Yun Hee Baek, Ok-Jun Lee, Chul-Joong Kim, Hyunggee Kim, Richard J Webby, Robert G Webster, and Young Ki Choi. 2009. "The Polymerase Acidic Protein Gene of Influenza A Virus Contributes to Pathogenicity in a Mouse Model." *Journal of Virology* 83 (23): 12325–35.

Song, Wenjun, Pui Wang, Bobo Wing-Yee Mok, Siu-Ying Lau, Xiaofeng Huang, Wai-Lan Wu, Min Zheng, Xi Wen, Shigui Yang, Yu Chen, Lanjuan Li, Kwok-Yung Yuen, and Honglin Chen. 2014. "The K526R Substitution in Viral Protein PB2 Enhances the Effects of E627K on Influenza Virus Replication." *Nature Communications* 5(May 2014): 5509.

Sonoda, Noriyuki, Mikio Furuse, Hiroyuki Sasaki, Shigenobu Yonemura, Jun Katahira, Yasuhiko Horiguchi, and Shoichiro Tsukita. 1999. "Clostridium Perfringens Enterotoxin Fragment Removes Specific Claudins from Tight Junction Strands: Evidence for Direct Involvement of Claudins in Tight Junction Barrier ." *The Journal of Cell Biology* 147 (1): 195–204.

Staehelin, L A. 1973. "Further Observations on the Fine Structure of Freeze-Cleaved Tight Junctions." *Journal of Cell Science* 13(3): 763–86.

Steel, John, Anice C Lowen, Samira Mubareka, and Peter Palese. 2009. "Transmission of Influenza Virus in a Mammalian Host Is Increased by PB2 Amino Acids 627K or 627E/701N." *PLoS Pathog* 5(1): e1000252.

Stevenson, B R, J M Anderson, D A Goodenough, and M S Mooseker. 1988. "Tight Junction Structure and ZO-1 Content Are Identical in Two Strains of Madin-Darby Canine Kidney Cells Which Differ in Transepithelial Resistance." *The Journal of Cell Biology* 107 (6): 2401–8.

Stevenson, Bruce R, Janet D Siliciano, Mark S Mooseker, and Daniel A Goodenough. 1986. "Identification of ZO-1: A High Molecular Weight Polypeptide Associated with the Tight Junction (zonula Occludens) in a Variety of Epithelia." *The Journal of cell biology* 103(3): 755–66.

Stirnimann, Christian U, Evangelia Petsalaki, Robert B Russell, and Christoph W Müller. 2010. "WD40 Proteins Propel Cellular Networks." *Trends in Biochemical Sciences* 35(10): 565–74.

Stöhr, Klaus. 2002. "Influenza—WHO Cares." *The Lancet Infectious Diseases* 2(9): 517.

Sturm-Ramirez, Katharine M, Trevor Ellis, Barry Bousfield, Lucy Bissett, Kitman Dyrting, Jerold E Rehg, Leo Poon, Yi Guan, Malik Peiris, and Robert G Webster. 2004. "Reemerging H5N1 Influenza Viruses in Hong Kong in 2002 Are Highly Pathogenic to Ducks." *Journal of Virology* 78 (9): 4892–4901.

Sui, Baoquan, Douty Bamba, Ke Weng, Huong Ung, Shaojing Chang, Jessica Van Dyke, Michael Goldblatt, Roxanne Duan, Michael S Kinch, and Wu-Bo Li. 2009. "The Use of Random Homozygous Gene Perturbation to Identify Novel Host-Oriented Targets for Influenza." *Virology* 387(2): 473–81.

Szymczak, Andrea L, Creg J Workman, Yao Wang, Kate M Vignali, Smaroula Dilioglou, Elio F Vanin, and Dario A A Vignali. 2004. "Correction of Multi-Gene Deficiency in Vivo Using a Single'self-cleaving'2A Peptide–based Retroviral Vector." *Nature biotechnology* 22(5): 589–94.

Tackett, Alan J, Jeffrey A Degrasse, Matthew D Sekedat, Marlene Oeffinger, Michael P Rout, and Brian T Chait. 2005. "I-DIRT , A General Method for Distinguishing between Specific and Nonspecific Protein Interactions Research Articles." *Journal of Proteome Research*: 1752–56.

Takeuchi, A. 1967. "Electron Microscope Studies of Experimental Salmonella Infection. I. Penetration into the Intestinal Epithelium by Salmonella Typhimurium." *The American Journal of Pathology* 50(1): 109–36.

Takizawa, Naoki, Ken Watanabe, Kaoru Nouno, Nobuyuki Kobayashi, and Kyosuke Nagata. 2006. "Association of Functional Influenza Viral Proteins and RNAs with Nuclear Chromatin and Sub-Chromatin Structure." *Microbes and infection / Institut Pasteur* 8(3): 823–33.

Taubenberger, Jeffery K, and David M Morens. 2008. "The Pathology of Influenza Virus Infections." *Annual review of pathology* 3: 499.

Tawaratsumida, Kazuki, Van Phan, Eike R Hrincius, Anthony a High, Richard Webby, Vanessa Redecke, and Hans Häcker. 2014. "Quantitative Proteomic Analysis of the Influenza A Virus Nonstructural Proteins NS1 and NS2 during Natural Cell Infection Identifies PACT as an NS1 Target Protein and Antiviral Host Factor." *Journal of virology* 88(16): 9038–48.

Teoh, Kim-Tat, Yu-Lam Siu, Wing-Lim Chan, Marc A Schlüter, Chia-Jen Liu, J S Malik Peiris, Roberto Bruzzone, Benjamin Margolis, and Béatrice Nal. 2010. "The SARS Coronavirus E Protein Interacts with PALS1 and Alters Tight Junction Formation and Epithelial Morphogenesis." *Molecular biology of the cell* 21(22): 3838–52.

Thomason, B. M., W. B. Cherry, and P. R. Edwards. 1959. "Staining Bacterial Smears with Fluorescent Antibody. VI. Identification of Salmonellae in Fecal Specimens." *Journal of bacteriology* 77(4): 478–86.

Tiselius, Arne. 1930. 7 *The Moving Boundary Method of Studying the Electrophoreses of Proteins*. Almquist & Wilksells boktryckeri-a.-b.

Tobita, K, A Sugiura, C Enomoto, and M Furuyama. 1975. "Plaque Assay and Primary Isolation of Influenza A Viruses in an Established Line of Canine Kidney Cells (MDCK) in the Presence of Trypsin." *Medical microbiology and immunology* 162(1): 9–14.

Tong, S., Y. Li, P. Rivailler, C. Conrardy, D. a. a. Castillo, L.-M. Chen, S. Recuenco, J. a. Ellison, C. T. Davis, I. a. York, a. S. Turmelle, D. Moran, S. Rogers, M. Shi, Y. Tao, M. R. Weil, K. Tang, L. a. Rowe, S. Sammons, X. Xu, M. Frace, K. a. Lindblade, N. J. Cox, L. J. Anderson, C. E. Rupprecht, and R. O. Donis. 2012. "A Distinct Lineage of Influenza A Virus from Bats." *Proceedings of the National Academy of Sciences*: 5–10.

Tong, Suxiang, Xueyong Zhu, Yan Li, Mang Shi, Jing Zhang, Melissa Bourgeois, Hua Yang, Xianfeng Chen, Sergio Recuenco, Jorge Gomez, Li-Mei Chen, Adam Johnson, Ying Tao, Cyrille Dreyfus, Wenli Yu, Ryan McBride, Paul J Carney, Amy T Gilbert, Jessie Chang, Zhu Guo, Charles T Davis, James C Paulson, James Stevens, Charles E Rupprecht, Edward C Holmes, Ian A Wilson, and Ruben O Donis. 2013. "New World Bats Harbor Diverse Influenza A Viruses." *PLoS Pathog* 9(10): e1003657.

Towbin, H, T Staehelin, and J Gordon. 1979. "Electrophoretic Transfer of Proteins from Polyacrylamide Gels to Nitrocellulose Sheets: Procedure and Some Applications." *Proceedings of the National Academy of Sciences of the United States of America* 76(9): 4350–54.

Trinkle-Mulcahy, Laura, Séverine Boulon, Yun Wah Lam, Roby Urcia, F.-M. Boisvert, Franck Vandermoere, Nick a. Morrice, Sam Swift, Ulrich Rothbauer, Heinrich Leonhardt, and Angus Lamond. 2008. "Identifying Specific Protein Interaction Partners Using Quantitative Mass Spectrometry and Bead Proteomes." *The Journal of Cell Biology* 183(2): 223–39.

Tripathi, S, J Batra, W Cao, K Sharma, J R Patel, P Ranjan, A Kumar, J M Katz, N J Cox, and R B Lal. 2013. "Influenza A Virus Nucleoprotein Induces Apoptosis in Human Airway Epithelial Cells: Implications of a Novel Interaction between Nucleoprotein and Host Protein Clusterin." *Cell death & disease* 4(3): e562.

Uetz, Peter, Yu-An Dong, Christine Zeretzke, Christine Atzler, Armin Baiker, Bonnie Berger, Seesandra V Rajagopala, Maria Roupelieva, Dietlind Rose, Even Fossum, and Jürgen Haas. 2006. "Herpesviral Protein Networks and Their Interaction with the Human Proteome." *Science (New York, N.Y.)* 311(5758): 239–42.

Uyeki, Timothy M, and Nancy J Cox. 2013. "Global Concerns Regarding Novel Influenza A (H7N9) Virus Infections." *New England Journal of Medicine* 368(20): 1862–64.

Vasin, A V, O A Temkina, V V Egorov, S A Klotchenko, M A Plotnikova, and O I Kiselev. 2014. "Molecular Mechanisms Enhancing the Proteome of Influenza A Viruses: An Overview of Recently Discovered Proteins." *Virus research* 185: 53–63.

Vijaykrishna, D, L L M Poon, H C Zhu, S K Ma, O T W Li, C L Cheung, G J D Smith, J S M Peiris, and Y Guan. 2010. "Reassortment of Pandemic H1N1/2009 Influenza A Virus in Swine." *Science* 328 (5985): 1529.

Viswanathan, V K, Athanasia Koutsouris, Sandra Lukic, Mark Pilkinton, Ivana Simonovic, Miljan Simonovic, and Gail Hecht. 2004. "Comparative Analysis of EspF from Enteropathogenic and Enterohemorrhagic Escherichia Coli in Alteration of Epithelial Barrier Function." *Infection and Immunity* 72 (6): 3218–27.

Wang, Qian, Xiao-Wei Chen, and Ben Margolis. 2007. "PALS1 Regulates E-Cadherin Trafficking in Mammalian Epithelial Cells." *Molecular Biology of the Cell* 18 (3): 874–85.

Wang, Zhongfang, Anli Zhang, Yanmin Wan, Xinian Liu, Chao Qiu, Xiuhong Xi, Yanqin Ren, Jing Wang, Yuan Dong, and Meijuan Bao. 2014. "Early Hypercytokinemia Is Associated with Interferon-Induced Transmembrane Protein-3 Dysfunction and Predictive of Fatal H7N9 Infection." *Proceedings of the National Academy of Sciences* 111(2): 769–74.

Watanabe, Ken, Takayuki Fuse, Issay Asano, Fujiko Tsukahara, Yoshiro Maru, Kyosuke Nagata, Kaio Kitazato, and Nobuyuki Kobayashi. 2006. "Identification of Hsc70 as an Influenza Virus Matrix Protein (M1) Binding Factor Involved in the Virus Life Cycle." *FEBS letters* 580(24): 5785–90.

Watanabe, Tokiko, Eiryo Kawakami, Jason E. Shoemaker, Tiago J.S. Lopes, Yukiko Matsuoka, Yuriko Tomita, Hiroko Kozuka-Hata, Takeo Gorai, Tomoko Kuwahara, Eiji Takeda, Atsushi Nagata, Ryo Takano, Maki Kiso, Makoto Yamashita, Yuko Sakai-Tagawa, Hiroaki Katsura, Naoki Nonaka, Hiroko Fujii, Ken Fujii, Yukihiko Sugita, Takeshi Noda, Hideo Goto, Satoshi Fukuyama, Shinji Watanabe, Gabriele Neumann, Masaaki Oyama, Hiroaki Kitano, and Yoshihiro Kawaoka. 2014. "Influenza Virus-Host Interactome Screen as a Platform for Antiviral Drug Development." *Cell Host & Microbe*: 1–11.

Watanabe, Tokiko, Shinji Watanabe, and Yoshihiro Kawaoka. 2010. "Cellular Networks Involved in the Influenza Virus Life Cycle." *Cell host & microbe* 7(6): 427–39.

Webster, R, W Bean, O Gorman, T Chambers, and Y Kawaoka. 1992. "Evolution and Ecology of Influenza A Viruses." *Microbiological reviews* 56(1): 152–79.

Webster, R G, D J Hulse-Post, K M Sturm-Ramirez, Y Guan, M Peiris, G Smith, and H Chen. 2007. "Changing Epidemiology and Ecology of Highly Pathogenic Avian H5N1 Influenza Viruses." *Avian Diseases* 51(s1): 269–72.

Weekes, Michael P., Peter Tomasec, Edward L. Huttlin, Ceri a. Fielding, David Nusinow, Richard J. Stanton, Eddie C Y Wang, Rebecca Aicheler, Isa Murrell, Gavin W G Wilkinson, Paul J. Lehner, and Steven P. Gygi. 2014. "Quantitative Temporal Viromics: An Approach to Investigate Host-Pathogen Interaction." *Cell* 157(6): 1460–72.

Wilhelm, Mathias, Judith Schlegl, Hannes Hahne, Amin Moghaddas Gholami, Marcus Lieberenz, Mikhail M Savitski, Emanuel Ziegler, Lars Butzmann, Siegfried Gessulat, Harald Marx, Toby Mathieson, Simone Lemeer, Karsten Schnatbaum, Ulf Reimer, Holger Wenschuh, Martin Mollenhauer, Julia Slotta-Huspenina, Joos-Hendrik Boese, Marcus Bantscheff, Anja Gerstmair, Franz Faerber, and Bernhard Kuster. 2014. "Mass-Spectrometry-Based Draft of the Human Proteome." *Nature* 509(7502): 582–87.

de Wit, Emmie, Vincent J Munster, Debby van Riel, Walter E P Beyer, Guus F Rimmelzwaan, Thijs Kuiken, Albert D M E Osterhaus, and Ron a M Fouchier. 2010. "Molecular Determinants of Adaptation of Highly Pathogenic Avian Influenza H7N7 Viruses to Efficient Replication in the Human Host." *Journal of virology* 84(3): 1597–1606.

Worobey, Michael, Guan-Zhu Han, and Andrew Rambaut. 2014. "Genesis and Pathogenesis of the 1918 Pandemic H1N1 Influenza A Virus." *Proceedings of the National Academy of Sciences* 111 (22): 8107–12.

Wu, Chung-Yi, King-Song Jeng, and Michael M-C Lai. 2011. "The SUMOylation of Matrix Protein M1 Modulates the Assembly and Morphogenesis of Influenza A Virus." *Journal of virology* 85(13): 6618–28.

van Wyke, K L, J W Yewdell, L J Reck, and B R Murphy. 1984. "Antigenic Characterization of Influenza A Virus Matrix Protein with Monoclonal Antibodies." *Journal of virology* 49(1): 248–52.

York, Ashley, Edward C Hutchinson, and Ervin Fodor. 2014. “Interactome Analysis of the Influenza A Virus Transcription/Replication Machinery Identifies Protein Phosphatase 6 as a Cellular Factor Required for Efficient Virus Replication.” *Journal of Virology* 88(22): 13284–99.

Zhang, Jie, and Robert A. Lamb. 1996. “Characterization of the Membrane Association of the Influenza Virus Matrix Protein in Living Cells.” *Virology* 225(2): 255–66.

Zhang, Leiliang, Nancy Y. Villa, Masmudur M. Rahman, Sherin Smallwood, Donna Shattuck, Chris Neff, Max Dufford, Jerry S. Lanchbury, Joshua LaBaer, and Grant McFadden. 2009. “Analysis of Vaccinia Virus-Host Protein-Protein Interactions: Validations of Yeast Two-Hybrid Screenings.” *Journal of Proteome Research* 8(9): 4311–18.

Zhirnov, O P, and H D Klenk. 1997. “Histones as a Target for Influenza Virus Matrix Protein M1.” *Virology* 235(2): 302–10.

Zihni, Ceniz, Maria S Balda, and Karl Matter. 2014. “Signalling at Tight Junctions during Epithelial Differentiation and Microbial Pathogenesis.” *Journal of Cell Science* 127(16): 3401–13.

Vita

Garrett Poshusta was born in the foothills of the Rocky Mountains in Choteau, Montana but grew up in in the foothills of the Cascades in Ellensburg, Washington. Garrett attended Gonzaga University, in Spokane, where he majored in Chemistry and was a member of the inaugural class of the Hogan Entrepreneurial Leadership Program. At Gonzaga Garrett conducted research on metabolism in *Rhodospirillum rubrum* under the direction of Dr. Jennifer Shepherd. Following his undergraduate studies, Garrett moved to Seattle where he worked on vaccine and diagnostic development for tuberculosis, leishmaniasis, and leprosy at the Infectious Disease Research Institute under the supervision of Dr. Rhea Coler and Jeffrey Guderian. Garrett earned his Doctor of Philosophy degree from the University of Washington, in 2015 through the Biophysics, Structure and Design Program and the Department of Biochemistry.