RESPIRATORY SYNCYTIAL VIRUS VACCINE DESIGN USING STRUCTURE BASED MACHINE LEARNING MODELS

by

Thomas C. McCarty
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Bioinformatics and Computational Biology

Committee:

_____	Dr. Iosif Vaisman, Committee Chair
_____	Dr. Saleet Jafri, Committee Member
_____	Dr. Peter Collins, Committee Member
_____	Dr. Dmitri Klimov, Committee Member
_____	Dr. Iosif Vaisman, Director, School of Systems Biology
_____	Dr. Donna M. Fox, Associate Dean, Office of Student Affairs & Special Programs, College of Science
_____	Dr. Fernando R. Miralles-Wilhelm, Dean, College of Science
Date: _____	Spring Semester 2021 George Mason University Fairfax, VA

Respiratory Syncytial Virus Vaccine Design Using Structure Based machine learning models

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Thomas C. McCarty

Director: Iosif Vaisman, Professor
School of Systems Biology

Spring Semester 2021
George Mason University
Fairfax, VA

DEDICATION

This is dedicated to my wonderful wife Ofelia, my children Natalie, Jacqueline, Shannon, and Hannah. I would also dedicate this to my parents Mary and Tom, brothers Kevin and Brian and sisters Mary and Clare.

ACKNOWLEDGEMENTS

I would like to thank the many friends, relatives, and supporters who have made this happen. I would like to thank all the professionals I know and work with who continue to exemplify how and why scientific discovery is accomplished. In particular I wish to thank all my colleagues in the RNA Virus Section of the Laboratory of Infectious Diseases (LID) at the National Institutes of Allergy and Infectious Diseases (NIAID), part of the National Institutes of Health (NIH). I would also like to thank the faculty and staff in the School of Systems Biology in the College of Science at George Mason University, particularly my committee members, Dr. Iosif Vaisman, Dr. Saleet Jafri, Dr. Dmitri Klimov as well as Ms. Kimberly Harris, Ms. Monique Sweeney, and Mrs. Diane St. Germain (retired). Lastly and most importantly I would like to express my deepest gratitude to Dr. Peter Collins, retired section head of RNA Virus Section, LID/NIAID/NIH who both encouraged and supported this work and without whose assistance this this work would not be possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

Alpha carbon ..... Cα
Amino acid ..... aa
Arginine ..... R
Aspartic acid ..... D
Attachment glycoprotein ..... G
AUTOmated server for predicting functional consequences of amino acid MUTations in protEins ..... AUTO-MUTE 2.0
Basic Local Alignment Search Tool ..... BLAST
Beta carbon ..... Cß
BLOcks SUbstitution Matrix sequences with ≥ 62% identity BLOSUM62
Bovine respiratory syncytial virus ..... bRSV
C-C motif chemokine ligand 5 ..... CCL5
Carbon dioxide ..... CO_2
Cluster analysis of the pairwise alignment; 3rd generation ..... ClustalW
Clustering Approach to Identify Near-Native Protein Folds ..... SPICKER
Coefficient of determination ..... R^2
Combined reverse transcriptase and quantitative polymerase chain reaction ..... qRT-PCR
Comparative Ct method ..... $2^{-\Delta\Delta Ct}$ method
Complementary DNA ..... cDNA
Confidence score ..... C-score
Cycle threshold ..... Ct
Dalton ..... Da
Degrees Celsius ..... °C
Deoxyribonuclease ..... DNase
Deoxyucleoside triphosphate mix ..... dNTP
Dithiothreitol ..... DTT
Dynamic programming protein alignment ..... Needleman-Wunsch
Elution buffer ..... EB
Enhanced green fluorescent gene/protein ..... eGFP
Enzyme-linked immunosorbent assay ..... ELISA
Fetal bovine sera ..... FBS
Fusion protein ..... F
GenBank sequence database ..... GenBank
Gene-end ..... GE
Gene-start ..... GS
Glutamine ..... Q
Glycosaminoglycans ..... GAG
Gonnet substitution matrix ..... Gonnet
Hendra virus ..... HeV

Human parainfluenza virus 3 ..... hPIV3
Human respiratory syncytial virus ..... RSV
Human respiratory syncytial virus with NS1 gene deletion dNS1
IFN-stimulated response elements ..... ISRE
Interferon beta 1 ..... IFNB1
Interferon lambda 1 ..... IFNL1
Interferon regulatory factor 9 ..... IFR-9
Interferon stimulated gene ..... ISG
Interferon-stimulated gene factor 3 ..... ISGF-3
Iterative Threading ASSEmbly Refinement ..... I-TASSER
Janus kinase ..... Jak
Large polymerase L
Leader ..... Le
Leucine L
Lower respiratory tract illness ..... LRI
M2-1 protein M2-1
M2-2 protein M2-2
Magnesium chloride ..... $MgCl_2$
Matrix protein ..... M
Measles ..... MeV
Messenger RNA ..... mRNA
Microgram ..... mg
Microliter ..... ul
Micrometer ..... um
Micromolar ..... uM
Milliliter ..... ml
Millimolar ..... mM
Mitochondrial antiviral-signaling ..... MAVS
Monoclonal antibody ..... mAb
Multiplicity of infection ..... MOI
Mumps ..... MuV
Murine pneumonia virus ..... MPV
Nanogram ..... ng
Nanometer ..... nm
Nanomolar ..... nM
Nipah virus ..... NiV
Nonstructural protein 1 ..... NS1
Nonstructural protein 2 ..... NS2
Normalized score ..... Z-score
Nucleoprotein ..... N
Open reading frame ..... ORF
Phosphoprotein ..... P
Picogram ..... pg
Polymerase chain reaction ..... PCR

Pre fusion F protein ..... preF
Protein Data Bank ..... PDB
Protein kinase RNA-activated ..... PKR
Protein-protein BLAST ..... BLASTP
Qualitative Model Energy ANalysis ..... QMEAN
Quantitative polymerase chain reaction ..... qPCR
Ras homolog family member A ..... RhoA
Retinoic acid-inducible gene I ..... RIG-I
Reverse transcriptase ..... RT
Ribonuclease H ..... RNase H
Ribonucleic acid ..... RNA
Ribosomal RNA ..... rRNA
RNA dependent RNA polymerase ..... RdRP
Root-mean-square-deviation ..... RMSD
Signal transducer and activator of transcription 1 ..... STAT1
Signal transducer and activator of transcription 2 ..... STAT2
Small hydrophobic protein ..... SH
Stability changes ..... ddG
Three-dimensional 3D
Trailer ..... Tr
Tryptophan ..... W
Tumor necrosis factor alpha ..... TNFα
Uniform resource locator ..... URL
Unit ..... U
Uracil-DNA glycosylase ..... UNG
Water ..... H_2O

ABSTRACT

RESPIRATORY SYNCYTIAL VIRUS VACCINE DESIGN USING STRUCTURE BASED MACHINE LEARNING MODELS

Thomas C. McCarty, PhD

George Mason University, 2021

Dissertation Director: Dr. Iosif Vaisman

Human Respiratory Syncytial Virus (RSV) disease remains a global concern for both young and old [1] and to date no approved vaccine capable of reducing or preventing RSV disease exists. Numerous vaccine candidates currently being tested [2,3] are legacies of classic vaccine design approaches. Current vaccine design underutilizes modern computational strategies to help identify or optimized new vaccine candidates. A combination of computational and traditional methodologies should be expected to optimize target selection and candidate prioritization. The fields of computational protein folding/function prediction and systems biology [4] allow researchers to develop virtual protein designs and then create them physically for functional testing. Systematically prioritized vaccine target/candidates can then be produced using more traditional virus generation strategies [5] and evaluated for efficacy.

Focusing on a specific RSV protein, nonstructural protein 1 (NS1), this work demonstrates a process for design, development, and testing of novel virus vaccines. NS1 of RSV is unique in that it is the initial virus encoded gene [6] and is the most abundant viral protein produced by RSV during viral infection [7], yet incompletely characterized [8]. Although there exist related viruses in other species (Cow, goat, mice, sheep) [9], there remain no known human virus relatives to either the NS1 gene sequence itself or the RSV viral genome organization. It has been reported that patients possess serum antibodies against NS1 [10], however these antibodies are likely not part of a protective host response to eliminate or reduce virus/disease. Therefore, NS1 is an appropriate target protein for computational analysis due to its significance as an evolutionary unique protein retained in the most proximal position in the human RSV virus genome.

NS1 of RSV has been associated with various processes including host cell apoptosis [11], interferon and interferon simulated gene regulation (and related processes) [12,13], and viral genome replication [14], but no confirmed mechanism detailing NS1 physical association with various interacting partners. The fact that NS1 is 1) a nonstructural cytosolic protein (i.e. not part of any known virus structure) and 2) only recently described in a crystal structure [15], contribute to its difficult study. NS1 has also been shown to be disposable such that an RSV virus in which the NS1 gene sequence has been deleted is viable [16] and a leading live attenuated RSV vaccine virus candidate (ClinicalTrials.gov Identifier: NCT03596801; RSV 6120/ΔNS1 and RSV 6120/F1/G2/ΔNS1) [17].

Balancing the level of virus growth (not too high, but high enough to invoke the appropriate immune response) and the level of antigenicity required for appropriate immune system induction [3] is complicated further when deletion vaccines are compared to full length wild-type virus since the initial growth conditions are not equivalent, one virus is shorter than the other. NS1 protein seems to participate in multiple measurable biologic phenomena that could be used as measurable functional readouts. Therefore, an analysis of mutation changes to the NS1 protein will provide a selection of vaccine candidates capable of reaching a variety of growth/antigenicity optima.

CHAPTER 1. LITERATURE REVIEW: HUMAN RESPIRATORY SYNCYTIAL VIRUS (RSV)

1.1 Introduction

Human respiratory syncytial virus (RSV) is considered the most important viral agent of lower respiratory tract illness (LRI) in infants worldwide [1] and an important cause of LRI in elderly and severely immunocompromised persons [18]. Globally RSV has been estimated to cause more than 3.5 million hospitalizations and 66,000 to 199,000 deaths annually [19]. Despite the overwhelming need for disease treatment, RSV has eluded the development of a safe and cost-effective vaccine. Since the 1970's researchers have struggled with a number of difficulties including 1) enhanced disease [20–22], 2) suboptimal balance between attenuation and immunogenicity [23] and 3) natural genomic instability of RNA viruses [24]. As a result, RSV remains an unvaccinated disease of worldwide importance.

1.2 Taxonomy and classification

RSV is a member of the genus Orthopneumovirus of the recently reclassified family Pneumoviridae [25] in the order Mononegavirales. This family is currently composed of the two genera, Orthopneumovirus and Metapneumovirus. Orthopneumovirus genus is exemplified by RSV, but also includes bovine RSV (bRSV) as well as murine pneumonia virus (MPV) (Figure 1). RSV has a single-stranded, nonsegmented, negative-sense RNA genome ranging in size from 15,191 to 15,277

nucleotides in length [18]. In addition, RSV can be classified into two antigenic and sequence based groups, A and B [26], based predominately on the sequence variation in the attachment glycoprotein (G).

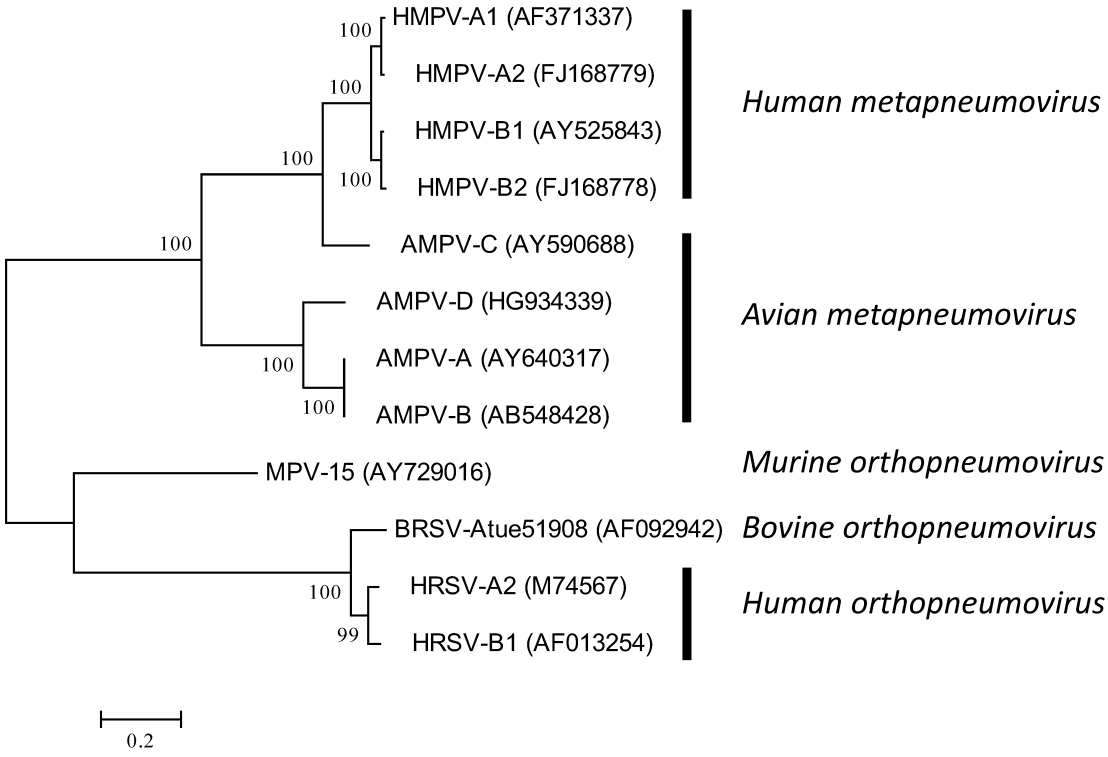


Figure 1 Pneumoviridae Phylogenic Distribution (aligned L amino acid sequences)
Multiple sequence alignment of L sequences from representative members of *Pneumoviridae* family. The percentage of trees in which the associated taxa cluster together is shown for values over 70%. Vertical bars (Black; right side) indicate species assignments of virus isolates [27,28].

Other families within the order Mononegavirales include highly infectious human pathogens such as measles (MeV) [29], mumps (MuV) [30], human parainfluenza virus 3 (hPIV3) [31–34]; as well as deadly zoonotic viruses Hendra virus (HeV) [35] and Nipah

virus (NiV) [36]. As there remains no effective vaccine to restrict or eliminate some of these viral pathogens, Pneumoviridae (and Paramyxoviridae) family members will continue to be responsible for a majority of respiratory tract diseases that result in significant burdens to global economic and health concerns [37].

1.3 Epidemiology and clinical features

Human respiratory syncytial virus (RSV) remains the most significant cause of pediatric respiratory infections and results in 90,000 emergency hospitalizations in the U.S. and 160,000 deaths worldwide each year, mostly in the developing world where access to supportive healthcare is limited [1]. Infants experiencing other medical conditions are at particularly high risk for developing severe RSV disease [38,39]. RSV infects individuals repeatedly throughout life [40] and causes significant disease in the elderly [41–43]. RSV infection exhibits a seasonal distribution typical of other viral induced lower respiratory disease, peaking in mid-winter and fewer illnesses during mid-summer [44]. The clinical features of RSV disease, particularly in children, appear as bronchiolitis and pneumonia [45,46]. In the elderly, the presentation of cough, wheezing, low-grade fever and nasal congestion due to RSV is less distinctive when influenza is cocirculating [41].

1.4 RSV virion structure and genome organization

RSV virions are found in either spherical or filamentous forms, but with a substantial range within these forms [47]. Filamentous particles have been measured between 70 to 190 nm in diameter and up to 2 um long. Spherical particles are

considered to be between 80 and 140 nm (but some forms are found sized at 250 to 600 nm) (Figure 2) [47].

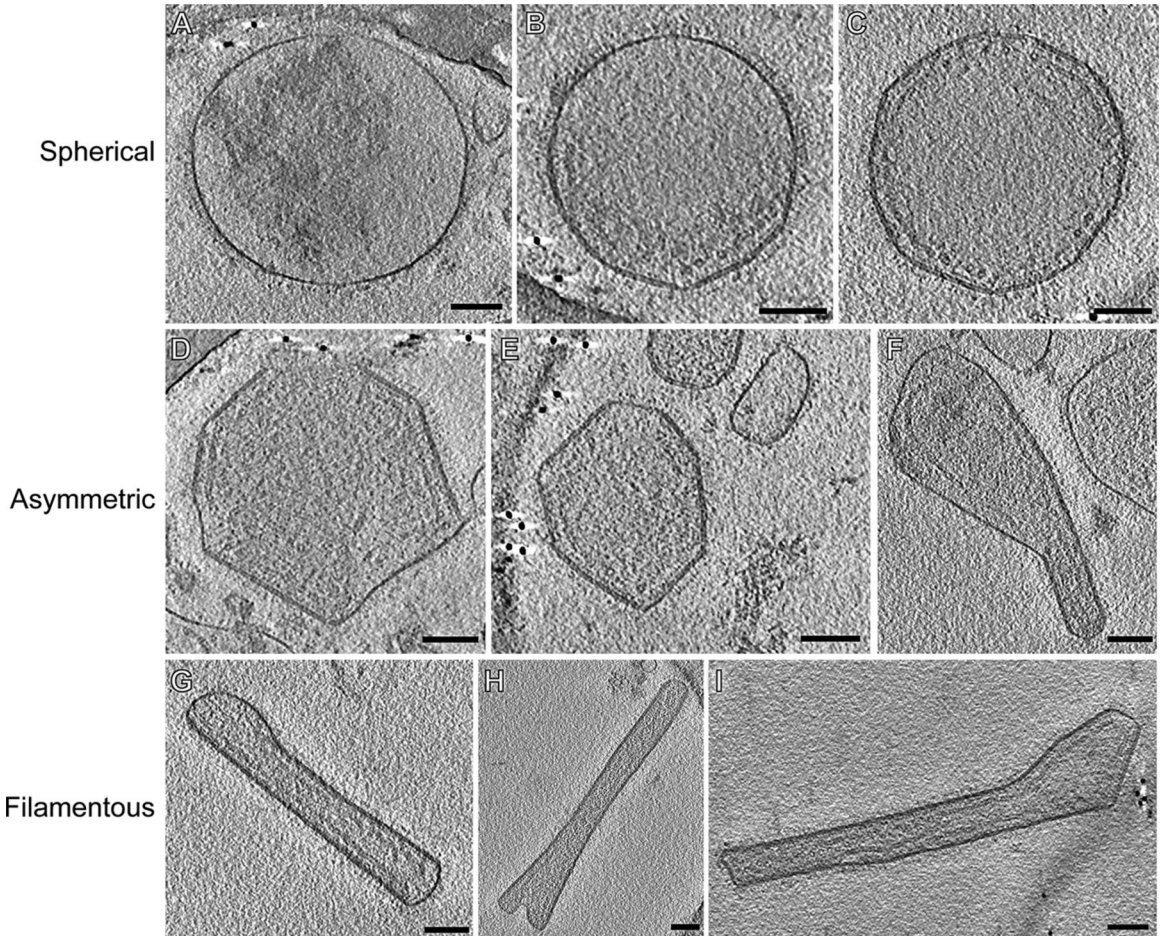


Figure 2 RSV Virion Morphologies
Morphologic characterization of three RSV shapes (Figure 2; [48]). Virions in panels A-C are spherical, panels D-F are asymmetric and panels G-I are filamentous. Images are 7.5 nm central slices from tomographic reconstruction. Horizontal bars (Black; bottom right) represent 100 nm.

RSV is a nonsegmented single-stranded negative-sense RNA virus. The viral genome is approximately 15.2 kb in length, containing 10 genes coding for 11 open

reading frames (ORFs) [18]. The genome has the following gene organization; leader (Le) sequence, nonstructural protein 1 (NS1), nonstructural protein 2 (NS2), nucleoprotein (N), phosphoprotein (P), matrix (M), small hydrophobic protein (SH), attachment glycoprotein (G), fusion glycoprotein (F), M2 (expressing M2-1 and M2-2 proteins encoded by two overlapping open reading frames withing the M2 gene), large (L) polymerase, and trailer (Tr) sequence are arranged in the order: 3'-Le-NS1-NS2-N-P-M-SH-G-F-M2-L-Tr5' [18]. Each gene begins with a highly conserved 9 nucleotide gene-start (GS) transcriptional signal and ends with a less conserved 12 to 13 nucleotide gene end (GE) signal (Figure 3) [49,50].

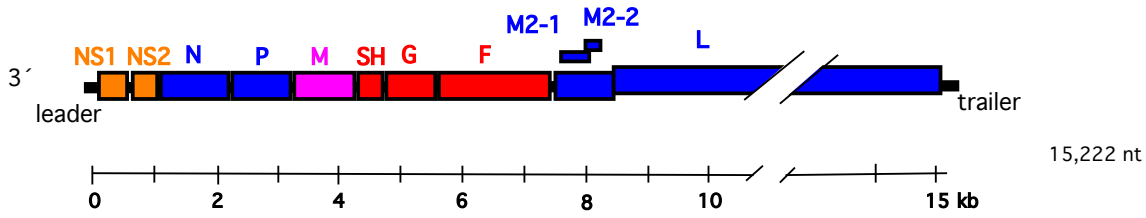


Figure 3 RSV Genome Organization
Negative-sense RNA genome (strain A2) depicted 3' to 5' showing the extragenic 3' leader (le) and 5' trailer (tr) regions and the intervening 10 viral genes (colored rectangles) that are each expressed as a separate mRNA. M2-1 and M2-2 are overlapping open reading frames of the M2 mRNA. The M2 and L gene coding regions overlap slightly. L mRNA is expressed by polymerase backtracking [51].

The RSV virion contains a host cell derived lipid envelope surrounding the M protein layer to which the RNA dependent RNA polymerase (RdRP) complex is linked by M2-1 [47]. Three surface glycoproteins, G, F, and SH form external projections from the host derived lipid envelope (Figure 4) [47,52]. Inside the virion, viral RNA (both

genomic and antigenomic) is coated (encapsidated) with N protein that protects the RNA from degradation and shields it from host cell pattern recognition receptors that can initiate innate immune responses. The encapsidated RNA genome (and RNA antigenome) further associates with P, L and M2-1 to form the RNA dependent RNA polymerase (RdRP) complex [18]. The RdRP complex associated with viral genome is a source for 1) templated message RNA (mRNA) transcription and 2) antigenome synthesis. The RdRP complex associated with antigenome allows replication of full length viral genome [53,54].

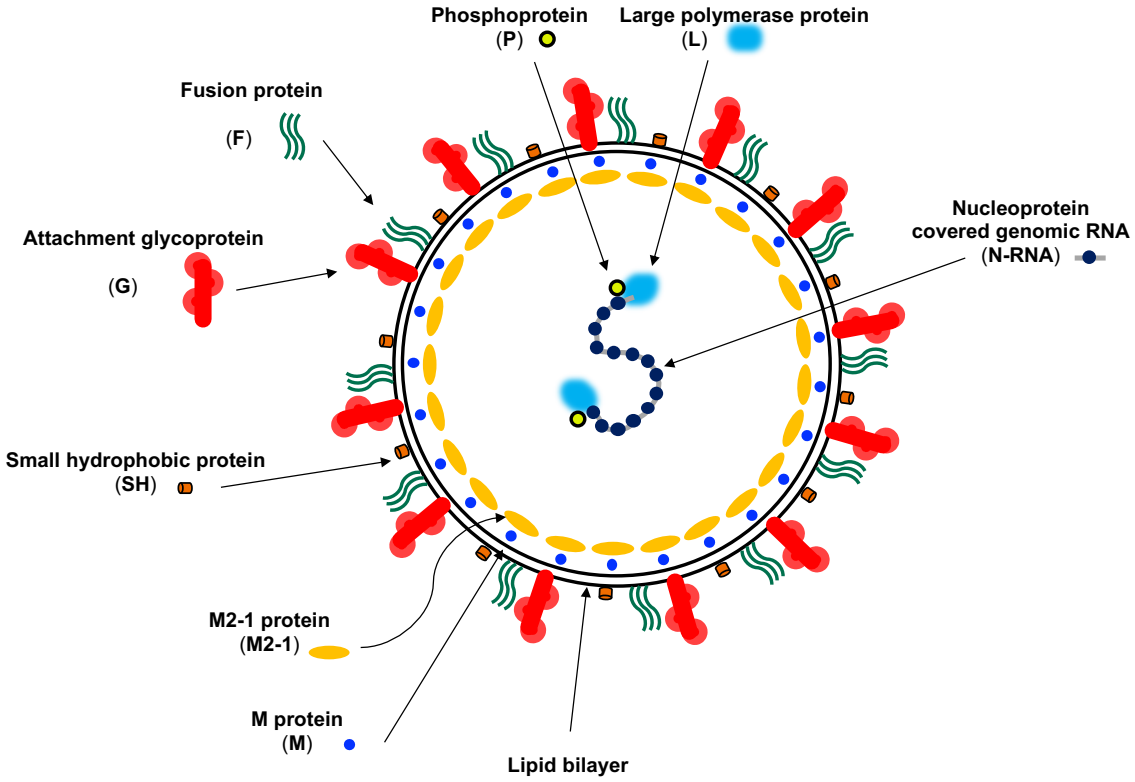


Figure 4 RSV Virion Structure
Graphic representation of spherical RSV virion. Negative-sense RNA genome (Grey line) is shown encapsidated by nucleoprotein (N; Dark Blue circles) with phosphoprotein (P; Light Green/Dark Blue bordered circle) and large polymerase protein (L; blurred Blue circle) attached at the genome termini. A cell derived lipid bilayer (double Black lines) are internally lined with matrix (M; Blue circle) and M2-1 (Yellow oval). The three glycoproteins, small hydrophobic (SH; Red cylinder), attachment glycoprotein (G; Red) and fusion peptide (Green curved lines) are each cell membrane anchored and externally facing.

1.5 RSV replication cycle

Initial attachment and uptake of extracellular RSV is suggested to occur by either micropinocytosis [55,56] or fusion at the plasma membrane [57] leading to subsequent RSV F glycoprotein mediated fusion of the virus envelope with the plasma membrane [58,59]. Once the virus is in the cytoplasm the nucleoprotein coated genome is released. The L polymerase (part of the RdRP complex) then encounters the viral genome at its 3' end and viral genes are transcribed into mRNAs in a start-stop-restart synthesis

mechanism guided by the gene-start (GS) and gene-end (GE) sequences [60]. This creates a polar transcription gradient in which genes at the 3' Leader (Le) end of the negative sense genome are transcribed more frequently than the genes which are downstream [4,61]. Also generated is a complete positive-sense RNA complement called the antigenome, which acts as a template for full length negative sense genome synthesis [62–64]. The genome and the antigenome are both coated with N proteins during synthesis, and the encapsidated forms serve as RNA synthesis templates [65,66]. The M protein regulates the assembly of the RSV particle by interacting with the intracellular domains of the surface glycoproteins F and G as well as with the nucleoprotein proteins N, P, and M2-1 [67]. Viral particle budding which is primarily filamentous *in vitro*, occurs after these newly synthesized proteins plus genome self-assemble at the inner cell surface allowing virus to acquire a lipid envelope from the infected-cell plasma membrane [47,68,69].

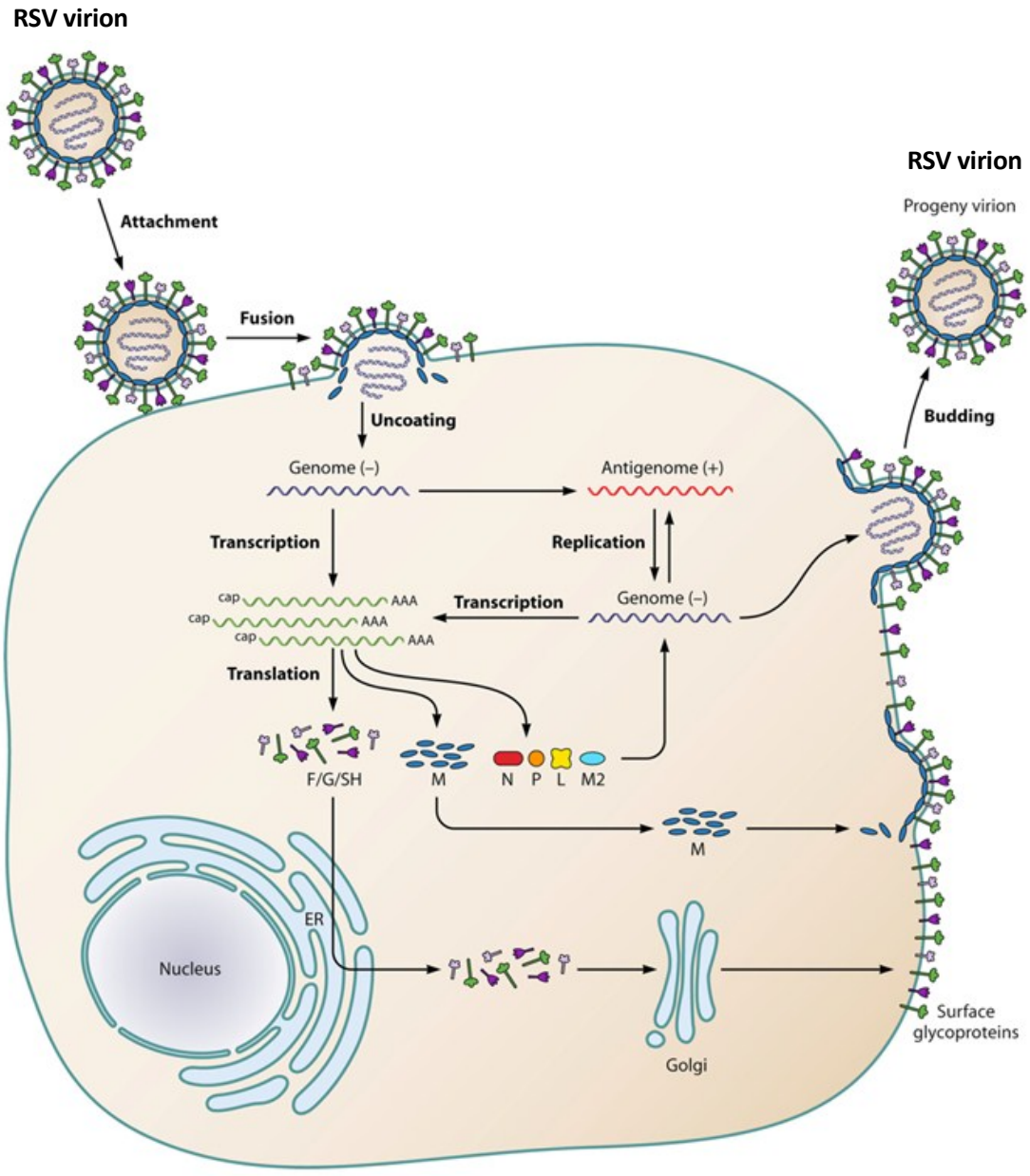


Figure 5 RSV Replication Cycle
Graphic representation of RSV virus replication including attachment, cytosolic viral gene expression/genome replication, virus assembly and virion budding.

1.6 RSV viral proteins

1.6.1. Nonstructural proteins 1 (NS1) and 2 (NS2)

NS1 and NS2 are small nonstructural proteins (139 and 124 amino acids in length respectively), encoded by the first and second genes in the RSV viral genome. They are unique to the Orthopneumovirus genus [25] as there are no known related sequence or structural counterparts in other human viruses. It is believed that NS proteins either singly or in larger protein complexes, target multiple signaling cascades related to the induction and signaling of type I and III interferons (IFNs) [13,15,70–73] that are key for cells to establish an antiviral state. Previous studies identified NS1 interactions with itself [74,75], NS2 [74], P [76] and M [74] proteins of RSV. A proteomic screen [77] revealed a list of cellular interacting partners for NS1. The NS1 protein has been shown to associate with MAVS (mitochondrial antiviral-signaling protein) in A549 cells, disrupting its interaction with RIG-I (retinoic acid-inducible gene I) [78–82], an essential early step in the induction of an antiviral response. This disruption interferes with downstream activation (translocation) of the IRF3 (interferon regulatory factor 3) and NFκB (nuclear factor kappa-light-chain-enhancer of activated B cells) [12,72,83,84] transcription factors that up regulate the expression of genes involved in antiviral activities including inflammation and apoptosis[11,12]. Separately, NS2 has been shown to bind to RIG-I, which similarly disrupts interaction with MAVS leading to an inhibition of IFN induction and IFN induced signaling pathways [83]. Lastly, NS1 and NS2 have been shown to interfere with cell apoptosis as well as dendritic and T lymphocyte cell responses [11,40,73,85]. With regard to replicating virus, RSV lacking the NS1 and/or

NS2 gene retains the ability to replicate in vitro and in vivo [16], although at reduced levels, and NS1 and NS2 deletion viruses presently are being evaluated in a clinical trial as candidate live vaccines (ClinicalTrials.gov Identifier: NCT03596801; RSV 6120/ΔNS1 and RSV 6120/F1/G2/ΔNS1 [17], ClinicalTrials.gov Identifier: NCT03387137; RSV 6120/ΔNS2/1030s [86].

1.6.2. Fusion glycoprotein (F)

The fusion glycoprotein (F) is one of three RSV surface glycoproteins (together with the G and SH proteins), is one of the two RSV neutralization as well as protective antigens (together with the G protein) [87–89] and is a major target of antiviral drug development. The F gene is located after G and before M2 in the RSV genome. As a type I integral membrane protein (single pass, C-terminally-proximal transmembrane anchor), the F0 precursor (574 amino acids) [90] is unable to cause fusion, but in the Golgi, trimerizes and processed by host cell proteases. Cleavage at two different furin-protease cleavage sites leads [91,92] to an F1 (438 amino acids) and F2 (87 amino acids) disulfide-linked prefusion protein preF [93,94]. The F protein mediates viral penetration. When triggered (by a mechanism that is poorly understood), the preF protein undergoes major refolding including insertion of the F1 N-terminal fusion protein (FP) into target cell membrane and subsequent association of heptad repeat regions A (HRA) and B (HRB) into a six-helix bundle (6HB) inducing fusion of viral and host cell membranes [95].

1.6.3. Attachment glycoprotein (G)

RSV G is an RSV surface glycoprotein that, with F, is one of the two virus neutralization and protective antigens. Compared to RSV F, RSV G is both smaller (289-299 amino acids) and substantially more-heavily glycosylated. It is a type II glycoprotein, with an N-terminally-proximal transmembrane domain [96]. There also exists a secreted or soluble form of RSV G (sG) that is 65-74 amino acids shorter at the N-terminus [97,98] due to translation initiation occurring at an internal secondary AUG followed by proteolytic trimming [99]. Both the anchored and soluble G proteins include a pair of mucin-like regions of high sequence variability flanking a highly conserved central region containing two pairs of disulfide-bonded cystine residues that form a cystine noose overlapping a sequence region similar to the CX3C chemokine fractalkine [100]. The extensive glycosylation of G and its possible interference with the CX3CR1 receptor on neutrophils may reduce protective immune responses [101,102]. Recent studies indicate that the binding of RSV G protein to the fractalkine receptor is an important step in viral attachment [103,104]. Interestingly, RSV from which the G gene has been deleted can replicate efficiently *in vitro* but is highly attenuated in experimental animals and humans [105].

1.6.4. Small hydrophobic surface protein (SH)

The RSV small hydrophobic protein (SH) (64-65 amino acids in length) monomer is attached to cell membranes due to an N-terminus transmembrane anchor sequence [106] similar to RSV G. The role SH protein plays in RSV infection is not completely know but may be analogous to viroporins which form proton or ion channels affecting

among other things membrane permeability [107,108]. SH is not required for RSV replication [109]. Although deletion of SH is shown to influence plaque size as well as virus replication [109], in animal models the same deletion demonstrates conflicting results regarding upper and lower respiratory tract attenuation [109,110]. In tissue culture as well as non-animal experiments, the SH protein is found in other forms including oligomeric structures [106,111], N terminal truncation and post-translational modifications [112].

1.6.5. Matrix protein (M)

The 256 amino acid matrix protein (M) of RSV is associated with multiple aspects of the viral infection cycle involving viral protein/complex assembly [67] affecting both viral [113] and cellular transcription [114]. Although associated with various membrane containing structures in cells, RSV M does not have the appropriate length hydrophobic residue sequence that defines most transmembrane anchors. As such M is likely associated with membranes, particularly the inner face of infected cell plasma membranes via an electrostatic charge interaction but also through M protein association with cytoplasmic ends of RSV G and F proteins [115]. RSV M can bind both accumulated RNPs and cellular actin [116] presumably to complete virus particle assembly. Other aspects of RSV M protein biology such as the role of its phosphorylation [96,115] or nuclear localization [114] during RSV infection remain unclear.

1.6.6. M2-1 and M2-2 proteins

RSV M2 gene codes for two novel RNA synthesis factors, M2-1 and M2-2, unique to RSV produced by a termination-dependent re-initiation mechanism [117,118]. M2-1 (194 amino acids) forms a homotetramer [119] that among other things binds to RSV P [120], N [121], M [122] proteins and viral RNA [123] and essential for viral replication [124,125]. A result of M2-1 interactions with these various molecules lead to its description as a polymerase processivity factor [125,126] affecting both viral gene transcription and viral genome replication [127]. The M2-2 protein is one of the smallest RSV proteins at 88 or 90 amino acids in length depending which ATG start site is used [128]. Recombinant RSV with deletion of M2-2 is shown to have a biased increase in viral mRNA synthesis at the expense of reduced genome replication [127].

1.6.7. Nucleoprotein (N)

The RSV nucleoprotein (N) is required for complete encapsidation of genomic and antigenomic viral RNA. The N protein is abundantly expressed early in infection and found to associate with viral genome with the idea that this masking will minimize detection of viral RNA by various host cell pattern recognition receptors [129]. This 391 amino acid nucleoprotein (N) is also known to associate with phosphoprotein (P) [130] during the formation of the RdRP complex. The N protein has also been shown to antagonize the cell innate immune response by interacting with protein kinase RNA-activated (PKR) [131].

1.6.8. Phosphoprotein (P)

The RSV phosphoprotein (P) is known to form a functional homotetramer as well as bind with N [132], L [133], M2-1 [122] and M [134] differentially in order to facilitate viral gene transcription [120,124,135] and viral genome replication. Much of the function of the 241 amino acid P protein is regulated through phosphorylation and oligomerization. As an essential co-factor of the large polymerase protein (L), the P protein interacts with newly synthesized N protein or with N on encapsidated viral RNA and in turn helps orient this with the L polymerase complex [130,136].

1.6.9. The large (L) polymerase protein

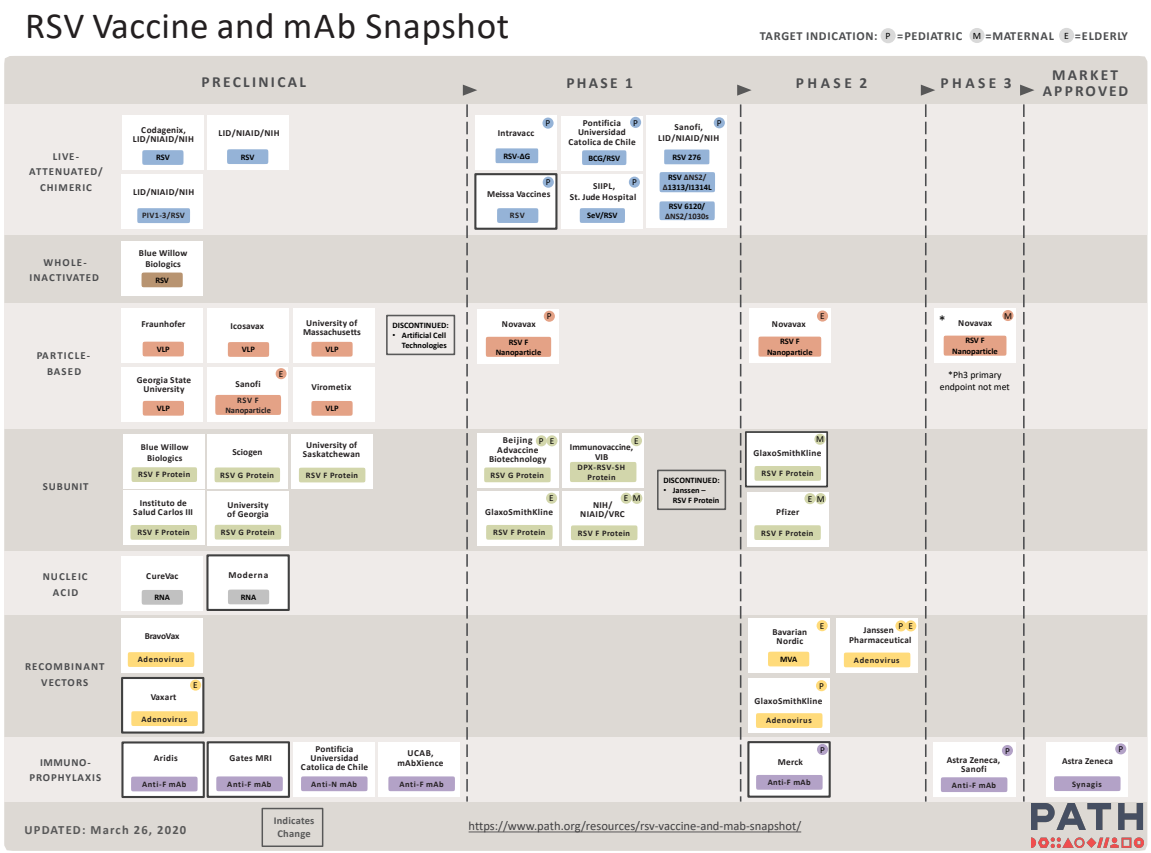
The large polymerase protein (L) is the virally encoded RNA dependent RNA polymerase (RdRP) responsible for both viral RNA transcription and viral RNA genome replication. As the largest (2165 amino acids) and least expressed viral protein, the polymerase activity of L requires intermittent coordination with N, P, M2-1 and M2-2 proteins. The polymerase will cap and polyadenylate viral mRNA but not viral genomes [137,138].

1.7 RSV vaccine development

Development of effective RSV vaccines have been hampered by the previous formalin-inactivated vaccine trial failure in the 1960's [20]. The realization that at least for naïve children (newborns), antigenic exposure for vaccination purposes is extremely context dependent. The deliver or RSV protein antigens alone without related host responses to replicating virus is not the proper context. Table 1 below lists the current

RSV vaccine and monoclonal antibody products as listed in the PATH database as of August 2019.

Table 1 RSV vaccine and monoclonal antibody snapshot
(https://www.path.org/resources/rsv-vaccine-and-mab-snapshot/)



Whereas most of RSV vaccine development has historically relied on more conventional methods, the introduction of reverse genetics offers greater precision in candidate engineering [4]. Yet, RSV remains as an unvaccinated viral disease of worldwide importance despite the substantial catalog of vaccine design strategies.

Confounding design improvements are the perpetual unknowns plaguing previous attempts, “what are the most appropriate targets of this precision engineering approach?” and “what are the most appropriate modifications of these target(s)?” This work introduces how computational biology and precise engineering (reverse genetics) can be united towards the goal of optimizing rationally designed vaccines.

CHAPTER 2. COMPUTATIONAL ALIGNMENT OF HUMAN RSV NONSTRUCTURAL PROTEIN 1 (NS1) AMINO ACID SEQUENCES

2.1 Abstract

Computational biology has helped drive forward the improvements of many biologic research fields [139,140]. In particular protein folding/protein modeling has benefitted dramatically as both computational hardware and algorithm quality improve [141–143]. Alignments of RSV NS1 sequences (both strain A and strain B as well as non-human RSV strains) were examined to see if there were any indications that certain sequences within NS1 are highly conserved or variable. These alignments indicated there was sequence information that would allow discrimination between strain A and strain B as well as from non-human NS1 virus sequences, however, there were no obvious sequence stretches that suggested conserved function (and therefore targets of modification).

2.2 Introduction

The need to improve virus vaccine candidate design is indisputable. There are historically more than 100 different RSV vaccine candidates alone [144–150]. The suggestion of introducing computational biology to improve the design process, is similarly not new. However, before relying on complicated methodology, it would seem informative to see whether a protein sequence alignment could reveal areas within a NS1 protein sequence that are obvious candidates for modification. Previous work uncovering protein structural motifs or active sites, from a primary amino acid sequence has helped guide this type of effort and is of great benefit when only the protein amino acid sequence

is available [151–157]. Therefore, by aligning the sequences it is feasible that any regions of interest would be revealed. This analysis should lead to protein structures that are 1) biologically interesting and 2) amenable to modification.

2.3 Materials and methods

2.3.1. NS1 amino acid sequence [GenBank: ALS35583.1]

```
MGSNSLSMIKVRLQNLFDNDEVALLKITCYTDKLIHLTNALAKAVIHTIKLNGIVFVHVITSSDI
CPNNNIVVKSNFTTMPVLQNGGYIWEMMELTHCSQPNGLLDDNCEIKFSKKLSDSTMTNYMNQLS
ELLGFDLNP
```

2.3.2. Basic Local Alignment Search Tool (BLAST) - blastp

A search for amino acid sequences matching NS1 sequence (GenBank: ALS35583.1) in a non-redundant protein sequence database was performed using the BLASTP program [155–157]. BLASTP is one of the many BLAST algorithms available through the BLAST website (https://blast.ncbi.nlm.nih.gov) which in turn is one of the most popular and easy to use sequence similarity bioinformatic search tools. There are more accurate/comprehensive sequence similarity search tools available, but BLAST is renowned for the speed at which it can search through very large numbers of sequences in a database and provide both statistical and visual outputs regarding the alignment between Query (input sequence) and returned matching sequences. The search returned 127 amino acid sequences that were identical or closely similar to some region of the input NS1 amino acid sequence (GenBank: ALS35583.1) of RSV. Specific algorithm parameters were set as follows:

General Parameters

- Max target sequences = 1000
- Short queries (Automatically adjust parameters for short input sequences)
- Expect threshold = 10
- Word size = 6
- Max matches in a query range = 0

Scoring Parameters

- Matrix = BLOSUM62
- Gap Costs = Existence: 11, Extension: 1
- Compositional adjustments = Conditional composition score matrix adjustment

Filters and Masking

- Filter (unchecked)
- Mask (unchecked)

After removing sequences that were incomplete or partial matches, 49 representative full length NS1 sequences remained.

2.3.4. CLUSTALW

The selected 49 amino acid sequences from the previous BLASTP search were aligned using ClustalW (version 2.1) [158–160] found in the MacVector software package (version 17.0.8) [161–163]. ClustalW continues to be oldest and one of the most widely used multiple sequence alignment programs available. Alignments are

progressive built up by merging larger and larger subalignments with the best scores satisfying the given gap penalties and amino acid weight matrix [164].

The default settings for ClustalW Multiple Alignment found in MacVector (version 17.0.8) were used to perform the alignment. These include Open Gap Penalty = 10, Extended Gap Penalty = 0.2 and an amino acid weight matrix = Gonnet.

2.4 Results

The BLASTP search returned 127 sequences that matched some portion of the NS1 query sequence (GenBank: ALS35583.1). After removing the sequences of low quality/partial matching, there remained 49 full length amino acid sequences that aligned meaningfully to NS1. Interestingly included in this set of well aligned full length NS1 sequence matches were sequences from non-human RSV strains. While these were not originally considered informative, they were retained in the event the multiple sequence alignment might reveal evolutionarily conserved/important sequence region(s). ClustalW was used to generate a multiple sequence alignment of these 49 NS1 amino acid sequences. The amino acid sequence alignment was interesting in that 1) the NS1 amino acid sequences aligned in groups agreeing with the RSV A or B serotypes, 2) revealed short stretches of amino acid sequences that accounted for these distinctions and 3) further revealed the amino acid differences between human and non-human RSV NS1 sequences.

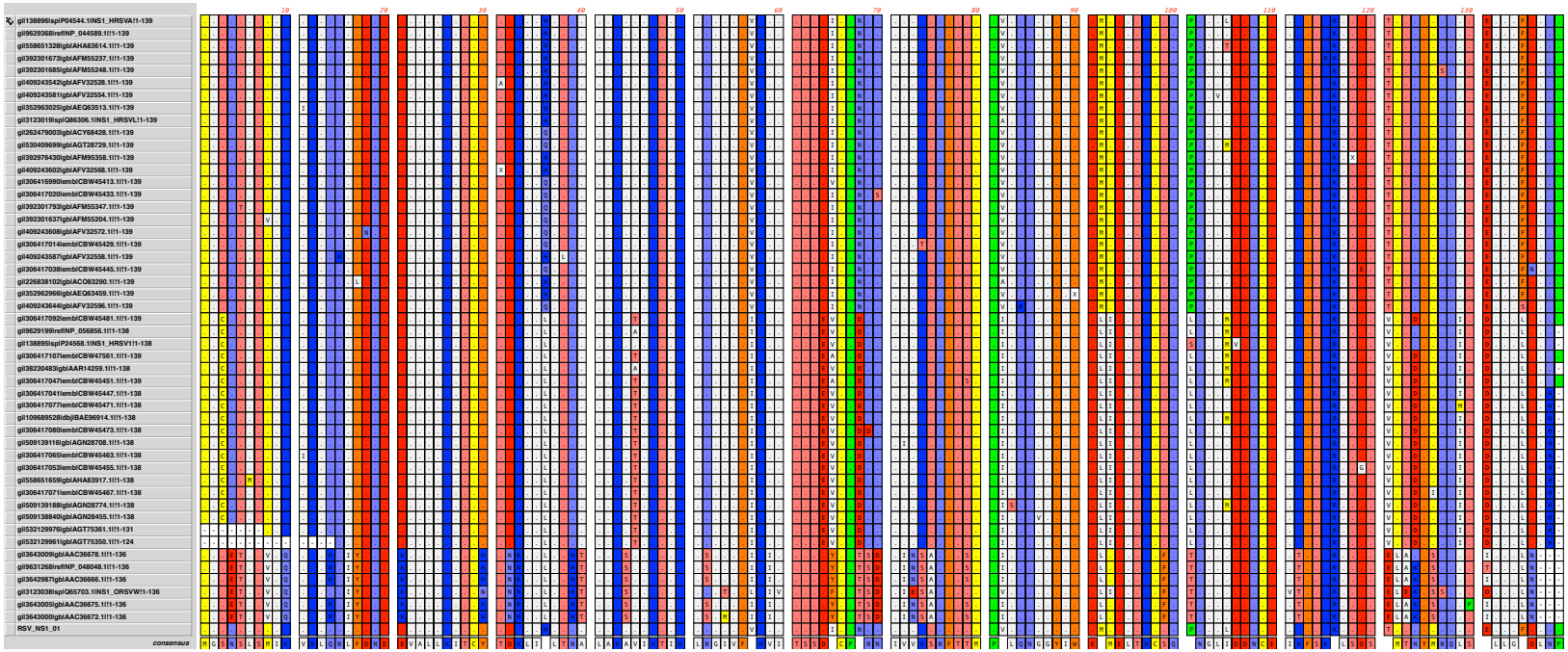


Figure 6 RSV NS1 Amino Acid Multiple Sequence Alignment
Multiple sequence alignment of 49 full-length RSV NS1 139 amino acid sequences using ClustalW. Far right column indicates the GenBank sequence identification. Colors indicate amino acid characteristic group; letter indicates a specific amino acid change from the reference amino acid sequence (bottom).

2.5 Discussion

The BLASTP search can return as many sequence matches as you request. The use of various algorithm parameters (such as “Max target sequences”, “Expect threshold”, etc.) can help focus the returned matches to suit the search requirements. The initial search returned 127 sequences matching some aspect of NS1, but visual inspection of the alignments revealed only 49 meaningful matches (high quality, full length). ClustalW was used to generate a multiple sequence alignment of these 49 sequences. The ClustalW alignment suggests there is sequence information that would allow discrimination between NS1 sequences of RSV strain A and strain B as well as from non-human NS1 virus sequences. Sequence alignment has been used in attributing function to unknown proteins [154,165], but again there are currently no known distant relatives to the RSV NS1 sequence at either the primary nucleotide, amino acid sequence or protein fold level [15] and there were no obvious sequence stretches that suggested a relationship with known protein functions (and therefore targets of modification).

CHAPTER 3. COMPUTATIONAL MODELING OF HUMAN RSV NS1 PROTEIN SEQUENCE

3.1 Abstract

Application of computational tools to improve vaccine design/development have unique requirements, in particular the protein structure. When the structure is not available, it becomes necessary to build a protein structure model computationally. Among the many protein structure model building programs, the Iterative Threading ASSEmbly Refinement (I-TASSER) [166] program is well regarded and uniquely so in creating de novo structural models when no similar protein structure is known.

3.2 Introduction

An alternative to multiple sequence alignment identification of interesting amino acid residues would be to computationally generate a protein fold model from the primary amino acid sequence and see 1) what does this model look like (is it believable) and 2) are there any region(s) that suggest a desired functional effect if modified. Generating an accurate protein structure in three-dimensional (3D) space is still an improving process [152,167,168]. If a primary amino acid sequence (or shorter functional stretches of amino acid sequences) is similar to any protein sequence/structure in the Protein Data Bank (PDB) [169–171], the likelihood of producing an appropriate 3D structural model is higher. However, if there are no related amino acid sequences with associated high-quality crystal (or other) structures, then this likelihood is unknown. Among the many *ab-initio* protein folding programs, I-TASSER program is efficient, highly regarded and easy to use. Computational methods used to predict three-

dimensional (3D) protein structures fall into one of three categories [166]. The first, comparative modeling, relies on the existence of highly homologous templates identified by sequence comparison in PDB. A second category relies on "threading" methods whereby protein sequences less well matched at the primary sequence level are nevertheless matched to 3D structures where certain folds are similar despite the lack of an evolutionary recognized primary sequence relationship. Lastly, protein sequences without any structurally related protein in PDB must have 3D structures modeled *ab-initio* [152,166].

Ab-initio model building was used in this effort as there was no available crystal structure for RSV NS1 at the outset of this research. Introduced in 2006, I-TASSER is a suite of programs [166,168,172] that in general, uses four steps: threading template identification, iterative structure assembly simulation, model selection and refinement, and structure-based function annotation. The first step, threading a query sequence through a non-redundant protein structure library, identifies any structural templates. This search for structural template matches relies on eight different fold-recognition programs. Any matches at the amino acid sequence level are noted and the corresponding protein fold structures are saved. This allows a larger variety of possible structural templates to guide model structure building. These results are divided into 1) threading-aligned and 2) threading-unaligned. Full-length model building is guided by computationally reassembling the continuously aligned fragment structures, in combination with the *ab-initio* folded structures derived from threading-unaligned regions.

The I-TASSER program outputs the top five models ranked in terms of 1) C-score estimate considering quality and convergence of threading alignments into structure clusters during assembly simulations [173], and 2) Cluster density [173] which is a measure of the number of structure decoy (replicas generated during reassembly simulations) that cluster with the model based on pairwise root-mean-square-deviation (RMSD) structure comparisons.

- C-score is a confidence score used by I-TASSER to estimate the quality of predicted models. It is calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. C-score is typically in the range of [-5,2], where a C-score of higher value signifies a model with a high confidence.
- Cluster density is defined as the number of structure decoys in a particular SPICKER [174] cluster. I-TASSER generates full length model of proteins by cutting continuous fragments from threading alignments and then reassembling them using replica-exchanged Monte Carlo simulations. Low temperature replicas (decoys) generated during the simulation are clustered by SPICKER and the top five cluster centroids are selected for generating full atomic models. A higher cluster density means the structure occurs more often in the simulation and therefore signifies a better-quality model.

3.3 Material and methods

I-TASSER on-line server (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) was used to generate 3D structure models of RSV NS1 amino acid sequence (GenBank:

ALS35583.1). The program was run using the default settings found on this website. Run time for a particular protein sequence can range from a few hours up to 60 hours. The NS1 sequence job took about 36 hours to complete. Results can be viewed directly in a web browser and/or downloaded as a compressed tar.bz2 file".

3.4 Results

Ribbon images for each of the five 3D structural models generated de novo for RSV NS1 are shown in Figure 7. Additional output files can be found in the folder generated by I-TASSER can be found in Supplemental Data 1.

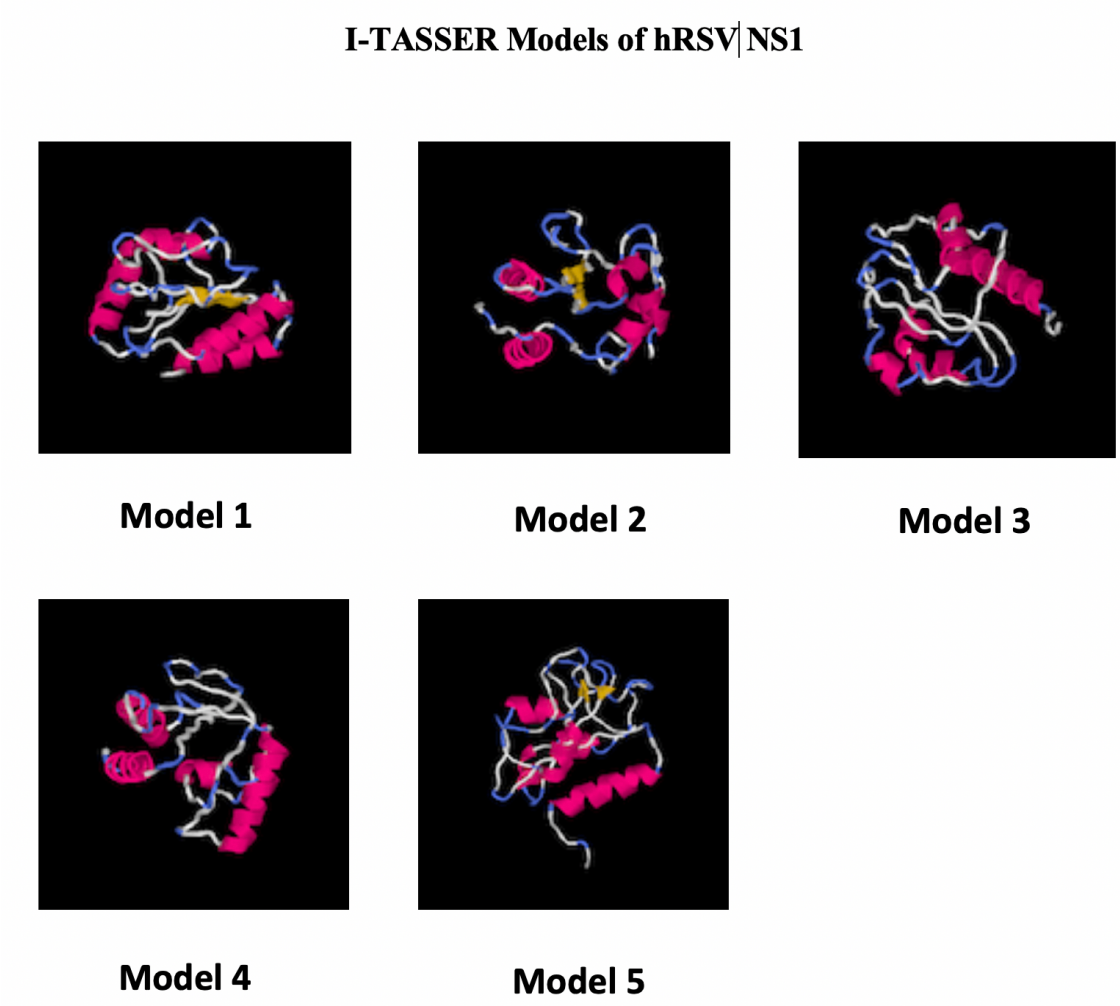


Figure 7 Graphical view of I-TASSER Models 1 to 5
Output ribbon diagrams of each NS1 protein structure model created from RSV NS1 sequence ALS35583.1. Secondary structures helix (Pink), sheet (Yellow) and coil (Blue) are shown highlighted.

Table 2 below contains the C-score and Cluster density measures for these same five models.

Table 2 I-TASSER Model C-score and Cluster density
C-score is a confidence score for estimating the quality of predicted models by I-TASSER typically ranging from -5.00 (lowest confidence) to 2.00 (highest confidence). Cluster density indicates the concentration of threading alignments that help guide structure model building. Better quality models are signified by higher cluster densities.

Name	C-score	Cluster density
Model 1	-3.18	0.0676
Model 2	-4.03	0.0290
Model 3	-4.57	0.0168
Model 4	-5.00	0.0088
Model 5	-5.00	0.0074

I-TASSER uses these metrics (C-score and Cluster density) to rank the output models. As shown in the table, Model 1 has the highest C-score (-3.18) and Cluster density (0.0676). While Model 1 was higher in rank than the next four models, none of the models possess metric values of such distinction that clearly separate one from the rest. However only a single model will be required as input to select modification region(s). For this application Model 1 was selected due to the highest C-score of -3.18 and Cluster density of 0.0676.

3.5 Discussion

Generating 3D protein structure models from primary amino acid sequences continues as an ever-improving effort. Among these programs, I-TASSER has been at the forefront. Many programs look to build 3D models based on the existence of previously solved structures and use this to guide the modeling of related sequences. However, the opportunity to create a meaningful model from a primary amino acid sequence when there is no related solved structure of any kind remains a challenge. The I-TASSER program is one of the better tools to create such a model structure. Based on C-score of -3.18 and Cluster density of 0.0676, Model 1 was selected as the preferred candidate structure for future computational analysis. The C-score and Cluster density values for all five models were well below values of high-quality models. Without any similar structure available in PDB, the model creation was not derived from any existing known structures. In spring of 2017, a crystal structure of RSV NS1 was solved and deposited in PDB, 5VJ2 [15] which was long after the I-TASSER generated models and mutant selections were made, but allowed a separate comparison between the I-TASSER models and the published NS1 crystal structure. Evaluation of the five models suggested that they may possibly represent various versions or energy states of the NS1 protein. Being a nonstructural protein, it is possible NS1 exists in more than one conformation as it knowingly partners with multiple host and or viral proteins to influence various processes in the infected cell.

CHAPTER 4. COMPUTATIONAL MODEL EVALUATION

4.1 Abstract

A critical question regarding computationally derived protein structure model is the degree of nativeness of the model. The I-TASSER models were reexamined using the SWISS-MODEL Structure Assessment Tool. Model quality is crucial to determine the utility of a computationally generated 3D protein structure model [175–180]. Quality is generally qualified in terms of local composite and global scoring functions [177]. These scoring functions rely on several statistical descriptors expressed as potentials of mean force measured in the protein structure model and how these relate to distributions of similar measures of experimental structures. Using a linear combination of these scores, a separate estimate of absolute model quality is provided [152,178].

4.2 Introduction

Model quality provides great insight into the relevance of a computationally derived 3D protein structure. The accurate physical orientation of amino acids in the model are critical to determining the suitability of the model for various biomedical applications. The SWISS-MODEL server pioneered the field of automated comparative modeling of 3D protein structures with its introduction in 1993 [181]. Beyond 2003, continuous server development has led to the introduction of a fully automated protein homology-modeling server [177,182]. When models are generated using different technologies, model quality evaluations provide a way to compare such models. When a 3D model of protein structure is derived from related existing known structures

(homology modeling), the model quality is related to the evolutionary distance between protein of interest sequence and available related sequence templates with known structures. Alternatively, models built *ab-initio* from amino acid sequences lacking existing related structures will not benefit from the relative ranking of models as compared to a known structure. Here a non-relative quality assessment is required. The estimates of quality incorporated into the SWISS-MODEL Structure Assessment website provide a better absolute scale by which any protein structure model can be evaluated [152].

4.3 Materials and methods

SWISS-MODEL Structure Assessment on-line server (https://swissmodel.expasy.org/assess) provided an independent measure of quality among the five models generated by I-TASSER. The program was run using the default settings found on this website. The SWISS-MODEL server provides an absolute quality estimate for a protein model structure using a linear combination of six structural descriptors that attempt to capture both local and global quality measures of a 3D protein structure. The first four measures of Qualitative Model Energy Analysis (QMEAN) [178] are statistical measures of mean force and look to capture local attributes of a model:

- C ß interaction energy
- All-atom pairwise energy
- Solvation energy
- Torsion angle energy

The remaining two measurement terms capture more global aspects of a 3D model [152] and include:

- Secondary structure agreement
- Solvent accessibility agreement

These values are further adjusted using a Z-score normalization that helps remove any amino acid sequence length bias present in the raw QMEAN score so that the QMEAN Z-score values are a better "degree of nativeness" estimate of the protein model [177].

4.4 Results

The first four scoring functions (C β interaction energy, All-atom pairwise energy, Solvation energy and Torsion angle energy) are considered local while the next two (Secondary structure agreement and Solvent accessibility agreement) represent the global measures. The local scoring functions indicate to what extent the adjacent amino acids in the model structure satisfy know biochemical properties of amino acid bond lengths, etc. The remaining scoring functions indicate to what extent longer range features, such as torsion angles between three consecutive amino acids, secondary (or higher order) protein structure violate known spatial relationships. The final scoring function, QMEAN6, is a composite of the previous six scores. The Z-score version of this scoring function is produced by normalizing the Raw Score values relative to the length of the protein sequence, thereby removing the inflated influence of longer protein sequences. Model 1 QMEAN Z-score is -2.78 on a scale that considers any score lower

than -4.0 a bad model [177]. The table below provide no strong indication that any model is significantly better than the others.

Table 3 SWISS-MODEL Structure Assessment Scores
Each row represents a specific energy potential measurement. C β interaction energy, All-atom pairwise energy, Solvation energy and Torsion angle energy are local atomic interactions measurements. Secondary structure agreement and Solvent accessibility agreement are global interaction measurements. QMEAN6 is a linear combination of these measurements. Raw Score is the potential energy measurement. Z-score is an amino acid length normalized version of the Raw Score [178].

	Model 1		Model 2		Model 3		Model 4		Model 5	
Scoring function term	Raw Score	Z-score	Raw Score	Z-score	Raw Score	Z-score	Raw Score	Z-score	Raw Score	Z-score
C β interaction energy	-69.4	-0.69	-94.3	0.08	-86.29	-0.11	-76.66	-0.46	-53.3	-1.12
All-atom pairwise energy	-2748.51	-1.43	-3372.91	-1.06	-2261.25	-1.83	-1653.29	-2.32	-1533.44	-2.27
Solvation energy	-7.08	-0.94	-18.75	0.93	-6.41	-1.05	-14.06	0.18	-2.48	-1.68
Torsion angle energy	1.53	-3.31	6.27	-3.71	10.35	-4.05	2.25	-3.38	7.53	-3.82
Secondary structure agreement	69.80%	-2.14	69.10%	-2.25	57.60%	-3.97	64.70%	-2.89	61.90%	-3.32
Solvent accessibility agreement	73.40%	-0.61	76.30%	-0.05	66.20%	-2.01	72.70%	-0.75	66.90%	-1.87
QMEAN6 score	0.493	-2.78	0.493	-2.78	0.303	-4.92	0.427	-3.52	0.343	-4.47

4.5 Discussion

While it is possible to mix and match various higher quality aspects of select models, 1) that was beyond the intent of this project and 2) there are no clearly better aspects of among the models that warrant consideration. In addition, the purpose was to rely on as many off the shelf programs as possible to enable as many researchers as possible to utilize the methodology. For example, a refinement of each model would require a series of subjective steps not suitable for someone inexperienced in the use of software required. Based on the SWISS-MODEL quality assessment, Model 1 was still considered an optimal model since no other models appeared more distinguished using the QMEAN measurements. The Z-scores for any measures, but especially for the QMEAN6 scoring function, further reinforce that Model 1 may be the least bad model of

the five. The output scores for each of the models are unremarkable and representative of the fact that *ab-initio* models, while improved, are difficult model structures in which to assess quality when no similar solved structures are available.

CHAPTER 5. COMPUTATIONAL MODEL COMPARISON TO HUMAN RSV NS1 PROTEIN CRYSTAL STRUCTURE (5VJ2)

5.1 Abstract

In spring of 2017, a crystal structure of RSV NS1 was solved and deposited in PDB, 5VJ2 (https://www.rcsb.org/structure/5vj2) [15]. The crystal structure PDB 5VJ2 of RSV NS1 provided a unique opportunity to compare the I-TASSER generated Model 1 of RSV NS1 with a solved crystal structure.

5.2 Introduction

Prior to the RSV NS1 crystal structure 5VJ2, Model 1 (from the five models generated by I-TASSER) was chosen and moved forward as the structure in which amino acid replacements could be measured on. While desirable to have had a solved protein crystal structure at the outset of this project, the availability after the fact allowed an unanticipated comparison between the two protein structures. Generally, assessing model quality is straight forward if there exists know solved structure(s) to compare to. The issue of model quality in the situation where no know structure (distantly related or otherwise) is available can be quite unpredictable. However, the situation where a protein sequence model is generated (and used in analysis) but later a crystal structure for the same protein is produced, presented a unique opportunity. A straightforward superpositioning of a model on the known structure allowed visual (and even numerical) relatedness, but this does not address the possibility that the model(s) and known structure are versions of one another. The MatchMaker tool in UCSF Chimera [183] is a useful program for overlaying proteins structures while the Morph Conformations [183]

tool helps to visualize what changes a given structure needs to match another structure. These tools each provide some visual insights regarding "quality" at least in terms of whether there are any structure similarities, unlike the quality metric values provided by SWISS-MODEL assessment of the I-TASSER ab-initio built Model 1. Lastly in order to provide a meaningful evaluation of "real" versus "model" it was necessary to apply the same scrutiny as was performed on Model 1 generated by I-TASSER to the published structure of RSV NS1 (PDB 5VJ2). This analysis allowed a different comparison of the structures in terms of amino acid replacement effects that might not be reflected in any of the previous assessments.

5.3 Material and methods

The following must be noted regarding each structure: 1) the published 5VJ2 crystal structure lacks alpha carbon (Cα) coordinates for amino acids 1-5 in the structure and 2) Model 1 structure does not contain Cα coordinates for amino acid 68. The structure quality assessment tools do not experience any issues based on these differences. However, an analysis based on amino acid change effects must restrict the comparison to amino acids 5-67 and 69-139 to directly compare the effects of a given amino acid change in each structure. The MatchMaker [183] tool in the UCSF Chimera software package (version 1.14) was used to superimpose Model 1 onto 5VJ2. The PDB structure file for Model 1 ("Structure to match") was aligned to 5VJ2 ("Reference structure") using the default settings in MatchMaker (Tools, Structure Comparison, MatchMaker) to see whether there were any region(s) of structural similarity (Figure 8).

The default settings used in MatchMaker include:

- Chain pairing – Best-aligning pair of chains between reference and match structure
- Alignment algorithm: Needleman-Wunsch (global)
- Matrix: BLOSUM-62

This allowed a superpositioning visual (structure overlap) to be generated revealing where overlaps exist between the two structures (Figure 8).

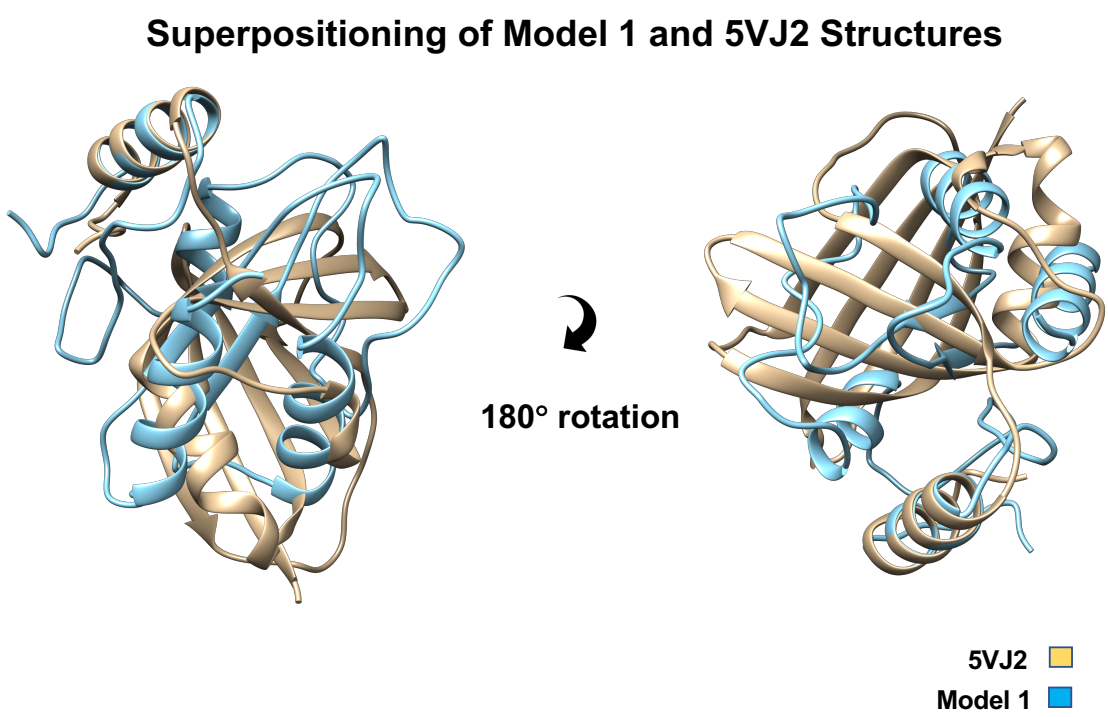


Figure 8 Superposition Alignment between Model 1 and 5VJ2
Superpositioning was performed between Model 1 (Blue) and 5VJ2 (Tan) using the MatchMaker tool in the Chimera software package [183].

Next the Morph Confirmation tool in Chimera was used to show the trajectory (positional changes) that Model 1 structure could follow to conform with 5VJ2 structure. A short movie was recorded to show this transformation (Figure 9).

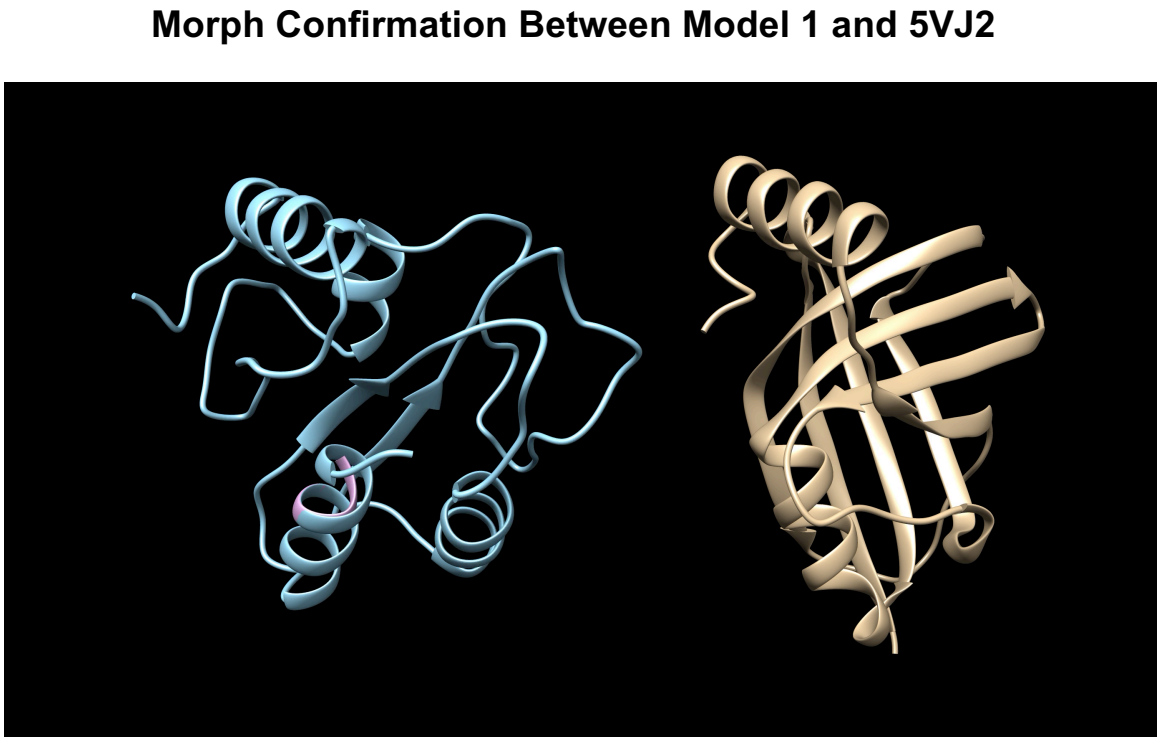


Figure 9 Morph Confirmation Alignment between Model 1 and 5VJ2
The Morph Conformation tool in Chimera software package was used to generate a video of an energetically favorable structural transformation between Model 1 (Blue) and 5VJ2 (Tan) [183]. MorphConfirmation.mp4 video found in Supplemental Data.

Morph Confirmation video provides a multi-dimensional view of how and where the pair of protein structures could morph from one to another. A final comparison relies more on amino acid substitution effects in each structure rather than the structure itself. These evaluations required Model 1 and 5VJ2 be processed using the

`stability_changes_ddG.pl` (with `-m2` switch) program in AUTO-MUTE 2.0 [184,185] in order to generate predicted numerical ddG quantities for comparison between Model 1 and 5VJ2. The numerical ddG values provide a basis through which a comparison between the structures can be made. Specifically, the ddG values for both Model 1 and 5VJ2 are graphed as x, y data points and linearly regressed. The X-Y plots of both an all-inclusive ddG values data set as well as filtered set (Figure 10 and Figure 11) indicate the aspects of correlation among the ddG values associated with the amino acid replacements tested in each structure.

5.4 Results

The superpositioning of Model 1 onto 5VJ2 can be seen in Figure 8. Most of the amino acid positioning do not appear to overlap greatly except the coil structure found from amino acid 119 to 139 in both structures. Whether NS1 may exist in a spectrum of structures is also a possibility and as such the lack of complete structure overlap is not surprising. Morph Confirmation tool shows visually to what extent Model 1 structure needs to transform to match structural aspects of 5VJ2. The video in Figure 9 shows this transformation. Again, the coil structure in Model 1 (amino acid 119-133) is linearly transformed onto 5VJ2. In addition, it appears that some aspects of the coil structure in Model 1 between amino acid 36-48 are transformable onto a coil structure in 5VJ2 between amino acid 33-43. Finally, each protein structure model was analyzed with AUTO-MUTE 2.0 generating ddG values for each amino acid change along the protein sequence. Since each structure is derived from the same amino acid sequence, the predicted values for the same changes at each shared position (133 out of 139 total amino

acids) can be visualized in an X-Y type plot and linear regression can quantify the extent of any relationship. When all shared amino acid replacements (2546 ddG values) are used, the linear correlation $R^2 = 0.4006$ (Figure 10). The inclusion of values from all shared amino acid position changes had evidently skewed the distribution of plotted values likely leading to low $R^2 = 0.4006$ value despite visually, there appearing to be a level of linear relatedness.

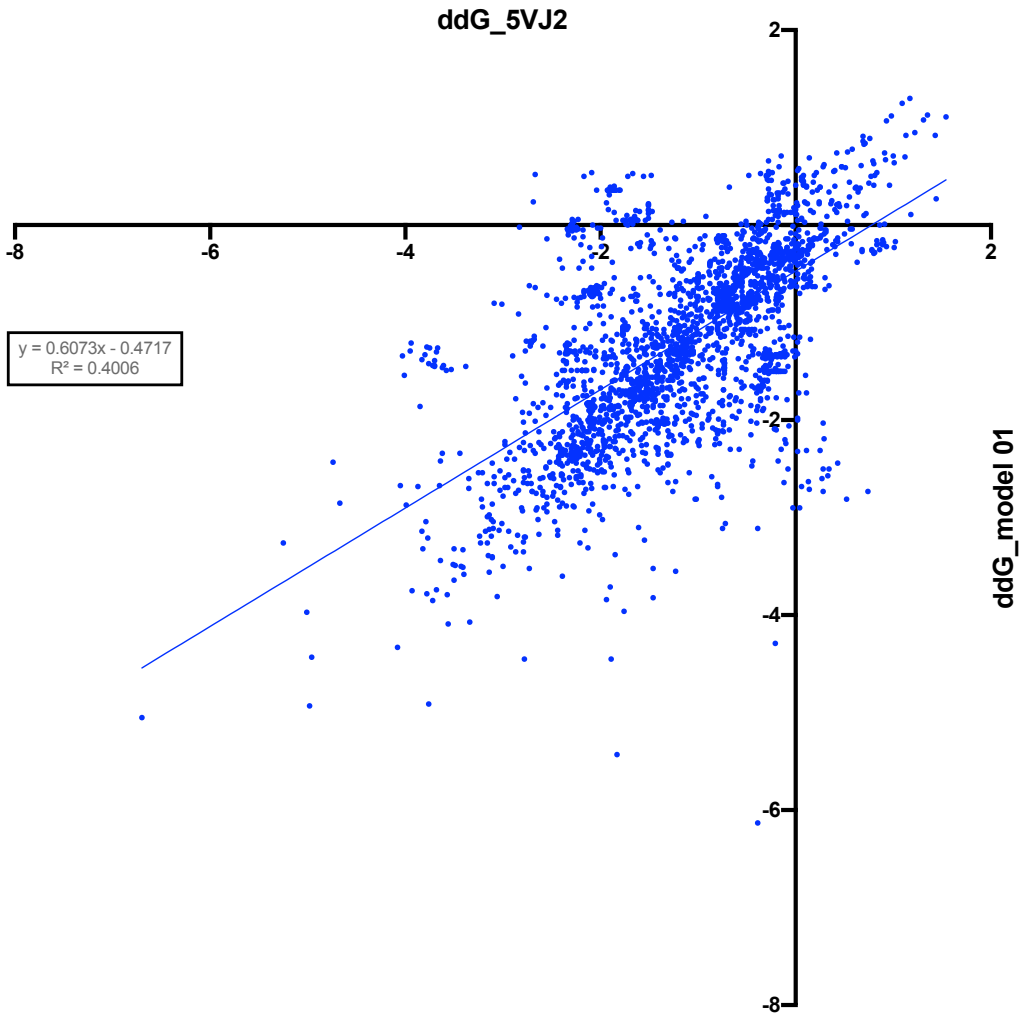


Figure 10 Linear Correlation of ddG Values Between 5VJ2 and Model 1
AUTO-MUTE 2.0 was used to calculate ddG values for all possible amino acid substitutions at every amino acid sequence position (except for positions 1-4 and 68, see text) in Model 1 and 5VJ2. The graph shows a dot plot of ddG values for all shared amino acid substitutions (n = 2546); the linear correlation R2 is 0.4006.

The idea that certain amino acids are confounding the recognition of a possible linear relationship was addressed by removing those shared amino acid changes that produced ddG values greater than 1.5. In fact, up to 92% of the shared amino acid

changes (2340 out of 2546) give ddG values that have a linear correlation with an R^2 value = 0.6568 (Figure 11).

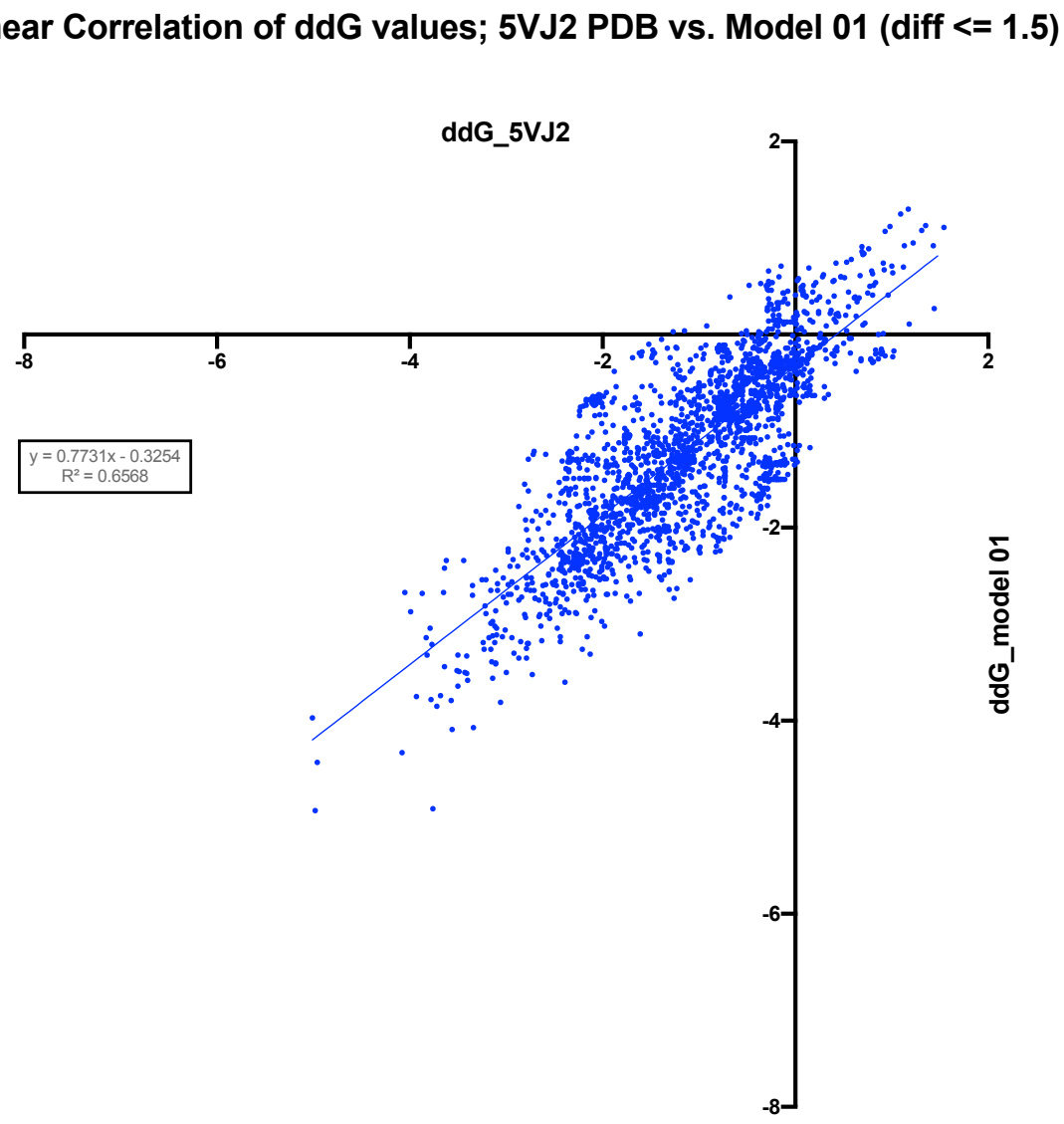


Figure 11 Linear Correlation of ddG Values Between 5VJ2 and Model 1 (diff <= 1.5)
AUTO-MUTE 2.0 was used to calculate ddG values for all possible amino acid substitutions at every amino acid sequence position (except for positions 1-4 and 68, see text) in Model 1 and 5VJ2. The graph shows a dot plot of ddG values following removal of amino acid replacements with ddG values greater than 1.5. The retained 92% of 311 the substitutions (2340 out of 2546 ddG values) have a linear correlation R2 value of 0.6568.

Additionally, we wanted to evaluate whether the ddG value correlation seen was related to local secondary structures despite a global structure similarity. Among predicted ddG values of all substitutions in each structure (5VJ2 and Model 1) for amino acids located in 5VJ2 Helix structures, we find that 83 % of these range within +/- 0.8 ddG (kcal/mol) units from a theoretical midline, Y = X; line of equality (Figure 12).

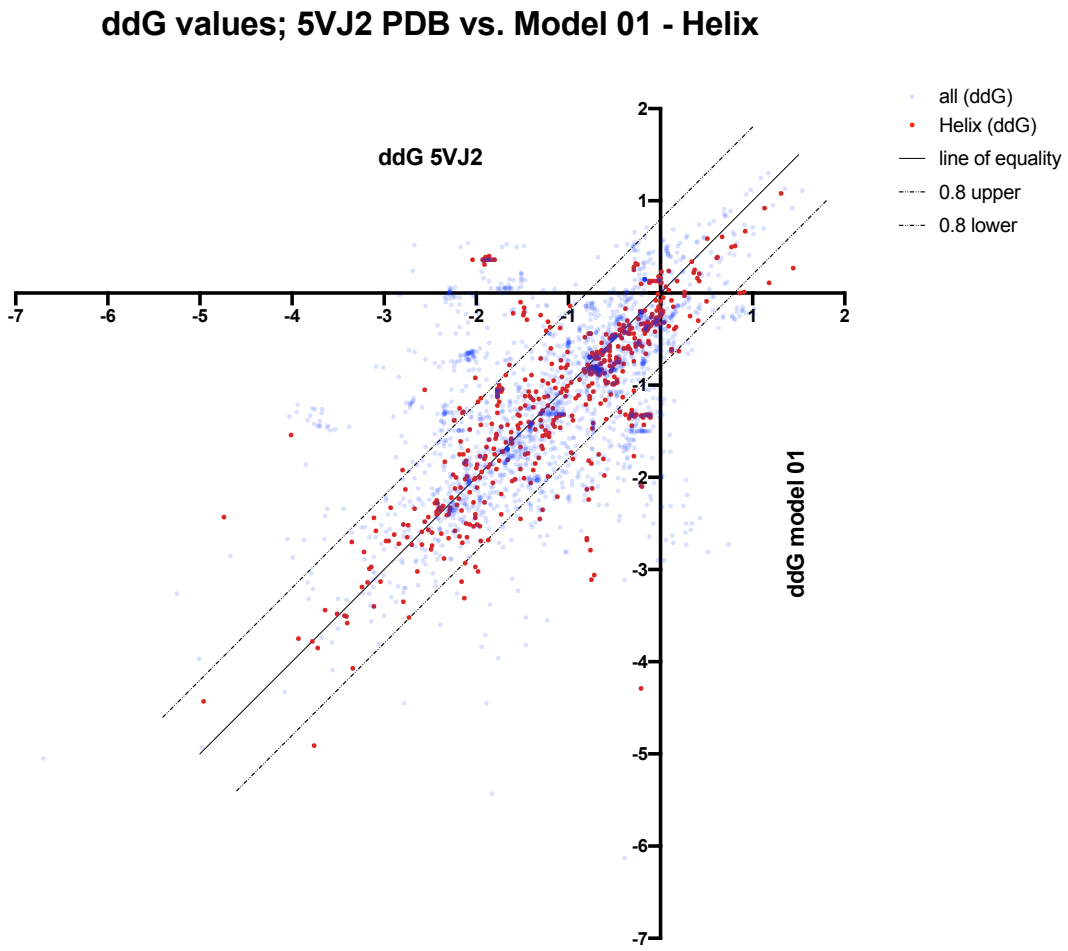


Figure 12 ddG values of Helix secondary structure amino acids
AUTO-MUTE 2.0 was used to calculate ddG values for all possible amino acid substitutions at every amino acid sequence position (except for positions 1-4 and 68, see text) in Model 1 and 5VJ2. The graph shows a dot plot of ddG values for all shared amino acid substitutions (n = 2546; Light Blue dots) and all amino acids identified within a Helix structure of 5VJ2 are shown in Red. As points of reference, a diagonal, Y = X solid line as well as 0.8 ddG (kcal/mol) upper and lower bound dashed lines are also shown.

Similarly, 77 % and 72 % of the predicted ddG values of all substitutions in each structure (5VJ2 and Model 1) for amino acids located in 5VJ2 Sheet and Coil structures respectively are shown in Figure 13 and 14.

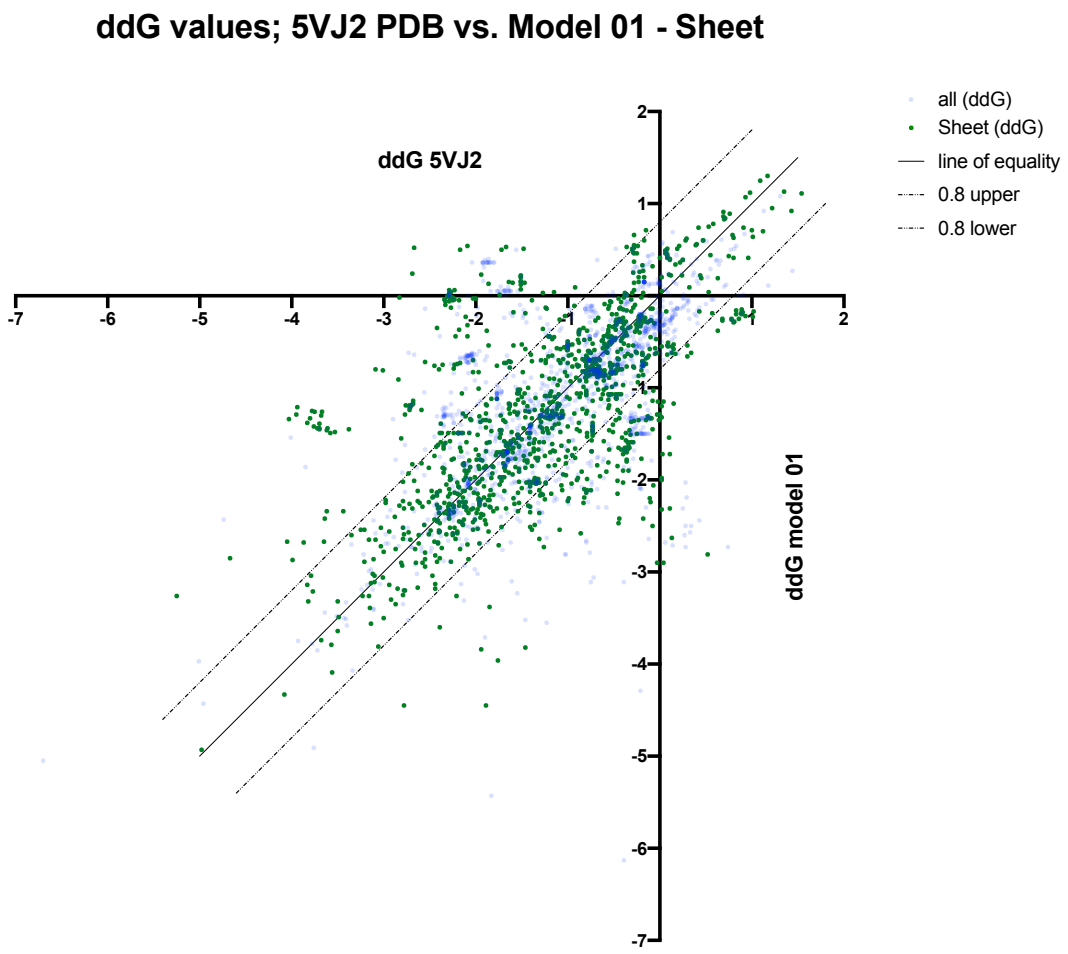


Figure 13 ddG values of Sheet secondary structure amino acids
AUTO-MUTE 2.0 was used to calculate ddG values for all possible amino acid substitutions at every amino acid sequence position (except for positions 1-4 and 68, see text) in Model 1 and 5VJ2. The graph shows a dot plot of ddG values for all shared amino acid substitutions (n = 2546; Light Blue dots) and all amino acids identified within a Sheet structure of 5VJ2 are shown in Green. As points of reference, a diagonal, Y = X solid line as well as 0.8 ddG (kcal/mol) upper and lower bound dashed lines are also shown.

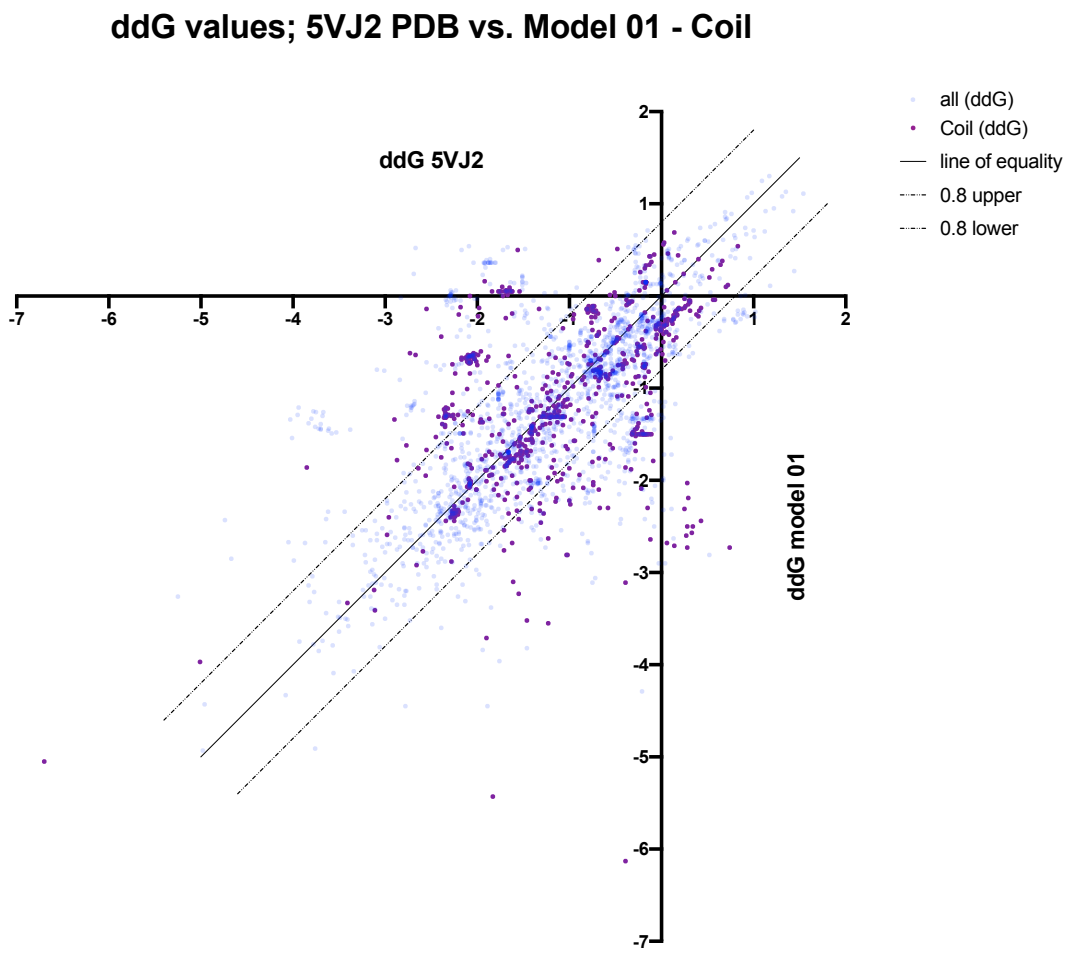


Figure 14 ddG values of Coil secondary structure amino acids
AUTO-MUTE 2.0 was used to calculate ddG values for all possible amino acid substitutions at every amino acid sequence position (except for positions 1-4 and 68, see text) in Model 1 and 5VJ2. The graph shows a dot plot of ddG values for all shared amino acid substitutions (n = 2546; Light Blue dots) and all amino acids identified within a Coil structure of 5VJ2 are shown in Purple. As points of reference, a diagonal, Y = X solid line as well as 0.8 ddG (kcal/mol) upper and lower bound dashed lines are also shown.

5.5 Discussion

The opportunity to contrast a known protein structure against one or more models of the same protein provided a unique opportunity. However, when a model structure is used as a surrogate in assessing amino acid replacements effects on the protein function,

an even more unique opportunity arises. How the validity of a model is measured may better be referred to as how the validity of model function is measured. The results here shed some light on this question. When a protein structure model is evaluated, it can be compared with structural peers and scored accordingly. If there is a lack of appropriate peers, then the model structure can be assessed based on well-established physical properties found in known protein structures. Both these paths rely on properties found among previously described protein structures. These are perfectly sensible approaches if there exists prior relevant knowledge. Alternatively, it is conceivable that a published protein structure is part of an ensemble of a flexible protein structures and that the model(s) may capture sufficient structure/function aspects along such a spectrum to be of value despite low score quality assessments of these modeled versions of a known structure. RSV NS1 has been associated with a spectrum of activities in an infected cell therefore, it is not unlikely the protein exists in multiple structural configurations and the lackluster quality scores from SWISS-MODEL and even I-TASSER itself might hint at this suggestion. The availability of a published protein structure for NS1 provided an opportunity to examine "real" versus "model" comparisons of these structures.

MatchMaker provided a quick visual assessment of whether there was any similarity between the structure of 5VJ2 and Model 1. The superpositioning indicates there was a near perfect agreement in the terminal 119-139 amino acid structure. However, most of the remaining sequence did not seem to share similarity under the direct overlay criteria. This type of assessment is not unlike the previously discussed quality assessments relying on previously established structure frameworks. If NS1 is a

flexible protein, then a tool like Morph Confirmation may be more appropriate for comparison. The video of Model 1 transforming into 5VJ2 seems to suggest a greater level of relatedness than seen by any other measure so far. This transformation again demonstrates the coil structure in Model 1 (amino acid 119-133) linearly transformed onto 5VJ2 but also reveals some aspects of the coil structure in Model 1 between amino acid 36-48 as transformable onto a coil structure in 5VJ2 between amino acid 33-43. This type of visual suggests that the NS1 structure may fluctuate to such an extent that both Model 1 and 5VJ2 are versions of one another. However, none of these comparisons address the issue of functional relatedness.

The comparison of shared amino acid replacement ddG values does. Here the predicted ddG values of the same amino acid replacements in Model 1 and 5VJ2 are shown to produce similar effects. When all amino acid replacements (2546) are used the linear correlation R^2 is 0.4006 (Figure 10). This is evidently skewed by certain amino acids that are likely poorly modeled and therefore providing a disproportionate effect. The idea being that if Model 1 and 5VJ2 share enough structural/functional aspects then there should be a correlation among the ddG values associated with the many amino acid replacements tested in each structure. The removal of a minimal confounding set of amino acid replacement measurements improved the visualization of similarity between these two structures as related to amino acid replacement effects. A plot of nearly 92% of the shared amino acid changes (2340 out of 2546) give ddG values that have a higher linear correlation R^2 value = 0.6568 (Figure 11).

Another consideration toward reconciling the apparent ddG correlation between the structures despite little global structure similarity likely resides in the local structures. Figures 12, 13 and 14 indicate that predicted ddG values for substitutions of Helix, Sheet or Coil 5VJ2 identified amino acids in both 5VJ2 and Model 1 are well correlated as indicated by the proximity to theoretical 1:1 diagonal line and further support the idea that useful models of soluble proteins like NS1, possibly existing in multiple confirmations, depend more on local secondary structure determination rather than the global orientation of these local secondary structures.

Whether there are likely more than a single protein structure associated with a spectrum of protein functions should be considered. Here the suggestion is that Model 1, while not ideal in terms of certain assessments, seems to perform as a sufficient surrogate for measuring the functional effect of specific amino acid substitutions. Model 1 is not inappropriate as a surrogate in terms of evaluating amino acid replacement effects despite revealing little to no similarity in terms of structure alignments. As a result, it is conceivable that the crystal structure published is simply a "version" of a flexible protein structure and that the model(s) may capture sufficient structure/function aspects along a structure spectrum to be of value.

CHAPTER 6. COMPUTATIONAL SELECTION OF MUTANTS

6.1 Abstract

Model 1 provided a 3D protein structure on which the AUTO-MUTE 2.0 [184,185] program can operate. The current generally accepted search for improved vaccine candidates relies on relatively subjective criteria [3,4]. Applying computational evaluations of 3D protein structures/models using a machine learning predictor, AUTO-MUTE 2.0, provided predictions of protein function. Based on these predictions, twelve sets of one or more residue replacements, falling into two groups, Group 1 and Group 2, were identified.

6.2 Introduction

Vaccine candidate design has relied on a historically subjective selection criterion. Earliest human viral vaccines were weakened versions of the disease causing virus [20–22] followed by the concept of virus inactivation which maintained necessary immunogenicity but eliminated virus growth [23]. Subsequent vaccine design strategy continued to progress (deletions, rearrangements, chimera, mutagenesis) [3,4,145,149], but to a large extent did not utilize fully the benefits of computational and structural biology. More recently there is an increasing level of computational objectivity being found in vaccine design methods (codon optimization, codon-pair optimization, protein structure selection) [89,186–189]. This movement towards more rational target selection/modification [190] helps reduce subjectivity.

Model 1 provided a 3D protein structure on which the AUTO-MUTE 2.0 program can operate. AUTO-MUTE 2.0 is a collection of programs that provide predictions regarding the effects specific amino acid change(s) will have on a proteins function. AUTO-MUTE 2.0 works by identifying the six closest amino acid residue positions that are structural neighbors to the amino acid residue undergoing substitution. Relevant attributes regarding the substituted residue as well as the 6 closest neighbors (< 12 angstroms away) make up individual component values of a unique feature vector characterizing each change. The program then employs a four-body knowledge based statistical potential function to provide structural perturbation values resulting in a subsequent residue replacement feature vector for each residue change. These residue replacement feature vectors are supplied to predictive models trained using statistical machine learning algorithms. Large sets of diverse mutations that have been studied experimentally for their functional effects and which occur in proteins that share low sequence similarity, were used to train these AUTO-MUTE 2.0 predictors. The following is a list of AUTO-MUTE 2.0 programs available:

- `stability_changes_ddG.pl`
- `stability_changes_ddG_H2O.pl`
- `stability_changes_dTm.pl`
- `activity_changes.pl`
- `human_nsSNPs.pl`

The stability change predictors offer two supervised classification models

- `-m 0 means RF (ddG and ddG_H2O) or AdaBoostM1/C4.5 (dTm) classification`
- `-m 1 means Support Vector Machine (SVM) classification (all 3 stability change programs)`

The stability change predictors offer two regression models (-m2 or -m3)

- `-m 2 means REPTree tree regression (all 3 stability change programs)`
- `-m 3 means Support Vector Regression (SVR, all 3 stability change programs)`

In short, the AUTO-MUTE 2.0 program provided predictions of protein function upon specific residue replacement for each amino acid in the supplied Model 1 of RSV NS1 protein structure.

6.3 Material and methods

Among the different free energy change calculations and machine learning classifiers available, we used the `stability_changes_dTm.pl` because of its thermal stability prediction labels "**Increased**" or "**Decreased**" and overall superiority described in a previously published classifier comparison [184,191,192] in order to get a

general categorization of all possible mutations. Additionally, we used `stability_changes_ddG.pl` to generate ddG prediction values for each possible substitution. Average ddG prediction values among specific amino acid groups (increased interior or protein-protein interface frequency) [193] at each amino acid position in Model 1 was used to highlight/identify specific substitution locations in predicted Model 1 secondary structures. The Model 1 3D structure from I-TASSER was used as input to AUTO-MUTE 2.0 and an output file containing either the amino acid stability change effect prediction or ddG value prediction was returned as output (files can be found in Supplemental Data). Elapsed real time for AUTO-MUTE 2.0 to make each type of prediction for all amino acid substitutions in Model 1 was about 72 hours.

6.4 Results

The `stability_changes_dTm.pl` program in AUTO-MUTE 2.0 was used to predict Increased or Decreased thermal stability [185] for all possible amino acid replacements at each of the 139 residues in Model 1 except residue 68, which was unable to satisfy the 6 "nearest neighbor" requirement. The majority of these changes were predicted as Decreased (78%, 2036/2621 **Decreased**; 22%, 585/2621 **Increased**; data not shown) which is a known computational bias found among protein stability prediction programs [194,195]. In addition, ddG values were predicted (calculated using `stability_changes_ddG.pl`) for all possible residue replacements in Model 1 with the exception of residue 68 as indicated above and the output data can be found in Supplemental Data. As a result, specific locations in the NS1 protein sequence were chosen for evaluation based on the overall predicted sensitivity (**Decreased**), average

ddG predicted values of groups of amino acids more frequently found within protein structure interior or protein-protein interfaces and location within predicted secondary structures in Model 1. Twelve mutants were identified and engineered into the original NS1 protein sequence. Eight of these mutants were collectively called Group 1 in which 1 to 3 amino acid assignments in NS1 were changed to alanine (A); specifically, residues 6 and 9 (NS1 06-09); residues 11, 13, and 15 (NS1 11-13-15); residue 29 (NS1 29); residue 47 (NS1 47); residues 54, 57 and 60 (NS1 54-57-60); residues 58 and 66 (NS1 58-66); and residues 98 and 109 (NS1 98-109). An additional mutant in Group 1, called NS1 98-104st, was inadvertently created due to a PCR error: in this sequence, amino acid 98 was correctly changed to A, but additionally, amino acids 101, 102, 103 and 104 also were changed (P>L, N>M, G>V, and L>Y) followed by an introduced stop codon after amino acid 104.

In a separate set of four mutants, collectively called Group 2, the leucine (L) residue at amino acid position 6 (6L) was replaced alternately with four amino acids predicted to decrease protein thermal stability and possess substantially different biochemical properties as compared to Leucine: namely, 6L was replaced with aspartic acid - 6D (NS1 06 D), glutamine - 6Q (NS1 06 Q), arginine - 6R (NS1 06 R) or tryptophan - 6W (NS1 06 W) which changed the non-polar (L) to either a polar-charged (D), neutral polar (Q), charged aliphatic (R) or non-polar aliphatic (W) amino acids.

Table 4 lists the mutants, the original and mutant amino acid assignments, their locations, the local secondary structure, the AUTO-MUTE 2.0 thermal stability

prediction and ddG value prediction for each amino acid substitution in Model 1 (and 5VJ2 as a matter of reference). Mutant NS1 98-104st is not shown in Table 4.

Table 4 Predicted properties of amino acid substitutions in the NS1 protein
The table includes the mutant virus names, original amino acid assignments, new mutant assignments, positions in the amino acid sequence, predicted local secondary structure, predicted effect on stability and predicted change in free energy (ddG) for both Model 1 and 5VJ2 structures. Groups 1 and 2 are identified by vertical bars on the left.

	Name	Amino Acid Original	Amino Acid New	Amino Acid Number	Secondary Structure in Model 1	Secondary Structure in 5VJ2	Auto-Mute 2.0 Predicted Stability Model 1	Auto-Mute 2.0 Predicted Stability 5VJ2	ddG Model 01 (kcal/mol)	ddG 5VJ2 (kcal/mol)
Group 1	NS1 6-9	L	A	6	Helix	Sheet	Decreased	Decreased	-2.51	-1.41
		I	A	9	Helix	Sheet	Decreased	Decreased	-2.89	-3.21
	NS1 11-13-15	V	A	11	Helix	Sheet	Decreased	Decreased	-0.91	-2.84
		L	A	13	Helix	Sheet	Decreased	Decreased	-2.42	-3.64
		N	A	15	Helix	Sheet	Increased	Decreased	-0.73	-1.24
	NS1 29	C	A	29	Sheet	Sheet	Increased	Decreased	0.20	-0.23
	NS1 47	H	A	47	Sheet	Sheet	Decreased	Decreased	0.53	-0.36
	NS1 54-57-60	I	A	54	Sheet	Sheet	Decreased	Decreased	-2.52	-2.79
		V	A	57	Sheet	Sheet	Decreased	Decreased	-2.31	-2.69
		I	A	60	Sheet	Sheet	Decreased	Decreased	-3.56	0.00
	NS1 58-66	H	A	58	Sheet	Sheet	Decreased	Decreased	0.41	-3.14
		C	A	66	Coil	Coil	Increased	Increased	-0.05	0.30
	NS1 98-109	C	A	98	Helix	Sheet	Increased	Increased	-2.63	0.13
		C	A	109	Coil	Coil	Decreased	Decreased	-2.50	0.34
Group 2	NS1 6 D	L	D	6	Helix	Sheet	Decreased	Decreased	-3.38	-1.85
	NS1 6 R	L	R	6	Helix	Sheet	Decreased	Decreased	-2.24	-1.96
	NS1 6 Q	L	Q	6	Helix	Sheet	Decreased	Decreased	-1.59	-1.26
	NS1 6 W	L	W	6	Helix	Sheet	Decreased	Decreased	-0.61	-0.58

Figures 4A-F show the highlighted location of each amino acid substitution in a 3D ribbon diagram of Model 1. Specifically, NS1 06-09 and NS1 11-13-15 (Figure 4A) are located in an amino terminal helix structure in Model 1 but not in 5VJ2 structure. NS1 29 and NS1 47 (Figure 4B and 4C) are in sheet structures flanking a helix found in both Model 1 and 5VJ2. NS1 54-57-60 (Figure 4D) is in a sheet structure, approximately

midpoint in the amino acid sequence, while NS1 58-66 (Figure 4D) straddles this same sheet structure and an adjacent coil region found in Model 1 and 5VJ2. Lastly, NS1 98-109 and NS1 98-104st (Figure 4E and 4F), situated in the last third of the NS1 protein, are in a coil structure preceding a carboxy terminal helix virtually identical between Model 1 and 5VJ2. Among Group 1 viruses, NS1 06-09 and NS1 11-13-15 were differentially described as located in a helix secondary structure in Model 1 and a Sheet in 5VJ2. However, alanine replacements of the selected amino acids at both locations were predicted to be generally destabilizing by both stability and ddG assessments (Table 4). The more negative ddG predicted values were, the greater the likelihood was of a predicted decreased stability. The secondary structure predictions in the remaining Group 1 sequences (NS1 29, NS1 47, NS1 54-57-60, NS1 58-66 and NS1 98-109) were in agreement between Model 1 and 5VJ2 as well as the relationship between negative ddG values and decreased stability with the exception of NS1 98-109. Here alanine substitutions of cysteines produced disparate predictions. Stability wise, these changes generated the same prediction per amino acid position (NS1 98, **Increased**, and NS 109, **Decreased**) in both Model 1 and 5VJ2 yet ddG predicted values were extremely negative at either position only in Model 1 (Table 4). Stability or ddG predictions among the single amino acid replacements at position 6 (Group 2 viruses) were all generally predicted as destabilizing with the exception of NS1 06 W where the ddG was barely negative. It is worth noting that the four Group 2 mutants had a range of increasingly negative ddG scores (NS1 06 W, -0.61; NS1 06 Q, -1.59; NS1 06 R, -2.24; and NS1 06 D, -3.38) in Model 1.

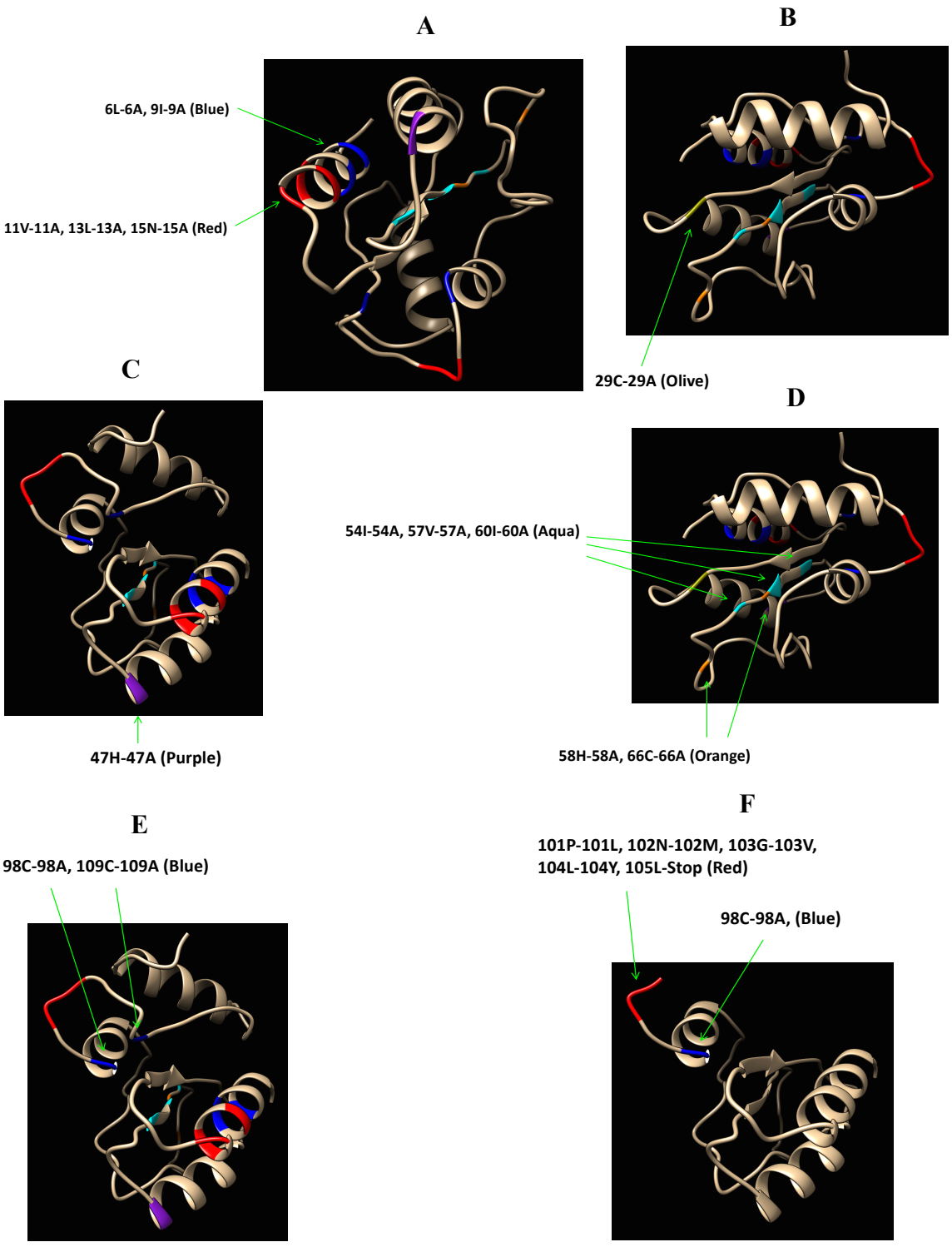

A
6L-6A, 9I-9A (Blue)
11V-11A, 13L-13A, 15N-15A (Red)
B
29C-29A (Olive)
C
47H-47A (Purple)
D
54I-54A, 57V-57A, 60I-60A (Aqua)
58H-58A, 66C-66A (Orange)
E
98C-98A, 109C-109A (Blue)
F
101P-101L, 102N-102M, 103G-103V,
104L-104Y, 105L-Stop (Red)
98C-98A, (Blue)

Figure 15 Location of NS1 substitution mutants in Model 1 (A-F)
A) NS1 06-09 and NS1 11-13-15. B) NS1 29. C) NS1 47. D) NS1 54-57-60 and NS1 58-66. E) NS1 98-109. F) NS1 98-104st.

6.5 Discussion

Computational tools can be expected to improve the empirical "trial and error" process of vaccine candidate selection. Such objectivity is apparent in more recent vaccine design methods (codon optimization, codon-pair optimization, protein structure selection) [89,186–189] and has helped the field move towards more rational target selection/modification [190] when developing new vaccines. In the mutational analysis of the Model 1 structure involving machine learning, the AUTO-MUTE 2.0 prediction algorithm combined knowledge-based four body statistical potentials with machine learning to generate predictions of the effects of single amino acid substitutions on protein structure stability. Thermal stability and free energy prediction programs were chosen to highlight specific substitution/locations in NS1. The mutations in Group 1 were chosen based on the predicted thermal stability change (**Increased** or **Decreased**) in combination with average ddG values among functionally specific amino acid groups. The selection was further restricted to locations in different secondary structures found within the first, second or third regions of the linear structure of Model 1. In general, the more negative ddG prediction values were, the more likely predicted stability was labeled **Decreased** except for NS1 98-104st. The discrepancy between mixed stability labels (NS1 98, **Increased**; NS1 109, **Decreased**) and extremely negative ddG predicted values at either position in Model 1 which may reflect a deficit in either determining the proper structure geometry and/or the machine learning applied to unstructured (Coil) protein

regions. The second group (Group 2) of mutations involved four different amino acid substitutions at a single position, amino acid 6 exhibited a range of increasingly negative ddG scores providing the possibility of assessing the effect of progressively increased thermal instability at this amino acid position. To the best of our knowledge, this is the first example in which the AUTO-MUTE 2.0 prediction algorithm was used to make protein structure predictions that were later characterized for actual protein function, using viable virus in which a number of specific activities could be assessed.

CHAPTER 7. VIRUS GROWTH (MULTICYCLE GROWTH)

7.1 Abstract

The AUTO-MUTE 2.0 algorithm [184,185] categorized the effect of amino acid changes on the stability of a modeled NS1 protein structure (Model 1) and a set of these selected amino acid changes were incorporated into the NS1 sequence of replicating virus for functional evaluation. While these amino acid changes can reasonably be equated to "greater than/equal to" or "less than" respectively in terms of affecting NS1 function during a virus infection, this has not been tested experimentally. Evaluation of NS1 function (which is inaccurately know at best) would include comprehensive measurements of virus phenotypes known to be associated with NS1. Multicycle growth was measured for each virus in both Vero (non-human, African green monkey kidney epithelial IFN incompetent cell line) and A549 (human Caucasian male adenocarcinoma human alveolar basal epithelial IFN competent cell line) to see the effects of NS1 amino acid substitutions on multicycle virus growth.

7.2 Introduction

The function of RSV NS1 during viral infection is likely to be multifaceted. It is the first gene coded for in the viral genome, is the earliest and most abundant protein produced during infection [7] and the association of NS1 with virus growth is well accepted [11–14]. An evaluation of mutant effects on NS1 function required the design, development, and production of RSV viruses each containing the identified amino acid substitutions. One of the functions associated with NS1 involves its ability to antagonize

the innate immune response during infection of tissue culture cells, therefore multicycle growth was measured for each virus in both Vero and A549 cell lines. The presence of functional NS1 is reported to help virus counter infected cell innate immune responses activated during RSV infection which are not antagonized during NS1 deletion virus infection, resulting in among other things reduced virus production. This work looks to measure and quantify the effects on multicycle virus growth in relation to the various NS1 amino acid substitutions.

7.3 Material and methods

NS1 amino acid substitutions

For the first set of substitutions, Group 1, DNA oligonucleotides were designed and used to introduce specific nucleotide sequence changes leading to single or multiple amino acid (aa) sequence substitutions from wild-type amino acid assignments to Alanine (A) a non-polar, aliphatic amino acid. The oligonucleotide sequences for these changes are listed below:

```
aa6 (L) and 9 (I) both to (A)
tm_NS1_053
ATTTAACTCCCTTGGTTAGAGATGGGCAGCAATTCAGCCAGTATGGCCAAAGTTAGATTACAAAATTTGTT
TGACAATGATGAAG
tm_NS1_054
CTTCATCATTGTCAAACAAATTTTGTAATCTAACTTTGGCCATACTGGCTGAATTGCTGCCCATCTCTAAC
CAAGGGAGTTAAAT

aa11 (V), 13 (L) and 15 (N) each to (A)
tm_NS1_055
CTTGGTTAGAGATGGGCAGCAATTCATTGAGTATGATAAAAGCCAGAGCCCAAGCCTTGTTTGACAATGAT
GAAGTAGCATTGTTAAAAATAACATG
tm_NS1_056
CATGTTATTTTTAACAATGCTACTTCATCATTGTCAAACAAGGCTTGGGCTCTGGCTTTTATCATACTCAA
TGAATTGCTGCCCATCTCTAACCAAG
```

aa54 (I), 57 (V) and 60 (I) each to (A)
tm_NS1_061
GCTAAGGCAGTGATACATACAATCAAATTGAATGGCGCCGTGTTTGCCCATGTTGCCACAAGTAGTGATAT
TTGCCCTAATAATAATATTGT
tm_NS1_062
ACAATATTATTATTAGGGCAAATATCACTACTTGTGGCAACATGGGCAAACACGGCGCCATTCAATTTGAT
TGTATGTATCACTGCCTTAGC

aa58 (H) and 66 (C) both to (A)
tm_NS1_049
GTTTGTGGCCGTTATTACAAGTAGTGATATTGCCCCTAATAATAATATTG
tm_NS1_050
CAATATTATTATTAGGGGCAATATCACTACTTGTAATAACGGCCACAAAC

aa98 (C) and 109 (C) both to (A)
tm_NS1_051
GAATTAACACATGCCTCTCAACCTAATGGTCTACTAGATGACAATGCCGAAATTAAATTC
tm_NS1_052
GAATTTAATTTCGGCATTGTCATCTAGTAGACCATTAGGTTGAGAGGCATGTGTTAATTC

In a second set, Group 2, DNA oligonucleotides were designed and used to introduce specific nucleotide sequence changes leading to a set of single amino acid sequence substitutions focused on amino acid number 6, Group 2. As a result, amino acid 6, Leucine (L) was changed to Tryptophan (W), Arginine (R), Aspartic Acid (D) or Glutamine (Q) which changes a non-polar (L) to either a non-polar aliphatic (W), charged aliphatic (R), polar-charged (D) or neutral polar (Q) amino acid. The oligonucleotide sequences for these changes are listed below:

aa06 (L-W)
tm_NS1_001a ATGGGCAGCAATTCATGGAGTATGATAAAAGTTAGATTACAAAAT
tm_NS1_001b ATTTTGTAATCTAACTTTTATCATACTCCATGAATTGCTGCCCAT

aa06 (L-R)
tm_NS1_002a ATGGGCAGCAATTCACGTAGTATGATAAAAGTTAGATTACAAAAT
tm_NS1_002b ATTTTGTAATCTAACTTTTATCATACTACGTGAATTGCTGCCCAT

aa06 (L-D)
tm_NS1_003a ATGGGCAGCAATTCAGATAGTATGATAAAAGTTAGATTACAAAAT
tm_NS1_003b ATTTTGTAATCTAACTTTTATCATACTATCTGAATTGCTGCCCAT

aa06 (L-Q)
tm_NS1_004a ATGGGCAGCAATTCACAGAGTATGATAAAAGTTAGATTACAAAAT
tm_NS1_004b ATTTTGTAATCTAACTTTTATCATACTCTGTGAATTGCTGCCCAT

An additional set of DNA oligonucleotides were designed and used to amplify PCR fragments from cDNA made from vRNA.

TM1399
ACGCGAAAAAATGCGTACAACAAAC
(RSV nucleotide 1 – 25)

TM1288
TCTGTCCAACTCTGCAGCTC
(RSV nucleotide 8446 – 8465)

TM969
TGCATCAATATCTCAAGTCAACG
(RSV nucleotide 7783 – 7805)

TM1400
ACGAGAAAAAAAGTGTCAAAAACTAATATC
(RSV nucleotide 14849 – 15878)

A plasmid containing RSV genome sequence from leader region through nonstructural protein 2 (NS2) and flanked by NotI (before leader) and KpnI (after NS2); pD46 6120 NS1//NS2, (gift from Dr. Buchholz) was used in the site-directed mutagenesis protocol (Stratagene) as recommended by the manufacturer. The NotI-KpnI fragment from each plasmid identified with appropriate nucleotide changes, was cut with Not I and KpnI (FastDigest-Thermo Scientific) as recommended by the manufacturer and ligated into a similarly restriction endonuclease digested plasmid containing full length RSV genome sequence with enhanced green fluorescent protein (eGFP) gene sequence inserted between phosphoprotein (P) and matrix (M) genes [73]. All full-length RSV genome plasmids containing the NS1 modifications are sequenced in their entirety to ensure complete sequence integrity. The sequencing oligos are listed below:

Name	Sequence	Gene
1F_WT	ACGCGAAAAAATGCGTACAACAAAC	
7F_WT	AAAAATGCGTACAACAAACTTGC	3' UTR
21F_WT	CAAACTTGCATAAACCAAAAAAATG	3' UTR
251F_WT	GAATGGCATTGTGTTTGTGC	NS1
607F_WT	ATCAATTCAGCCAACCCAAC	NS1-NS2
1120F_WT	GAAGATGGGGCAAATACAACC	NS2-N
1647F_WT	TAAATTAGCAGCAGGGGACAG	N
1705F_WT	GTCCTAAAAAATGAAATGAAACGTTAC	N
2013F_WT	CCAAAAATTGGGTGGTGAAG	N
2419F_WT	AAGGGCAAATTCACATCACC	P
2850F_WT	TGCCATGGTTGGTTTAAGAG	P
3190F_WT	AGCCAATCCAAAACTAACCAC	P-M
3554F_WT	CGCTAATGTGTCCTTGGATG	M
3830F_WT	CAAATGCAAAAATCATCCCTTAC	M
4034F_WT	CTTTTTCCTCTACATCAGTGTGTTAATTC	M
4220F_WT	GGGGCAAATAATCATTGGAG	SH gene start
4564F_WT	GGCAAATGCAAACATGTCC	G gene start
4745F_WT	GCAGCCATCATATTCATAGCC	G
5213F_WT	AAAGATCCCAAACCTCAAACC	G
5536F_WT	TGGGGCAAATAACAATGGAG	F gene start
5689F_WT	CTCTGAGAACTGGTTGGTATAC	F
5999F_WT	TGGCGTTGCTGTATCTAAGG	F
6348F_WT	ACAAATGATCAGAAAAAGTTAATGTCC	F
6754F_WT	TAAGCAGCTCCGTTATCACATC	F
7018F_WT	CATCAATATCTCAAGTCAACGAG	F
7589F_WT	CCCATGCACTGCTTGTAAGAC	M21
7608F_WT	CCCATGCACTGCTTGTAAGA	M21
7949F_WT	TGCCAGCAGACGTATTGAAG	M21
8068F_WT	AAATGACCATGCCAAAAATAATG	M21/M22
8323F_WT	CTCAATTCTAACACTCACCACATC	M22-L
8808F_WT	CCAACAATGGACAAGATGAAG	L
9280F_WT	CTTAGGCTTAAGATGCGGATTC	L
9611F_WT	AAGCTTGCAGGTGACAATAACC	L
9949F_WT	CTATCGTGAGTTTCGGTTGC	L
10486F_WT	TCTCAGCAAATTCAATCAAGC	L
10842F_WT	AACCAATCAGACTCATGGAAGG	L
11200F_WT	TCTACAATTAAAAAATCATGCATTATG	L

11622F_WT	TGGCAGTTACAGAGGTTTTGAG	L
11877F_WT	GAGCCACTGAGATGATGAGG	L
12137F_WT	TTAACACGTGGTGAGAGAGGAC	L
12422F_WT	CCATGTGAATTCCCTGCATC	L
12871F_WT	TTTAGCTGGACATTGGATTCTG	L
13240F_WT	AGAATTCACAGTTTGCCCTTG	L
13648F_WT	GTCTGCAATGATTAGAACCAATTAC	L
14013F_WT	GAAGTCTGAAAGATTGCAATGATC	L
14404F_WT	CTTAGTCCTTACAATAGGTCCTG	L
14807F_WT	AACAGCTTGACAACCAATGAAC	L
322R_WT	TGGCATTGTTGTGAAATTGG	NS1
672R_WT	GACAACGGTCTCATGTCTGTG	NS2
1203R_WT	TCCGTTGGATGGTGTATTTG	N
1731R_WT	TCCTTGGGTAGTAAGCCTTTG	N
2151R_WT	TGATTCCTCGGTGTACCTCTG	N
2557R_WT	GGCTTGTTCCCTGCAGTATC	P
2794R_WT	GCACTTGCCACTACTAATGTGTG	P
3015R_WT	TTCCCTTCCAATAGGTTGTTC	P
3288R_WT	CTGTGTATGTGGAGCCTTCG	M
3378R_WT	TCTGCTGGCATAGATGATTGG	M
3971R_WT	GTGCTTCCAATTTGTGGTAACAT	M
4287R_WT	TCCATTGGTTGATTCTGTATGG	SH gene start/SH
4621R_WT	TGATTGAGAGTGTCCCAGGTC	G
4869R_WT	AGACGGATTAGAGGGACTGATTC	G
5008R_WT	GTTGTTTTGTGGTGGGCTTG	G
5418R_WT	TGATGGGTACTCGGATGTTG	G
5854R_WT	TCTGGCTCGATTGTTTGTTG	F
6374R_WT	GGACATTAACTTTTTCTGATCATTTGT	F
6577R_WT	AAAGATACTGATCCTGCATTGTCAC	F
6819R_WT	TCCACGATTTTTATTGGATGC	F
7033R_WT	GCTCTGGTTAATCTTCTCGTTGA	F
7679R_WT	TCTGTCCAACTCTGCAGCTC	M21
7833R_WT	GCTCTTCATTGTCCCTCAGC	M21
8045R_WT	TATTTTTGGCATGGTCATTTG	M21/M22
8244R_WT	ATCAATTCTTGAGAGGTCCAATG	M22
8677R_WT	GCAATCTGTTCTGACGAGGTC	L
8978R_WT	TAACCATGATGGAGGATGTTG	L
9466R_WT	GCATCTGTGATGTTGTTGAGC	L

9696R_WT	TAACAGCATCCATGGCTTGTC	L
10205R_WT	CCACATGATTAGGGTTGTTGAG	L
10618R_WT	GGGGTGCATGCCTATATGTG	L
11047R_WT	CCACGGTCCCACTCTTAGG	L
11390R_WT	AGCCTCTGTGAGGAAGTCAGG	L
11880R_WT	TTTTTCCTCATCATCTCAGTGG	L
12263R_WT	TACACCCAATCCAATTTTGC	L
12644R_WT	ATGGGAGGTTTCATCAAATG	L
13028R_WT	GCTTTGCTTTGCCATAACC	L
13250R_WT	TAACAACCCAAGGGCAAAC	L
13647R_WT	TTGGTTCTAATCATTGCAGACG	L
13803R_WT	GCATGCAGTAAAGTGATGTGC	L
14192R_WT	TGTTACAGACAATTCGGCATC	L
14484R_WT	AGGCATGATGAAATTTTTGG	L
14809R_WT	AAGTTCATTGGTTGTCAAGCTG	L
15089R_WT	CTAATATCTCGTAATTTAGTTAATACAC	5'-trailer
TM1400	ACGAGAAAAAAAGTGTCAAAAACTAATATC	5'-trailer

Reverse Genetic System Recovery of RSV Containing NS1 substitutions

The RSV reverse genetics system [124,144,196] was used to recover live virus containing NS1 modifications. Helper plasmids expressing the following RSV gene products; nucleoprotein (N) pTM1-N, phosphoprotein (P) pTM1-P, M2 ORF 1 protein (M2-1) pTM1-M2-1 and large protein (L) pTM1-L are transfected along with each NS1 modified full-length viral genome plasmid into BSR T7/5 cells, baby hamster kidney 21 (BHK21) that constitutively express T7 RNA polymerase [197]. This cell plus virus material was scrapped and transferred to Vero cells for virus expansion. When recovering/expanding virus on Vero cells in T225 flasks, infections in 10 ml volumes were left on cells for 1-2 hours and then an additional 15 ml of OptiPRO SFM, 4 mM L-glutamine media was added (final volume of 25 ml per T225 flask) and left at 37°C; 5%

CO_2 until virus harvest. Virus grown on Vero cells are recovered, tittered and used to generate working virus stocks with PFU/ml values above 10^6 [198].

Cell Lines

Cell lines used include BSR T7/5 cells, a baby hamster kidney 21 (BHK21) cells that constitutively express T7 RNA polymerase [197]. These cells were maintained in Glasgow minimal essential medium (GMEM) (Gibco-Life Technologies) supplemented with 10% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies) and 2% MEM amino acids (Gibco-Invitrogen). Every other passage, the medium was supplemented with 2% Geneticin (Gibco-Invitrogen) to ensure cells retain the T7 polymerase construct. Vero cells (normal adult African green monkey kidney cell) (ATCC CCL-81) were maintained and infected in OptiPRO SFM (Gibco-Life Technologies), 4 mM L-glutamine (Gibco-Life Technologies). A549 cells (adenocarcinoma human alveolar basal epithelial cell) (ATCC CCL-185) were maintained in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 10% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies).

Virus Infections

For multicycle virus infections, virus stocks were diluted to a multiplicity of infection of 0.1 (MOI = 0.1) pfu/cell in OptiPRO SFM (Gibco-Life Technologies), 4 mM L-glutamine (Gibco-Life Technologies) and used to infect duplicate or triplicate wells in a 12 well plate of either A549 or Vero cells. Infections were left on cells for 1-2 hours with periodic rocking at 37°C; 5% CO_2. Subsequently the infectious volume is removed, monolayers washed with OptiPRO SFM, 4 mM L-glutamine media and refed with either

F-12K, 2% FBS, 2 mM L-glutamine media (A549 cell infection) or OptiPRO SFM, 4 mM L-glutamine media (Vero cell infection). Infected cells were left at 37°C; 5% CO_2 until sample collection by scraping [198].

Virus Titration

For all virus titrations, virus stocks were tittered as described. In short 24 well plates of Vero cells were infected with ten-fold virus dilutions in OptiPRO SFM, 4 mM L-glutamine. Infections were left on cells for 1-2 hours with periodic rocking at 37°C; 5% CO_2. Subsequently approximately 1 ml of 0.8% Methyl cellulose overlay [198] was added to each well and returned to incubator at 37°C; 5% CO_2 until visible plaque formation and/or green fluorescent protein expression in virus infected cells was detectable. Growth levels of virus stocks were qualified by counting eGFP expressing virus plaques (syncytia) from scanned images generated using an Amersham Typhoon Imaging System (GE Healthcare Life Sciences) as recommended by the manufacturer. Image files were analyzed using ImageJ (Version 1.52d NIH/Wayne Rasband) to count the number of green fluorescent protein expressing plaques as recommended. These counts were used to calculate the amount of viable virus in each sample and recorded as plaque forming units per ml (PFU/ml).

Site-Directed Mutagenesis

Site-directed mutagenesis of specific nucleotides in Group 1 and Group 2 viruses was accomplished using QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies) as recommended by the manufacturer.

Thermocycle Reactions

All thermocycler reactions were done using a ProFlex PCR System (Applied Biosystems) as recommended by the manufacturer.

Plasmid DNA Extraction

Plasmid DNA were either extracted using QIAprep Spin Miniprep Kit (Qiagen) or NucleoBond Xtra Maxi EF Endotoxin-free plasmid DNA purification (Macherey-Nagel) for 2 ml or 150 ml bacteria cultures respectively as recommended by the manufacturer. Plasmid DNA was Sanger sequenced (Eurofins Genomics Sequencing) to verify sequence integrity.

Viral RNA Extraction

Recovered viruses had viral RNA (vRNA) extracted using QIAamp Viral RNA Mini Kit (Qiagen) as recommended by the manufacturer. The vRNA was converted to complementary DNA (cDNA) using Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Scientific) as recommended by the manufacturer. The cDNA was used to generate two large PCR products using SequalPrep Long PCR Kit with dNTPs (Invitrogen) as recommended by the manufacturer. Fragment A is between PCR primers TM1399 and TM1283 (genome nucleotides 1 – 8564). Fragment B is between PCR primers TM969 and TM1400 (genome nucleotides 7783 – 15878). These PCR products were Sanger sequenced (Eurofins Genomics Sequencing) to verify sequence integrity.

7.4 Results

During the site directed mutagenesis work to create NS1_98-109, an additional virus, NS1_98_101-102-103-104st, was inadvertently created due to a site directed

mutagenesis PCR error. Sequencing indicated that the codon for aa 98 was correctly changed but due to the PCR error, aa 101, 102, 103 and 104 were also recoded (P-L, N-M, G-V, and L-Y) followed by a stop codon after aa 104. Therefore, in this virus NS1 protein is recoded at aa 98, 101, 102, 103, 104 and truncated after aa 104. In the remainder of the manuscript the virus will be referred to as NS1_98-104st.

Multicycle virus growth measured in Vero cells peaked on day 4 (Group 1) or 5 (Group 2) among all viruses. RSV wt (wt) grew to a peak titer of 3.6×10^7 pfu/ml followed by Group 1 viruses NS1 29 and NS1 98-109 with wt-like peak growth titers of 2.3×10^7 and 1.6×10^7 pfu/ml respectively. Peak virus titers in Vero cells among an intermediate growth set of Group 1 viruses, NS1 47 and NS1 58-66, was 9.8×10^6 and 9.3×10^6 pfu/ml respectively, about 0.4 $\log_{10}$ lower than the wt-like group. The remaining Group 1 viruses, NS1 06-09, NS 11-13-15, NS1 54-57-60 and NS1 98-104st as well as ΔNS1, exhibited growth restriction in Vero cells on day 4 with virus titers at 6.3×10^6, 5.6×10^6, 4.5×10^6, 3.3×10^6 and 7.5×10^6 pfu/ml respectively and were about 0.5-0.9 $\log_{10}$ lower than peak wt-like virus titers.

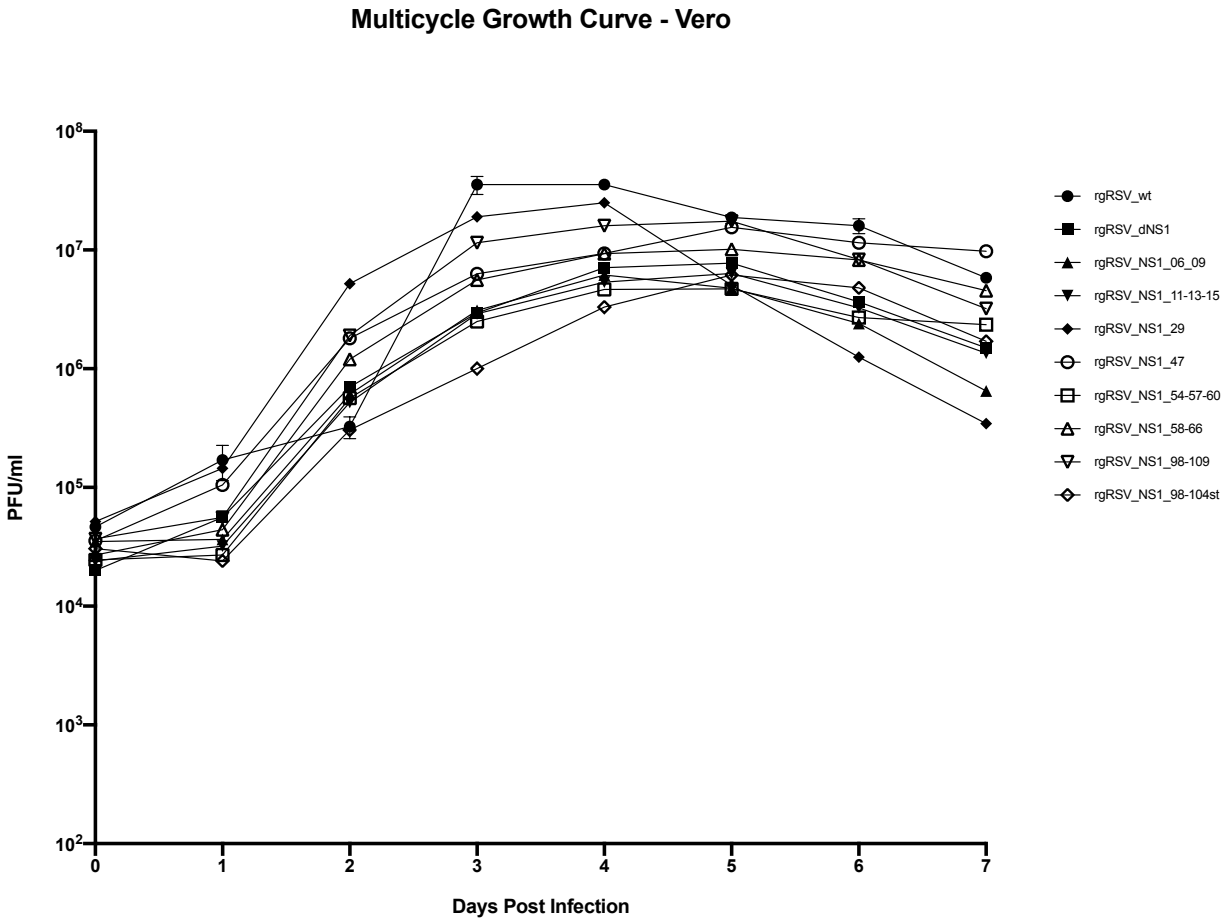


Figure 16 Multicycle growth curve of Vero cells infected with Group 1 NS1 substitution viruses
Monolayer cultures of Vero cells were infected in triplicate with the Group 1 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 0.1 PFU/cell and incubated at 37ºC for 7 days. Infected cells were scraped into the growth medium. After a brief vortex, clarified medium supernatants were analyzed by virus titration of 10-fold dilution series in Vero cells under methylcellulose and quantified by counting eGFP plaques. Data points represent an average of the plaque counts; standard deviation is show as horizontal lines above and below each data point. X-axis is days post infection and Y-axis is $\log_{10}$ scaled PFU/ml.

Separately, Group 2 viruses NS1 06-Q, NS1 06-R and NS1 06-W grew as well as wt (day 5 titer of 1.2 x 10^7 pfu/ml) with peak titers at day 5 of 1.3 x 10^7, 1.8 x 10^7 and 1.3 x 10^7 pfu/ml respectively. Only NS1 06-D grew as poorly in Vero cells as ΔNS1 with

day 5 titers of 2.5 x 10^6 and 1.9 x 10^6 pfu/ml respectively or about 0.8 $\log_{10}$ lower than the wt-like viruses.

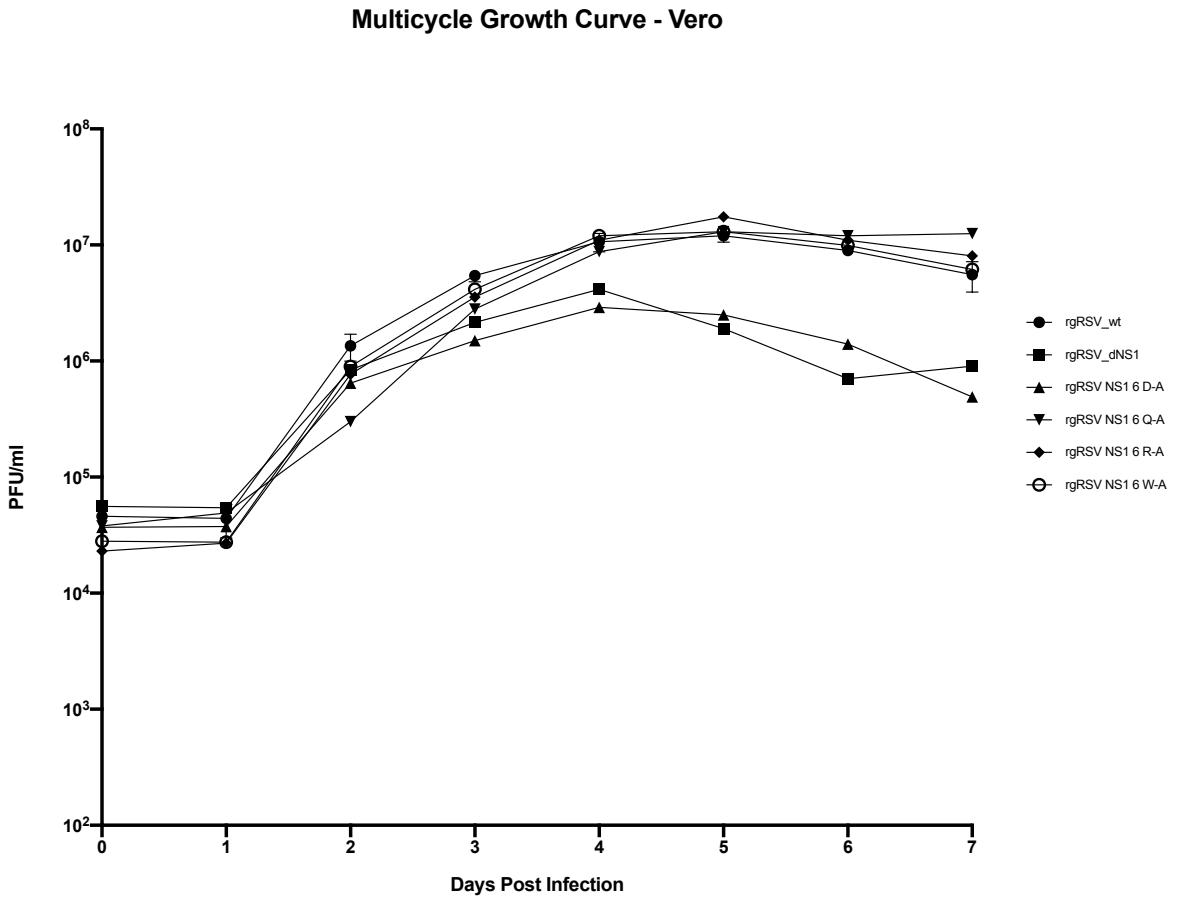


Figure 17 Multicycle growth curve of Vero cells infected with Group 2 NS1 substitution viruses
Monolayer cultures of Vero cells were infected in triplicate with the Group 2 NS1 mutant viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 0.1 PFU/cell and incubated at 37ºC for 7 days. Infected cells were scraped into the growth medium. After a brief vortex, clarified medium supernatants were analyzed by virus titration of 10-fold dilution series in Vero cells under methylcellulose and quantified by counting eGFP plaques. Data points represent an average of the plaque counts; standard deviation is show as horizontal lines above and below each data point. X-axis is days post infection and Y-axis is $\log_{10}$ scaled PFU/ml.

Multicycle growth among all viruses was also measured in a more relevant, interferon competent cell type, A549. Despite inconsistent day 2 virus titers, days 3-4

appeared more indicative of any consequence a defective NS1 would have on virus growth. RSV wt grew optimally through day 4 with a peak titer at 1.7 x 10^6 pfu/ml in A549 cells. There was an anomalously low virus titer for RSV wt at day 2, but likely not representative. Like wt virus, Group 1 virus NS1 29 had a peak titer at day 2 (1.6 x 10^6 pfu/ml) that was maintained through day 4 (1.4 x 10^6 pfu/ml), followed by NS1 98-109 virus with a day 2 peak titer of 5.0 x 10^5 pfu/ml dropping to 3.4 x 10^5 pfu/ml by day 4, below RSV wt virus growth level in A549. The remaining Group 1 viruses, NS1 47, NS1 58-66, NS1 11-13-15, NS1 06-09, NS1 98-104st, NS1 54-57-60 as well as ΔNS1, had day 2 titers of 1.6 x 10^5, 1.0 x 10^5, 6.0 x 10^4, 4.8 x 10^4, 3.8 x 10^4, 3.4 x 10^4, and 1.5 x 10^4 pfu/ml respectively. The peak growth at day 2 of these viruses was reduced 1-2 $\log_{10}$ compared to wt growth and continued to decline further over time (day 5 titer reductions are 1.7, 1.9, 2.5, 1.7, 2.3, 2.9 and 2.8 $\log_{10}$ respectively).

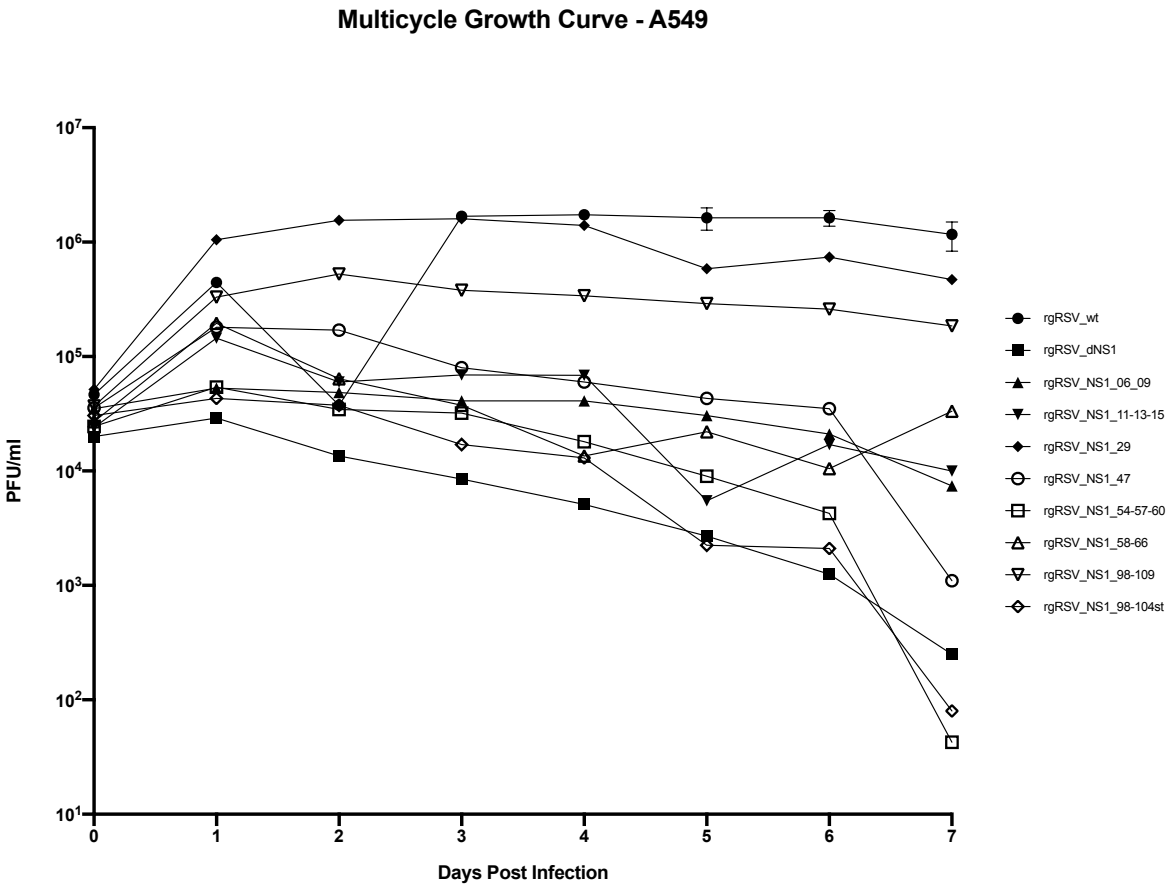


Figure 18 Multicycle growth curve of A549 cells infected with Group 1 NS1 substitution viruses
Monolayer cultures of A549 cells were infected in triplicate with the Group 1 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 0.1 PFU/cell and incubated at 37ºC for 7 days. Infected cells were scraped into the growth medium. After a brief vortex, clarified medium supernatants were analyzed by virus titration of 10-fold dilution series in Vero cells under methylcellulose and quantified by counting eGFP plaques. Data points represent an average of the plaque counts; standard deviation is show as horizontal lines above and below each data point. X-axis is days post infection and Y-axis is $\log_{10}$ scaled PFU/ml.

Multicycle growth in Group 2 viruses were tested separately relative to RSV wt and ΔNS1 viruses for growth in A549 cells. RSV wt virus had a day 2 peak virus titer of 1.5 x 10^5 pfu/ml while Group 2 viruses NS1 06-W, NS1 06-R, NS1 06-Q and NS1 06-D) grew almost as well as RSV wt on day 2 (1.3 x 10^5, 1.1 x 10^5, 8.5 x 10^4 and 8.2 x 10^4

pfu/ml respectively), but became moderately reduced about 0.5 log_{10} by day 5 compared to wt. In contrast, ΔNS1 was reduced almost 1.3 log_{10} by day 5.

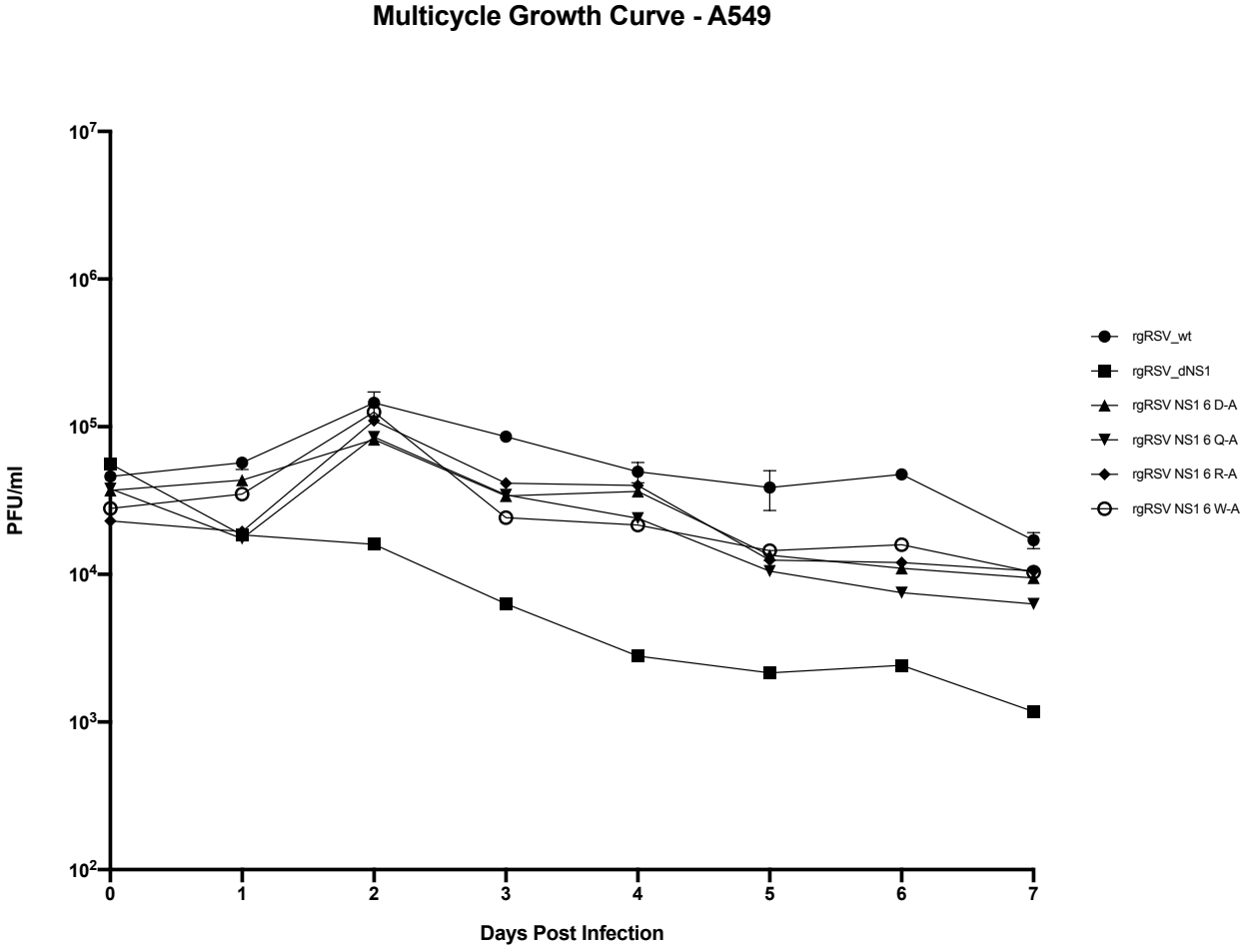


Figure 19 Multicycle growth curve of A549 cells infected with Group 2 NS1 substitution viruses
Monolayer cultures of A549 cells were infected in triplicate with the Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 0.1 PFU/cell and incubated at 37ºC for 7 days. Infected cells were scraped into the growth medium. After a brief vortex, clarified medium supernatants were analyzed by virus titration of 10-fold dilution series in Vero cells under methylcellulose and quantified by counting eGFP plaques. Data points represent an average of the plaque counts; standard deviation is show as horizontal lines above and below each data point. X-axis is days post infection and Y-axis is log_{10} scaled PFU/ml.

7.5 Discussion

The function of RSV NS1 during viral infection is likely to be multifaceted as it is the first gene coded for in the viral genome and is the earliest and most abundant protein produced during infection [7]. The presence of functional NS1 is reported to help virus counter innate immune responses in cells activated during RSV infection which during NS1 deletion virus infection are not antagonized and result in among other things reduced virus production [16,110,199–201]. As such, in vitro criteria used to characterize previous RSV vaccine candidates seem appropriate to characterize these new designs and whether they correlate with machine learning predictions. Multicycle growth in cell culture systems of Vero or A549 cells was used as an initial in vitro characterization of growth attenuation levels among the vaccine candidates. Virus growth in Vero cell cultures affords insights as to whether any profound growth defects may exist such that the attenuated growth is below a minimum level required for vaccine manufacture. As expected, RSV wt (positive control) virus exhibited peak growth titer of 3.6×10^7 pfu/ml (Figure 5A) at day 4 followed closed by NS1 29 and NS1 98-109 (2.3×10^7 and 1.6×10^7 pfu/ml respectively). In the case of NS1 29, its growth level seemed to agree with the prediction that such an amino acid change to NS1 did not significantly lessen its growth ability in Vero cells. This result is suggestive that at least a prediction of **Increased** thermal stability of NS1 does not lead to less than wt like virus growth levels. The other virus, NS1 98-109, with almost wt like growth levels in Vero cells, was predicted to be a combination of **Increased** (amino acid 98 substitution) and **Decreased** (amino acid 109 substitution) in terms of protein thermal stability. While the level of this virus growth was above 10^7 the heterogeneously predicted

amino acid substitutions appeared to slightly lessen this virus's ability to grow in Vero cells. Another set of viruses, NS1 47 and NS1 58-66, follow with an intermediate day 4 peak level of growth in Vero cells (9.8 x 10^6 and 9.3 x 10^6 pfu/ml respectively). In this case the NS1 47 virus growth level seemed to agree with the NS1 47 substitution prediction of **Decreased**. As seen before with NS1 98-109, NS1 58-66 had a heterogeneous thermostability prediction of **Decreased** (amino acid 58 substitution) and **Increased** (amino acid substitution 66). But in this substitution set, the **Decreased** prediction seemed to be associated with decreased virus growth in Vero cells unlike NS1 98-109 virus.

Among the final members of Group 1 substitutions, NS1 06-09, NS 11-13-15, NS1 54-57-60 and NS1 98-104st, with all or mostly **Decreased** thermal stability predictions, the day 4 peak virus growth levels were more similar to ΔNS1 (negative control) virus growth levels (7.5 x 10^6, 6.3 x 10^6, 5.6 x 10^6, 4.5 x 10^6, and 3.3 x 10^6 pfu/ml respectively). In this final set of Group 1 it appeared the predicted **Decreased** thermal stability prediction was better associated with decreased virus growth in Vero cells. While the thermal stability prediction and actual virus growth appeared associated for the various Alanine substitutions made among the various regions throughout NS1 structure, this agreement appeared less well supported in Group 2. All predictions for 6L replacements (6D, 6R, 6Q and 6W) were labeled **Decreased**, yet NS1 06-Q, NS1 06-R and NS1 06-W all had RSV wt like day 5 peak titers (1.3 x 10^7, 1.8 x 10^7, 1.3 x 10^7 and 1.2 x 10^7 pfu/ml respectively). Only NS1 06-D had reduced virus growth similar to ΔNS1 (2.5 x 10^6 and 1.9 x 10^6 pfu/ml respectively). Whether this contradiction in Group 2 amino acid substitutions predictions was unique to the single amino acid 6 of NS1 was unclear but may relate to previous reports

of NS1 influence on RSV replication [14].

Virus growth in A549 cells however is more likely a better measure of any impacts NS1 substitutions may have on virus growth. A549 cells possess an intact interferon beta locus unlike Vero cells and therefore more fully initiate a functional innate cell immune response to viral infection approximating what most human RNA negative sense nonsegmented single strand viruses that infect the lungs and upper airways have evolved against. Combined with the suggestion that RSV NS1 protein is considered a virally encoded interferon antagonist, multicycle growth measured in interferon competent A549 cells would be yet another important initial in vitro characterization assay. As there was an anomalously low virus titer for RSV wt at day 2, but likely not representative, the RSV wt day 3-5 average titer of 1.7×10^6 pfu/ml was used as an optimal growth upper bound. As seen in Vero cells, NS1 29 and NS1 98-109 had peak titers at day 2 (1.6×10^6 and 5.0×10^5 pfu/ml respectively). NS1 29 growth level would further seem to agree with the prediction that such an amino acid change to NS1 did not significantly lessen its growth ability in A549 cells. NS1 98-109, with lower than wt like growth levels in A549 cells, offered clarification to the NS1 substitution prediction meaning in the context of virus growth but in an interferon competent environment. The heterogeneous NS1 98-109 prediction, both **Increased** (amino acid 98 substitution) and **Decreased** (amino acid 109 substitution) seemed to trend towards decreased virus growth relative to RSV wt, furthering what was seen in Vero cells. Among the remaining members of Group 1 (NS1 47, NS1 58-66, NS1 11-13-15, NS1 06-09, NS1 54-57-60, NS1 98-104st and ΔNS1), the majority had thermal stability labels of **Decreased**. While NS1 58-66 had a heterogeneous

thermostability prediction like NS1 98-109, it grew considerably less well in A549 cells as compared to NS1 98-109. In fact, peak growth at day 2 in this second set of Group 1 viruses (with **Decreased** thermal stability prediction) was reduced 1-2 $\log_{10}$ compared to wt growth and continued to decline further over time (day 5 titer reductions are 1.7, 1.9, 2.5, 1.7, 2.3, 2.9 and 2.8 $\log_{10}$ respectively).

Among the Group 2 substitution set, in which all predictions for 6L replacements (6D, 6R, 6Q and 6W) were labeled **Decreased**, the levels of virus growth were close to, but below the RSV wt day 2 peak virus titer of 1.5×10^5 pfu/ml (Figure 5D). NS1 06-W, NS1 06-R, NS1 06-Q and NS1 06-D all had peak growth titers on day 2, but became progressively more reduced, about 0.5 $\log_{10}$, by day 5 compared to wt. None of Group 2 substitutions however grew as poorly in A549 cells as ΔNS1 (reduced almost 1.3 $\log_{10}$ by day 5). The A549 data provided additional support that there seems to be an association between the substitution prediction **Decreased** and actual growth attenuation. This gave further context regarding the prediction meanings regarding NS1 functional contribution to virus growth. The alignment between **Decreased** thermal stability prediction for NS1 amino acid substitutions and viral growth attenuation needed further evaluation by additional assays used to characterized previous vaccine candidates.

CHAPTER 8. VIRAL MESSENGER RNA (MRNA) EXPRESSION

8.1 Abstract

It would follow that differences in viable virus production among the various NS1 amino acid substitutions could be related to changes in expressed viral messenger RNA (mRNA) levels [12,13,202,203]. Based on the multicycle growth data in A549 cells, along with controls RSV wt and ΔNS1, the samples have been reordered into an wt-like growth group (RSV wt, NS1 29 and NS1 98-109) and intermediate-low growth group (NS1 47, NS1 11-13-15, NS1 06-09, NS1 58-66, NS1 54-57-60, NS1 98-104st and ΔNS1). Also included as a separate group in this data are the collection of single amino acid substitution viruses in which amino acid 6 (aa 6; Leucine L - non-polar) was substituted for single amino acids with completely different chemical properties (Aspartic Acid D - polar-charged, Arginine R - charged aliphatic, Glutamine Q - neutral polar and Tryptophan W - non-polar aliphatic). Their inclusion allows the examination of whether single amino acid changes alone correlate with a thermal stability prediction. In this work we evaluate any relationship between amino acid substitutions predicted by AUTOMUTE 2.0 and levels of viral mRNA transcription.

8.2 Introduction

The wide range of virus growth among the different substitution viruses suggests a relationship between NS1 and growth. The mixture of prior experiments attesting to NS1 function(s) [204–206] are abstracted from the many viral factors that are present during a viral infection and with which NS1 has evolved. Therefore, examination of

measurable steps during virus infection of tissue culture cells may provide insightful 1) to identify effective/defective NS1 function and 2) to illuminate possible NS1 roles during viral infections. The variable level of virus growth in an IFN competent cell line A549, the known association of IFN production with mRNA transcription and NS1 suggested role as an IFN response antagonist might be exhibited in differential levels of viral mRNA [12,13]. This work will examine a number of important viral gene transcripts using custom designed TaqMan assays to evaluate whether changes in viral gene transcription associate with viral growth and AUTOMUTE 2.0 generated NS1 amino acid substitution predictions.

8.3 Material and methods

Cell Line

A549 cells (adenocarcinoma human alveolar basal epithelial cell) (ATCC) were maintained in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 10% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). After virus infections, A549 cells were maintained in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 2% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). A549 cells provide a relevant experimental system to examine virus growth in a fully IFN competent human lung epithelial host cell system.

Virus Infections

Triplicate experiments of A549 cells in 12 well dishes were infected with virus and samples collected at 18 and 24 hours post infection. For these virus infections, virus stocks were diluted to a multiplicity of infection of 3 (MOI = 3) pfu/cell and in OptiPRO

SFM (Gibco-Life Technologies), 4 mM L-glutamine (Gibco-Life Technologies) and added to cells in appropriate volumes. Infections were left on cells for 1-2 hours with periodic rocking at 37°C; 5% CO_2. Subsequently the infectious volume is removed, monolayers washed with OptiPRO SFM, 4 mM L-glutamine media and refed with F-12K, 2% FBS, 2 mM L-glutamine. Infected cells were left at 37°C; 5% CO_2 until sample collection.

Quantitative RT-PCR

Triplicate experiments of infected A549 cell cultures were collected, and total RNA was isolated with RNeasy Mini Kit (Qiagen) as recommended by the manufacturer, including on-column DNase digestion using a QIAcube sample preparation machine (Qiagen). The total RNA was eluted in 50 ul of EB solution and stored at -80°C until needed. To measure gene transcript levels, a fixed volume of total RNA was combined with a specific TaqMan Gene Expression Assay using the TaqMan RNA-to-Ct 1-Step Kit (Life Technologies) as recommended by the manufacturer. The TaqMan gene expression assay 18S rRNA (Hs99999901_s1) was used as a normalization control in conjunction with the custom TaqMan assays listed below. The custom designed TaqMan assays include:

```
eGFP (EGFP-5)
Fwd   TGTCGCCCTCGAACTTCAC
Rev   GAGCGCACCATCTTCTTCAAG
Probe ACGACGGCAACTACA

RSV N (RSV-N-5)
Fwd   TGGCATGTTATTAATCACAGAAGATGCT
Rev   TTCTCTTCCTAACCTAGACATCGCATA
Probe AACCCAGTGAATTTATG
```

```
RSV P (RSV-P-5)
Fwd   TAAATTCCTAGAATCAATAAAGGGCAAATTCA
Rev   TCTTTGGTTACTTCTATATCTATTGAGTTGACAGA
Probe CATCACCCAAAGATCC

RSV NS1 (RSV-NS1-5)
Fwd   GAAGTAGCATTGTTAAAAATAACATGCTATACTGATAAA
Rev   TGCACAAACACAATGCCATTCAATT
Probe CTGCCTTAGCCAAAGCA

RSV F (RSV-F-5)
Fwd   CTTCTGGTCAAAACATCACTGAAGAATTT
Rev   ACCAACCAGTTCTCAGAGCAC
Probe CAGTTAGCAAAGGCTATCTTA

RSV G (RSV-G-5)
Fwd   CCTGGGACACTCTCAATCATTTATTATTCA
Rev   AGTGAAGTTGAGATTATCATTGCCAGAA
Probe TAGCACAAATCACATTATCC
```

Below are the reaction volumes used for a single qRT-PCR reaction.

2X TaqMan RT-PCR Mix	5 ul
40X TaqMan RT Enzyme	0.25 ul
20X TaqMan Gene Expression Assay	0.5 ul
Total RNA sample	0.05 ul
Water (H_2O)	4.2 ul

The reaction volume is 10 ul. The cycle parameters are:

Reverse transcription	48°C	15 minutes
Enzyme activation	95°C	10 minutes
Denaturation	95°C	15 seconds
Annealing/extension	60°C	1 minute (repeat 40 times)

Each RNA sample/TaqMan Gene Expression Assay is set up in triplicate and assay reactions were analyzed on the 7900HT Fast Real-Time PCR system (Applied Biosystems, Foster City, CA). The threshold cycle (Ct) for each reaction was determined by the SDS RQ Manager program (Applied Biosystems). The relative changes in each transcript level were calculated by the Comparative Ct method ($2^{-\Delta\Delta Ct}$ method) [207,208]

and reported relative to the RSV wt virus sample collected 18 hours after infection. On a per sample basis, each assay Ct value was first adjusted using their 18S measured Ct value. Then on an assay basis, each "18S adjusted" Ct value is made relative to the "18S adjusted" Ct value of RSV wt sample at 18 hours. These comparisons are converted to fold change using the Comparative Ct method for each sample/TaqMan assay tested. These values are averaged between three repeat experiments and shown in the bar graphs where samples are reordered based on growth level groupings evident in multicycle A549 virus growth data.

8.4 Results

The following graphs show specific viral mRNA expression levels during infection of A549 cells at a multiplicity of infection of 3 (MOI = 3). Samples in each graph are reordered by virus growth levels found in infected A549 cells. The left side contains three samples, RSV wt and wt-like viruses (NS1 29 and NS1 98-109) that exhibited wt or almost wt levels of virus growth. The middle set contains six more amino acid substitution virus samples as well as the ΔNS1 virus all having intermediate to low levels of growth. The right side contains viruses with the four single amino acid replacement for amino acid number 6 (Leucine) in NS1. These four viruses are grouped separately to show the impact a single amino acid change can have across different measures.

The expression of eGFP in the wt-like growth group of viruses all show increased mRNA from 18 to 24 hour period (2.3-2.4 fold increase). Interestingly, the two amino acid substitution viruses in this group seem to produce greater amounts of eGFP

transcripts at 18 hours relative to wt, but ultimately are equivalent by 24 hours (these dynamics would not be seen in multicycle growth curves). Members of the intermediate-low growth group produce eGFP message at 18 hours similar to RSV wt (range between 0.7 and 1.6 fold increase). By 24 hours this same group appears to have two different eGFP expression profiles. One set (NS1 06-09, NS1 58-66 and NS1 98-104st) has a more increased level of eGFP expression (1.5, 1.8 and 1.6 fold increase respectively) while the other set (NS1 47, NS1 11-13-15, NS1 54-57-60) has a more ΔNS1 like level of eGFP expression (1.1, 1.2, 1.3 and 0.9 fold increase respectively). The single amino acid substitution virus set reveal a similar dichotomy regarding eGFP transcript expression. All viruses in this group transcribe eGFP message about the same level through 18 hours. However, by 24 hours, only NS1 06D is unable to increase eGFP message up to RSV wt 24-hour levels (1.1 versus 2.3 fold respectively). The other single amino acid substitution viruses (NS1 06Q, NS1 06R and NS1 06W) are capable of increasing eGFP message similar to wt (2.0, 2.3, 1.6 and 2.3 fold respectively).

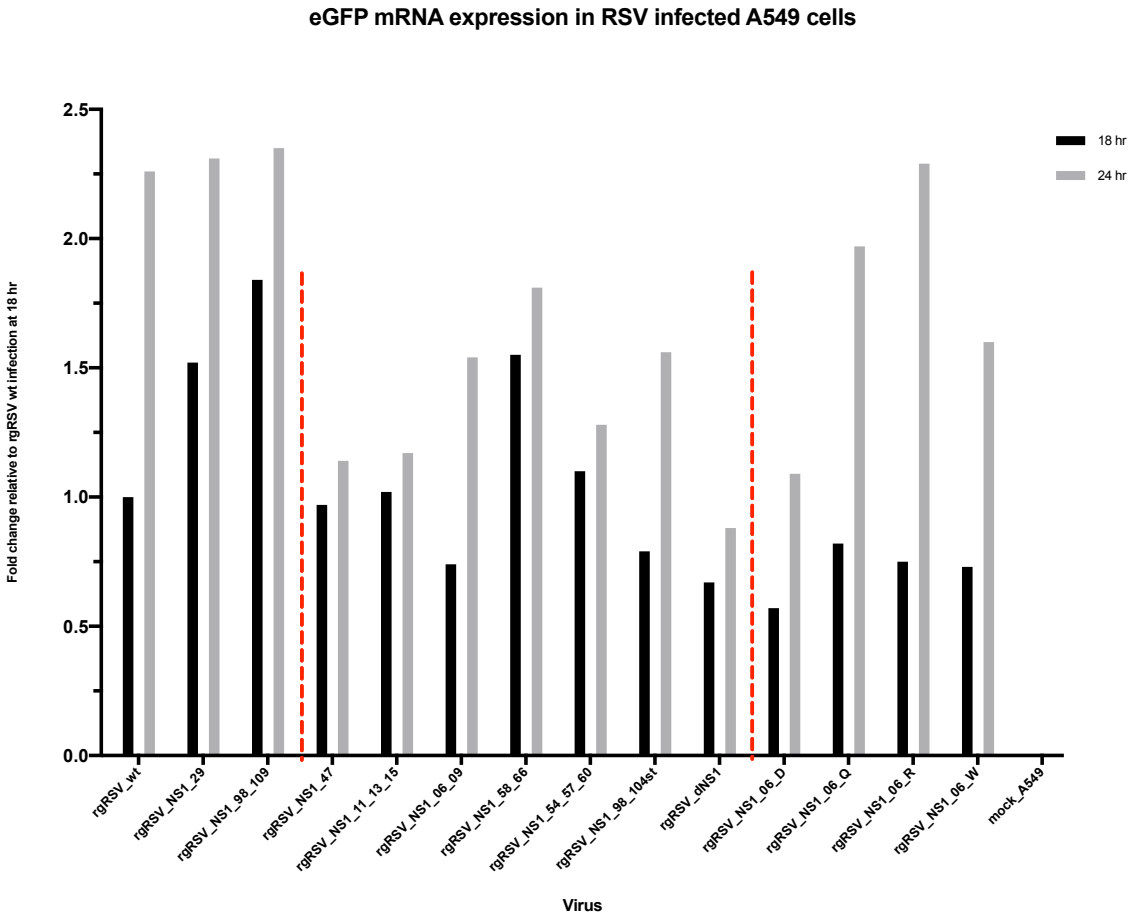


Figure 20 Viral eGFP gene expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Viral gene expression levels of eGFP were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a custom eGFP TaqMan assay. Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

RSV wt expresses N transcripts 1.8 fold at 24 hours over its level at 18 hours. Unlike eGFP, N mRNA doesn't universally increase among the samples from 18-24 hours. In the other wt-like growth group of viruses (NS1 29 and NS1 98-109), the detected level of N transcripts is 1) highest at 18 hours and 2) remained elevated at 24

hours (1.6 and 1.9 fold respectively; wt is 1.8 fold at 24 hours). The intermediate-low growth set of viruses again appear to have two sets of N expression profiles. One set of viruses (NS1 47, NS1 11-13-15, NS1 58-66 and NS1 54-57-60) all show RSV N mRNA levels decrease between 18-24 hours (1.1-0.9, 1.2-1.1, 1.9-1.5 and 1.3-1.0 fold respectively). The other N transcript expression profile set of this group of viruses (NS1 06-09, NS1 98-104st and ΔNS1) all show N transcription increasing between 18 and 24 hours (0.9-1.3, 1.0-1.2 and 0.8-1.0 fold increase respectively). The remaining group of single amino acid replacement viruses exhibit 18-24 hour RSV N mRNA expression profiles falling into either the intermediate-low growth group (NS1 06-D; 0.6-1.0 fold increase) or wt-like (NS1 06-Q; 0.9-1.9, NS1 06-R; 0.8-1.9 and NS1 06-W; 0.8-1.6 fold increase respectively).

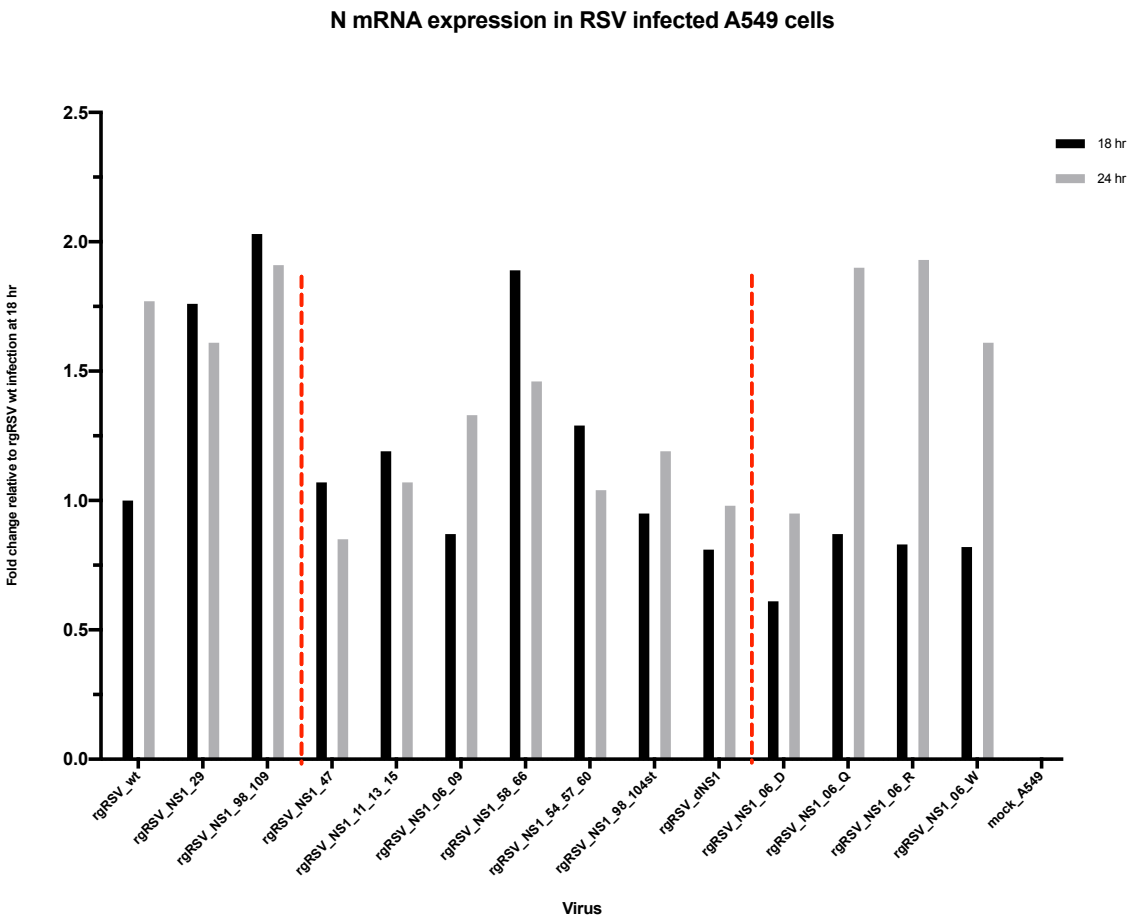


Figure 21 Viral RSV N gene expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Viral gene expression levels of RSV N were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a custom RSV N TaqMan assay. Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

RSV wt expresses P at about a 1.8 fold at 24 hours over its level at 18 hours. The other wt-like viruses (NS1 29 and NS1 98-109) have much greater levels of P mRNA transcription at 18 hours compared to wt (as seen previously with N and eGFP

expression). Interestingly the level of RSV P mRNA expression in the intermediate-low growth virus group appears to be associated with the two expression profile groupings similar to N mRNA transcription. One set of viruses (NS1 47, NS1 11-13-15 and NS1 58-66) all show RSV P mRNA levels decrease between 18-24 hours (1.1-0.9, 1.6-1.4 and 2.1-1.9 fold respectively). The other P transcript expression profile set (NS1 06-09, NS1 54-57-60, NS1 98-104st and ΔNS1) all show P transcription increasing between 18 and 24 hours (1.1-1.8, 1.5-1.9, 1.3-2.0 and 1.1-1.3 fold increase respectively). The remaining group of single amino acid replacement viruses exhibit 18-24 hour RSV P mRNA expression profiles falling into either the intermediate-low growth group (NS1 06-D; 0.6-1.1 fold increase) or wt-like (NS1 06-Q; 0.8-1.6, NS1 06-R; 0.9-2.1 and NS1 06-W; 0.7-1.4 fold increase respectively).

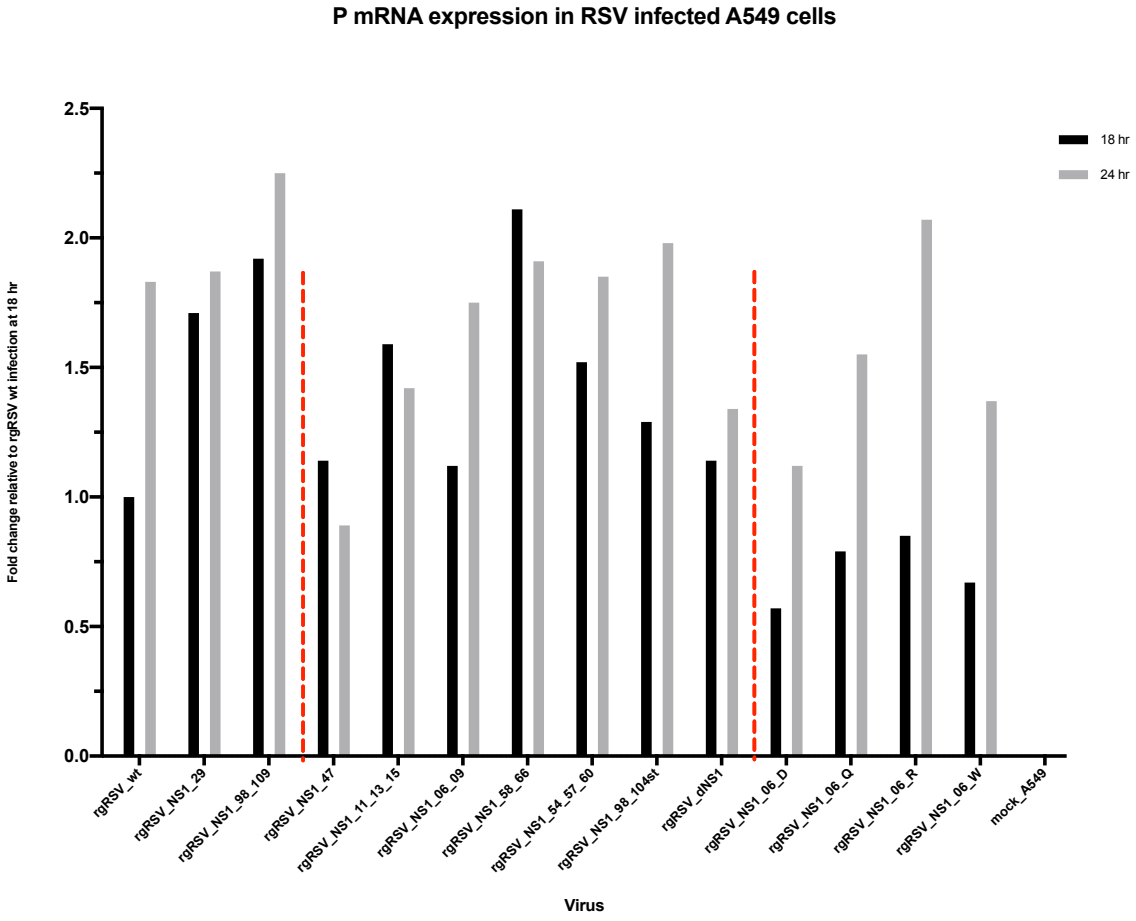


Figure 22 Viral RSV P gene expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Viral gene expression levels of RSV P were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a custom RSV P TaqMan assay. Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

For each virus, the NS1 protein sequence has been specifically modified, therefore presence and level of NS1 mRNA expression in each is paramount. The level of NS1 mRNA is 1.3 fold higher at 24 hours relative to 18 hours in the RSV wt sample.

A single member of the wt-like virus growth group (NS1 98-109) approaches or exceeds this level of expression (0.8-1.4 fold increase) during the same time period. The other virus, NS1 29 expresses NS1 mRNA at a lower but still similar to wt level (0.6-0.9 fold increase). The level of NS1 mRNA expression among the intermediate-low virus growth group seem to fall into two sets of mRNA expression. One set (NS1 47, NS1 54-57-60 and NS1 98-104st) all show reduced levels of NS1 mRNA expression at 18 and 24 hours (0.5-0.5, 0.5-0.5 and 0.5-0.7 fold increase respectively). Of the remaining intermediate-low virus growth members, NS1 11-13-15 expresses NS1 mRNA at RSV wt levels at 18 hours (1.1 fold) dropping slightly by 24 hours (0.8 fold). The other two viruses, NS1 06-09 and NS1 58-66, express NS1 mRNA at both 18 and 24 hours relative to RSV wt NS1 mRNA expression slightly lower (0.9-0.8 and 0.7-0.7 fold respectively). Levels of NS1 mRNA in the ΔNS1 virus is not detected as expected. Interestingly, the NS1 98-104st sample, which was accidentally created during the amino acid modification of NS1 98-109 sample, produces an mRNA with a protein translation stop sequence at codon position 105. This sample expression data indicates the NS1 mRNA is produced and detectable. This is in contrast to the ΔNS1 virus in which the complete NS1 gene sequence is missing and as expected produces no detectable NS1 mRNA. The remaining group of single amino acid replacement viruses exhibit 18-24 hour NS1 mRNA expression profiles very similar to RSV P, RSV N and eGFP, falling into either the wt-like growth group (NS1 06-Q; 0.7-1.3, NS1 06-R; 0.6-1.2 and NS1 06-W; 0.6-1.0 fold increase) or intermediate-low growth group (NS1 06-D; 0.5-0.8 fold increase).

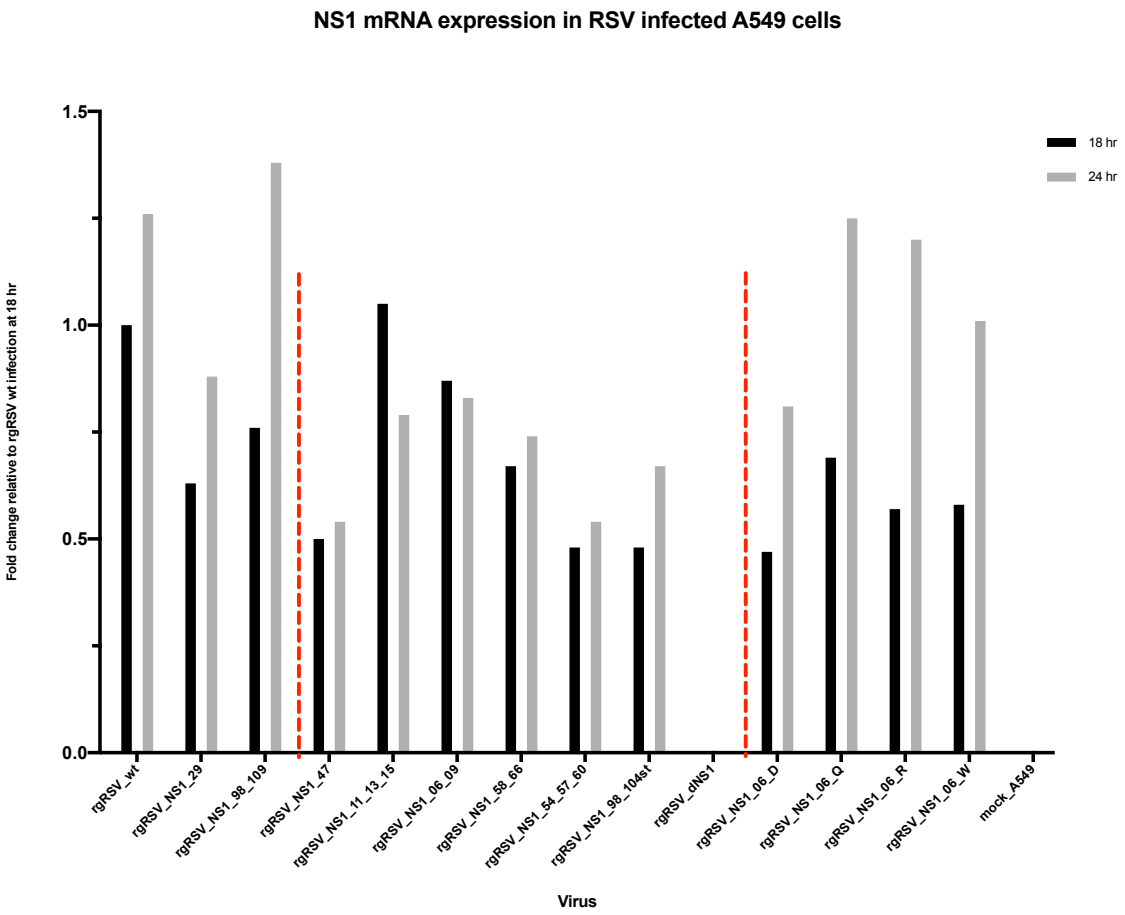


Figure 23 Viral RSV NS1 gene expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Viral gene expression levels of RSV P were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a custom RSV P TaqMan assay. Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

While RSV N and P transcripts code for essential building blocks of viable virus production, it was of interest to see whether levels of RSV glycoproteins, F and G mRNA related to levels of virus growth. Both proteins are produced during infection and are the primary antigenic targets based on levels of serum antibodies against each detected in

infected populations [40,209,210]. However, the relationship between functional NS1 and the production of these glycoproteins is not known. The level of F mRNA in RSV wt infection increased 2.2 fold by 24 hours. Two other members of the wt-like growth group (NS1 98-109 and NS1 29) express F mRNA similar to RSV wt (1.4-1.8 and 1.1-1.7 fold increase respectively) between 18 and 24 hours. The level of F mRNA in the intermediate-low growth virus group (NS1 47, NS1 11-13-15, NS1 06-09, NS1 58-66, NS1 54-57-60, NS1 98-104st and ΔNS1) fluctuated around the wt 18 hour mRNA level (0.8-0.7, 0.8-0.8, 0.5-1.1, 1.2-1.1, 0.8-1.1, 0.7-1.2 and 0.6-0.7 fold increase respectively). The remaining group of single amino acid replacement viruses exhibit 18-24 hour F mRNA expression profiles very similar to RSV N, P, NS1 and eGFP, falling into either the wt-like growth group (NS1 06-Q; 0.7-1.5, NS1 06-R; 0.7-1.7 and NS1 06-W; 0.6-1.2 fold increase respectively) or intermediate-low growth group (NS1 06-D; 0.5-0.9 fold increase).

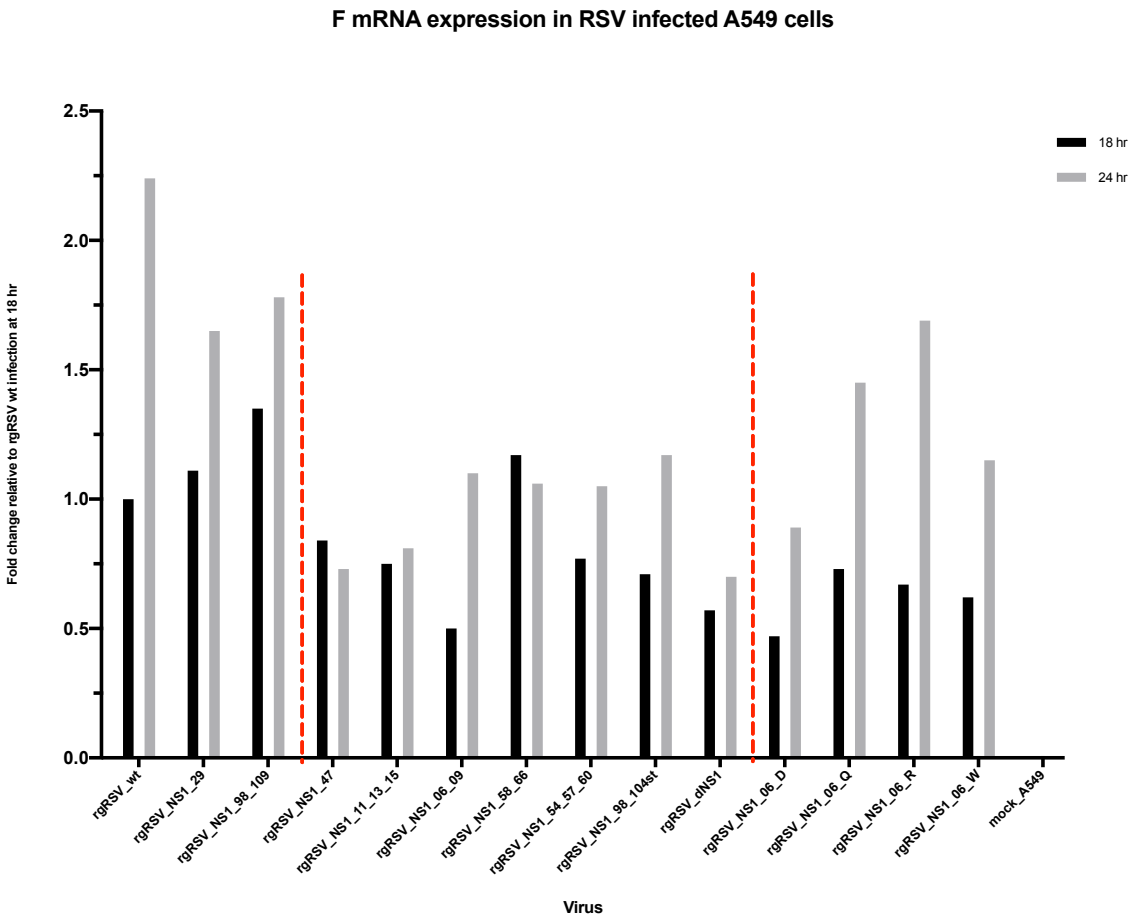


Figure 24 Viral RSV F gene expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Viral gene expression levels of RSV F were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a custom RSV F TaqMan assay. Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

The level of G mRNA in RSV wt infection increased 1.9 fold by 24 hours. The amounts of G mRNA made by each of the viruses is well correlated with F mRNA levels for the same virus infection at either 18 or 24 hour (R^2 = 0.9177 and 0.9387 respectively, data not shown). Two members of the wt-like growth group (NS1 98-109 and NS1 29)

have G mRNA expression profiles similar to eGFP, N and P mRNA of these same viruses between 18 and 24 hours. Notably the 18 hour levels of transcription are measurably higher than RSV wt (1.7 and 1.4 fold increased respectively). The level of G mRNA (like the level of F mRNA) in the intermediate-low growth virus group (NS1 47, NS1 11-13-15, NS1 06-09, NS1 58-66, NS1 54-57-60, NS1 98-104st and ΔNS1) fluctuated around the wt 18 hour mRNA level (0.7-0.7, 0.8-0.8, 0.6-1.0, 1.2-1.0, 0.9-0.9, 0.7-1.0 and 0.5-0.6 fold increase respectively). The remaining group of single amino acid replacement viruses exhibit 18-24 hour G mRNA expression profiles very similar to RSV F, N, P, NS1 and eGFP, falling into either the wt-like growth (NS1 06-Q; 0.6-1.3, NS1 06-R; 0.7-1.4 and NS1 06-W; 0.6-1.1 fold increase respectively) or intermediate-low growth group (NS1 06-D; 0.4-0.7 fold increase).

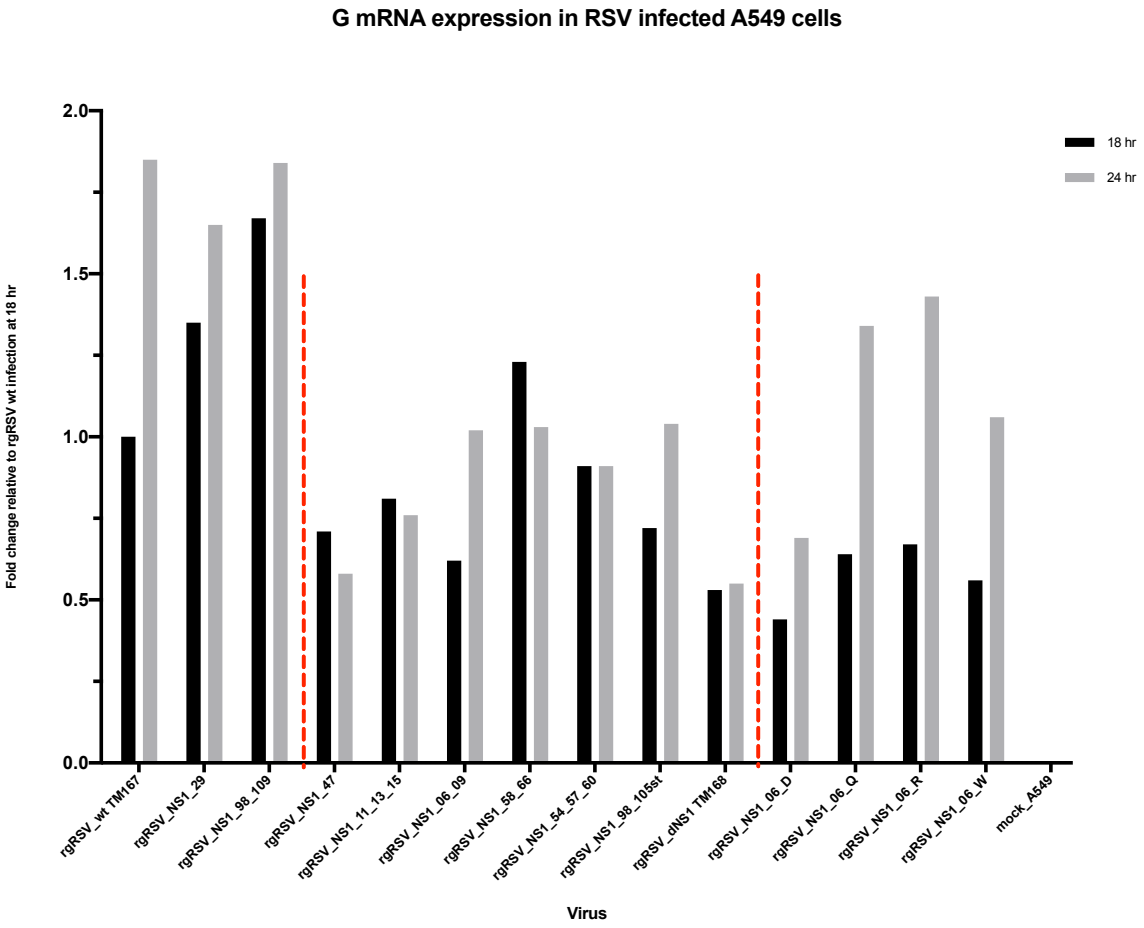


Figure 25 Viral RSV G gene expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Viral gene expression levels of RSV G were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a custom RSV G TaqMan assay. Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

8.5 Discussion

Any relationship between AUTOMUTE 2.0 predictions and subsequent mechanistic effects of NS1 regarding the range of virus growth among the different

substitutions is unknown. The suggestion(s) of NS1 protein function during an actual viral infection is derived from mixtures of experiments, some in which NS1 is abstracted from the other viral factors that exist during a viral infection [204–206] and with which NS1 protein function has evolved. The variable level of virus growth in an IFN competent cell line like A549, the known association of IFN with host mRNA transcription and NS1 suggested role as an IFN response antagonist may influence levels of viral mRNA expression [12,13]. It would follow that differences in viral mRNA expression could relate to differences in viable virus.

Unlike the multicycle growth experiments where the MOI = 0.1, here we are considering the primary effect of NS1 protein on transcription levels given that approximately all cells (MOI = 3) are infected at approximately the same time and that this effect should be manifested acutely in time (18-24 hours) unlike multicycle growth that takes days to develop. This data may also afford insight as to whether NS1 exhibits a dose dependency regarding its influence on RSV infection process. However, there is a "chicken . . . egg" type question. Are any NS1 substitutions primarily affecting viral mRNA transcription which lead to different viral growth levels or the other way around (NS1 substitutions primarily affect viral growth which in turn lead to different viral mRNA transcription levels)?

The level of eGFP transcription among the NS1 substituted viruses agreed well with the grouping of viruses by growth level. This result is not surprising since the level of virus growth detected in the multicycle growth experiments was measured as the number of eGFP expressing plaque forming units. However, this surrogate gene is not a

natural part of RSV viral gene repertoire and therefore may be less informative in assessing the impact NS1 may have on regulating transcription of authentic viral genes with which it has an evolutionary relationship with. RSV N and P gene transcription, providing a more appropriate evolutionary relationship context, are likely better indicators of any NS1 influence on viral gene transcription. N and P are both proteins essential to successful virus replication [124] (unlike eGFP) yet the general level of mRNA expression among eGFP, N and P show similar kinetics within each growth group suggesting that the expression of these transcripts are influenced by virus growth. Whether NS1 protein function is primarily affecting viral gene transcription or more generally related to viral infection is not clear. Equally important but unlikely related to viral transcription regulation is the level of NS1 transcription itself. Do any NS1 mRNA sequence changes (due to amino acid recoding) influence mRNA stability and therefore NS1 mRNA expression? The level of NS1 mRNA expression does seem related to groupings of virus growth, however it remains unclear whether the different NS1 mRNA levels are due to stability or just a reflection of virus replication (and related virus gene transcription).

Finally looking at the immunologically important viral glycoproteins F and G [211], again we see that the levels of these transcripts seem to correspond well to the groupings of virus growth. An interesting consideration for these products relates to the high MOI used in these experiments. Particularly regarding F (but also likely G), it is important to remember that virtually all cells are infected at a MOI = 3 and therefore the relevance of producing (and consequently detecting) increasing amounts of these mRNA

may not be accurately reflected in this saturating experiment design. Lastly as evident among the single amino acid substitution viruses, it is interesting to see that a single amino acid change can have considerable variation on both virus growth as well as mRNA transcription. These data imply that the general reduction in all viral gene transcripts in NS1 06-D compared to the other single amino acid replacements is likely related primarily to NS1 substitutions. Among these viruses, the difference in A549 infected virus growth is less pronounced than the differences in viral gene transcription. This suggests that at least among these viruses the impact of NS1 function (resulting from changes in its structure) led to consequences impacting mRNA transcription. The various NS1 amino acid substitutions are likely related to a change in the level of expressed viral mRNAs as well as differences in viable virus production and therefore lends support to the concept that AUTOMUTE 2.0 predictions are not random.

CHAPTER 9. VIRAL GENOME REPLICATION (STRAND SPECIFIC QRT-PCR)

9.1 Abstract

In addition to the influence NS1 seems to have on viral mRNA transcription, it has been reported that NS1 can affect levels of viral RNA genome synthesis [14]. Since NS1 amino acid substitutions seem to influence the level of viral mRNA synthesized, it is not unreasonable to examine whether a similar influence would be seen in the amount of viral genomic RNA synthesized. In this work, the same set of viruses with various amino acid substitutions in NS1 were evaluated for differences in the amount of genomic viral RNA detected using customized TaqMan assays that are specifically designed to detect viral genomic sequences not found in mRNA.

9.2 Introduction

The previous qRT-PCR data suggests the levels of expressed viral mRNA and virus growth are associated with amino acid substitutions in the NS1 protein sequence. Negative-sense single-stranded viruses such as RSV rely on a virally encoded RNA dependent RNA polymerase (RdRP) to synthesize viral mRNA transcripts from the negative-sense RNA genome [4]. However, upon some unknown mechanism, this same RdRP can ignore unique virus transcriptional gene start and gene stop signals in order to synthesize full length positive-sense complementary copies (referred to as positive-sense viral RNA genome) of the negative-sense viral RNA genome [49,127]. Lastly by yet another incompletely described mechanism, the same RdRP is able to begin using the newly synthesized full length positive-sense viral genomic RNA as a template to

synthesize new negative-sense viral genomic RNA required for viable virus assembly and subsequent infection [64,212]. Using this positive-sense version of the viral genome as a template, the RdRP will synthesize complementary versions starting at the 3' trailer end of positive-sense viral genomic RNA template generating negative sense viral RNA genomes in a 5' to 3' fashion [212]. The detection of negative-sense genomic sequences is reliant on strand (sense) specific reverse transcription (RT) priming and subsequent PCR only capable of detecting cDNA made during the strand specific RT reaction [187]. It has been reported that NS1 can affect the level of RdRP produced negative-sense viral genomic RNA [14]. Combining this methodology with custom designed TaqMan assays allows quantification of strand specific viral genomic RNA and this work will use the same set of NS1 amino acid substitution viruses to address whether any of the amino acid substitutions appear to associate with differential levels of viral genomic RNA.

9.3 Material and methods

Cell Line

A549 cells (adenocarcinoma human alveolar basal epithelial cell) (ATCC CCL185) were maintained in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 10% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). Infected A549 cells were in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 2% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies).

Virus Infections

Triplicate experiments of A549 cells in 12 well dishes were infected with virus. Samples were collected at 18- and 24-hours post infection. For these virus infections, virus stocks were diluted to a multiplicity of infection of 3 (MOI = 3) pfu/cell in OptiPRO SFM (Gibco-Life Technologies), 4 mM L-glutamine (Gibco-Life Technologies) and added to cells in appropriate volumes. Infections were left on cells for 1-2 hours with periodic rocking at 37°C; 5% CO_2. Subsequently the infectious volume is removed, monolayers washed with OptiPRO SFM, 4 mM L-glutamine media and refed with F-12K, 2% FBS, 2 mM L-glutamine. Infected cells were left at 37°C; 5% CO_2 until sample collection).

Thermocycle Reactions

All thermocycler reactions were done using a ProFlex PCR System (Applied Biosystems) as recommended by the manufacturer.

Strand Specific Reverse Transcriptase Reaction (RT)

Triplicate experiments of infected A549 cell cultures were collected, and total RNA was isolated with RNeasy Mini Kit (Qiagen) as recommended by the manufacturer, including on-column DNase digestion using a QIAcube sample preparation machine (Qiagen). The total RNA was eluted in 50 ul of EB solution and stored at -80°C until needed. To measure strand specific viral genome levels, a fixed volume of total RNA was combined with a strand specific reverse transcriptase (RT) primer (each RT primer is paired with a specific custom TaqMan Expression Assay) and used in a SuperScript III First-Strand Synthesis System (Invitrogen) reaction as recommended by the manufacturer

to generate complimentary DNA (cDNA). Below are the volumes used for a given RT reaction.

Negative-sense viral genome detection

Strand specific primer [2 uM]	1 ul
Total RNA sample	0.05 ul
dNTP [10 mM]	1 ul
Water (H_2O)	7.95 ul

18S detection

Random Hexamer primer [50 ng/ul]	1 ul
Total RNA sample	0.05 ul
dNTP [10 mM]	1 ul
Water (H_2O)	7.95 ul

The reaction volume is 10 ul. The cycle parameters are:

RNA denaturation	65°C	5 minutes
Cooling/Anneal	ice	1 minute

RT reaction

10X First Strand Buffer	2 ul
25 mM $MgCl_2$	4 ul
0.1 M DTT	2 ul
RNaseOUT [40 U/ul]	1 ul
SuperScript III [200 U/ul]	1 ul

This 10 ul volume is added to the previously heated/cooled RNA + primer mix. The cycle parameters are:

cDNA Synthesis	50°C	50 minutes
Deactivate RT enzyme	85°C	5 minutes

Add 1 ul of RNase H (to remove RNA templates) and incubate at 37°C for 20 minutes.

Remove from thermocycler and bring 20 ul reaction up to a final of 200 ul with H_2O.

Quantitative PCR (qPCR)

Strand specific viral genomic sequences are detected with custom designed TaqMan Expression Assays that work in conjunction with paired RT primers used in the strand specific cDNA reaction. The TaqMan Gene Expression Assay 18S rRNA (Hs99999901_s1) was used as a normalization control in conjunction with the custom TaqMan assays designed to detect three genome areas; RSV 49 trailer (nucleotides 15733-15856; exclusively in trailer sequence (15718-15872)), RSV M2-1 (nucleotides 8248-8308) and RSV 57 leader-NS1 (nucleotides 13-104; second half of leader sequence through the adjacent non-coding region of NS1 and into the first 6 nucleotides of NS1 open reading frame (ORF)). The custom designed TaqMan assays include:

[The sequence in **BOLD** font indicates the nucleotide sequence complement to the Fwd primer in each custom TaqMan assay that ensures only RT primed cDNA sequences are further amplified during the PCR process; the sequences in **RED** are unique to a given location in the genome sequence]

RSV 49 trailer (TM_RSV_49)
RT primer **CGGTCATGGTGGCGAATAA**TTTTTTAAATAACTTTTAGTGAACTAATCCTAAAG

Fwd CGGTCATGGTGGCGAATAA
Rev CAAAAACTAATATCTCGTAATTTAGTTAATACACATATAAACCAAT
Probe TTGGAGGAATAAATTTAAACCC

RSV M2-1 (RSV_M2_Gen)
RT primer **CGGTCATGGTGGCGAATAA**GGGCAAATATGTCACGAAGGA

Fwd CGGTCATGGTGGCGAATAA
Rev CCATTTAAGCAATGACCTCGAAT
Probe TCCTTGCAAATTTG

```
RSV 57 leader-NS1 (TM_RSV_51)
RT primer   CGGTCATGGTGGCGAATAAGCGTACAACAAACTTGCATAAACCA

Fwd         CGGTCATGGTGGCGAATAA
Rev         GCCCATCTCTAACCAAGGGAGTT
Probe       AAGTGGTACTTATCAAATTC
```

Below are the reaction volumes used for a given qPCR reaction.

2X Universal Master Mix (No UNG)	5 ul
20X TaqMan Expression Assay	0.5 ul
cDNA sample	2 ul
Water (H_2O)	2.5 ul

The reaction volume is 10 ul. The cycle parameters are:

Enzyme activation	95°C	10 minutes
Denaturation	95°C	15 seconds
Annealing/extension	60°C	1 minute (repeat 40 times)

Each cDNA sample/TaqMan Expression Assay is set up in triplicate and assay reactions were analyzed on the 7900HT Fast Real-Time PCR system (Applied Biosystems, Foster City, CA). The threshold cycle (Ct) for each reaction was determined by the SDS RQ Manager program (Applied Biosystems). The relative changes in each detected sequence level were calculated by the Comparative Ct method ($2^{-\Delta\Delta Ct}$ method) [207,208] and reported relative to the same detected sequence in the RSV wt virus sample collected 18 hours after infection. On a per sample basis, each assay Ct value is first adjusted with their 18S measured Ct value. Then on an assay basis, each "18S adjusted" Ct value is made relative to the "18S adjusted" Ct value of RSV wt sample at 18 hours. These comparisons are converted to fold change measures for each sample/TaqMan assay tested. These values are averaged between three repeat experiments and shown in the bar

graphs in which samples are reordered based on growth level groupings evident in multicycle A549 virus growth data.

9.4 Results

Beginning with the trailer sequence assay, RSV 49, the level of negative strand genome detected in RSV wt infected cells is virtually unchanged between 18-24 hour period (1.0-0.9 fold increase respectively). The other wt-like growth viruses (NS1 29 and NS1 98-109) show a greater increase in detectable genome between 18-24 hours (1.3-2.1 and 0.8-1.7 fold increase respectively) as compared to RSV wt infection. The intermediate-low growth virus group (NS1 47, NS1 11-13-15, NS1 06-09, NS1 58-66, NS1 54-57-60, NS1 98-104st and ΔNS1) as well as the single amino acid substitution group viruses (NS1 06D, Q and W) all show lower viral genome expression relative to RSV wt at 18 hours and achieve wt-like viral genome expression by 24 hours. Only two viruses have a different genome expression profile, NS1 11-13-15 and NS1 06R. Interestingly NS1 11-13-15 genome expression profile is very much like wt-like viruses (0.9-1.6 fold increase) between 18-24 hours. The other virus, NS1 06R appears to have a much higher level of genome expression at 18 hours than any other virus tested, and it remains elevated through 24 hours.

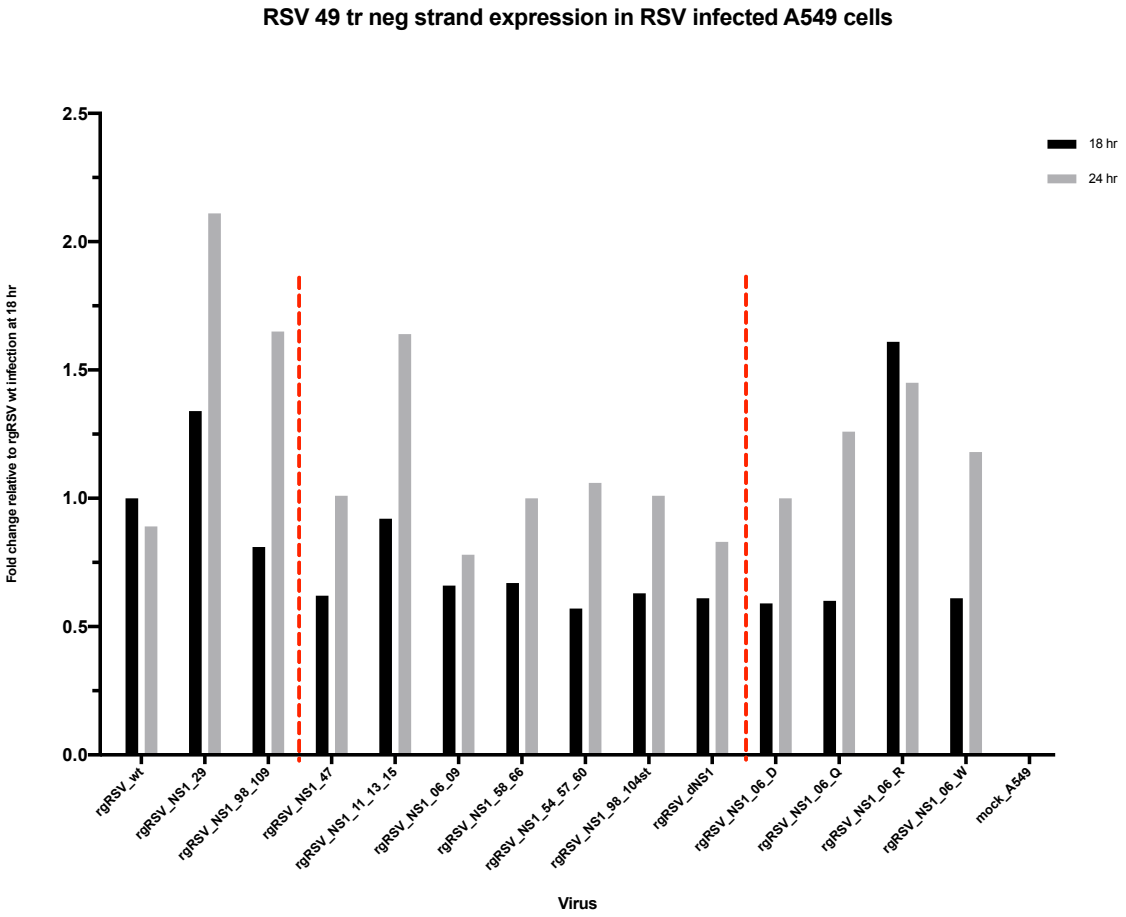


Figure 26 Viral RSV genome expression of strand-specific Trailer sequence in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Strand-specific viral genome expression levels of RSV Trailer sequence were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a custom reverse transcriptase (RT) complimentary DNA (cDNA) primer and custom RSV Trailer TaqMan assay (RSV 49 tr neg strand). Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

A previously reported assay, RSV M2-1 [187], was also used to measure negative strand genome copy levels. As with the previous described RSV trailer sequence assay, the level of negative strand genome detected by this internal position sequence assay in

RSV wt infected cells is virtually the same between 18-24 hour period. As seen with the RSV 49 trailer assay detected genome levels, the other wt-like growth viruses (NS1 29 and NS1 98-109) show a greater increase in detectable genome between 18-24 hours (2.0-3.5 and 1.2-2.3 fold increase respectively) than RSV wt infection. The intermediate-low growth virus group (NS1 47, NS1 11-13-15, NS1 06-09, NS1 58-66, NS1 54-57-60, NS1 98-104st and ΔNS1) all show viral genome expression approximately similar to RSV wt levels at 18 hours and increase by 24 hours. The genome expression profiles of the single amino acid substitution group viruses (NS1 06D, Q and W) as detected with the RSV M2-1 sequence assay all show lower viral genome expression relative to RSV wt at 18 hours and achieve wt-like viral genome expression by 24 hours. The remaining virus, NS1 06R appears to have a higher level of genome expression relative to RSV wt at 18 and 24 hours.

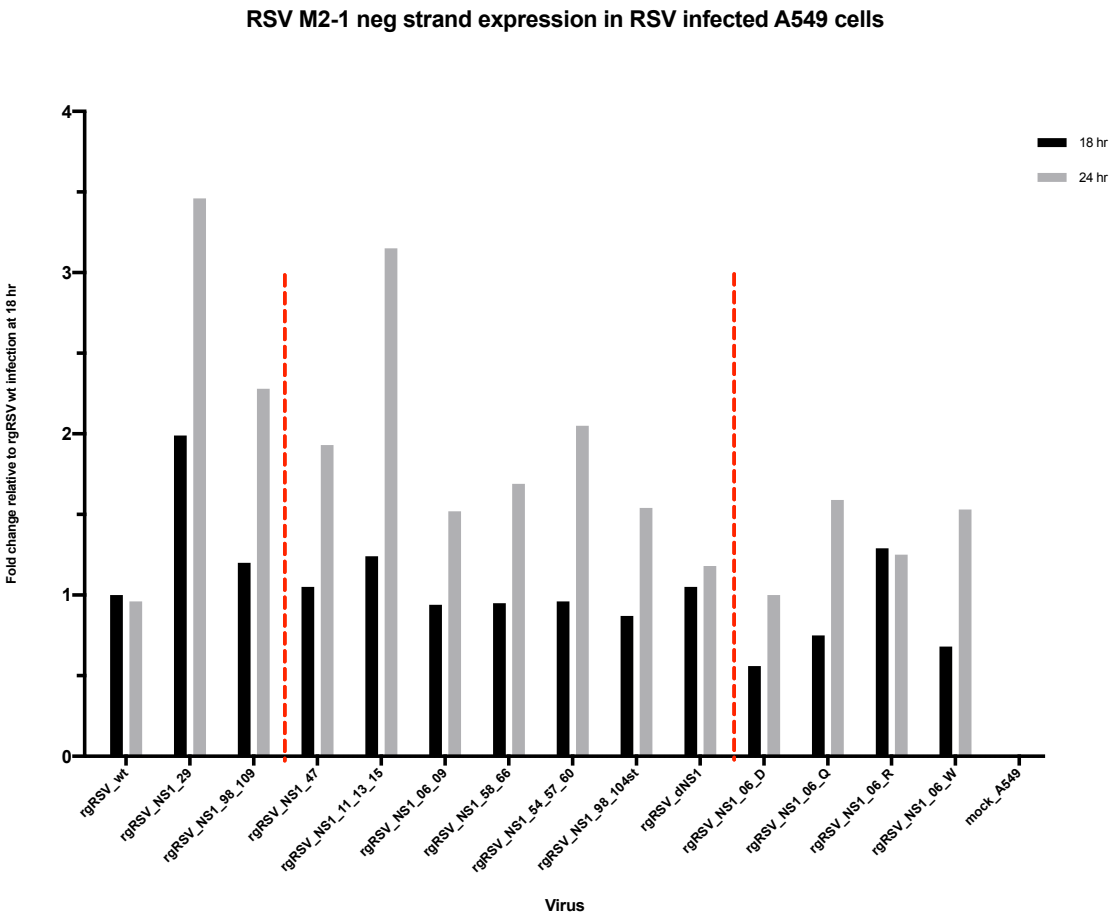


Figure 27 Viral RSV genome expression of strand-specific M2-1 sequence in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Strand-specific viral genome expression levels of RSV M2-1 sequence were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a custom reverse transcriptase (RT) complimentary DNA (cDNA) primer and custom RSV M2-1 TaqMan assay (RSV M2-1 neg strand). Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

The third negative strand genome assay (RSV 57 leader-NS1) was designed to detect viral genomic nucleotide sequence (nucleotide 13-104) spanning leader through the non-coding region of NS1 and into the first 6 nucleotides of NS1 open reading frame

(ORF). Similar to the RSV 49 trailer and RSV M2-1 assay, the level of negative strand genome detected by this assay in RSV wt is virtually the same between 18-24 hour period. The other wt-like viruses, NS1 29 and NS1 98-109, as well as all the single amino acid substituted viruses, NS1 06D, Q, R and W, have 18-24 hour expression level fold changes virtually matching the kinetics of detectable genome replication across all the negative sense assays designed (1.4-2.7, 1-1.5, 0.8-1.7, 1.0-1.7, 1.8-1.3 and 1.2-1.8 fold change respectively). Among the remaining viruses, RSV 57 (leader-NS1) detected genome levels were approximately similar to RSV wt levels at 18 hours (2.1, 1.4, 1.1, 1.4, 1.3, 1.1 and 0.8 fold increase respectively) with the exception of NS1 47, which had 2.1 fold higher level of genome expressed at 18 hours as compared to RSV wt. All members of this virus growth group had a substantial increase of expressed genome by 24 hours (6.0, 6.7, 3.0, 3.8, 4.6, 3.1 and 3.0 fold increase respectively).

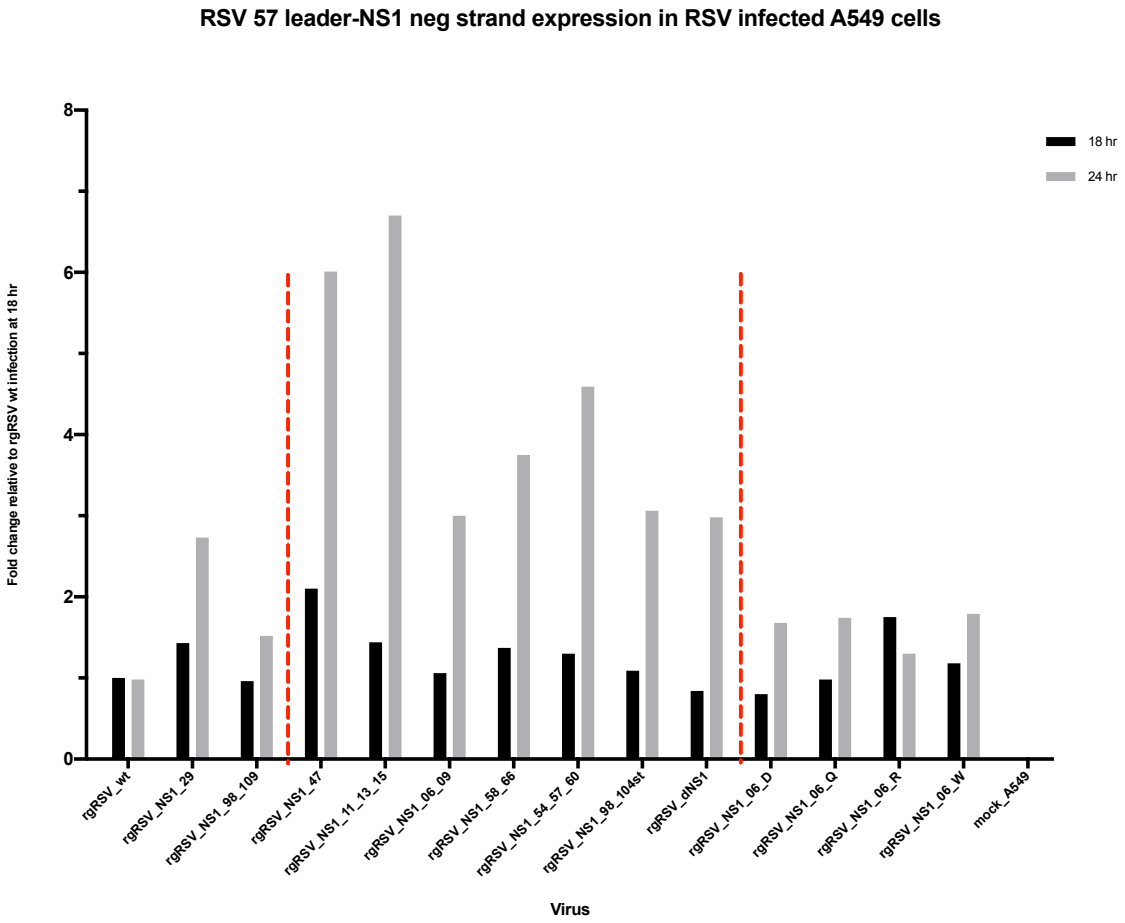


Figure 28 Viral RSV genome expression of strand-specific Leader sequence in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Strand-specific viral genome expression levels of RSV Leader sequence were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a custom reverse transcriptase (RT) complimentary DNA (cDNA) primer and custom RSV Leader TaqMan assay (RSV 57 leader-NS1 neg strand). Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

9.5 Discussion

Previous work introduced strand specific assays [187] to detect negative strand genome sequences. In this work we introduce additional assays to further evaluate

differential RSV genome copy levels. The levels of detected negative strand genome sequences relative to RSV wt sample at 18 hours for each virus is generally correlated among the assays however there does seem to be unique initial variation in levels of genome expression among viruses. For instance, the detected genome expression relativity between the viruses in the wt-like growth group are maintained across the negative strand detection assays and timepoints. The same appears true for the single amino acid substitution viruses providing a sense that the assays work and can detect the relative amounts of genome sequences regardless of sequence location. However, the detected genome expression relativity among the viruses in the intermediate-low growth group are maintained among the strand specific assays and timepoints with the exception of RSV 57 (leader-NS1). At 18 hours, the relative levels of expressed genome detected by RSV 57 (leader-NS1) assay among these samples is like relative genome levels found using the other strand specific assays. Surprisingly the levels detected by this same assay at 24 hours are substantially higher but relatively similar.

These results indicate that although there are relative differences in detected negative strand genome levels between the various viruses, it may simply be primarily related to variable levels of A549 virus growth. This is suggested where we see consistent directional relativity of genome expression between the viruses regardless of the detection assay used (the genome segment being detected). In addition, the detection of higher-than-expected levels of genome sequence containing leader into NS1 noncoding sequence in the intermediate-low growth virus sample group may indicate that NS1 function related to amino acid substitutions associates with a greater abundance of

leader-NS1 containing sequences despite the lower growth levels relative to RSV wt. More intriguing is whether this relates to the previously noted influence NS1 has had on RdRP viral gene transcription-genome replication [14]. Furthermore reports of RSV (and other negative sense single strand RNA viruses) ability to generate defective genomes may being inadvertently detected here [213–215]. As there is no mechanistic description regarding such a mechanism, this is speculation. What has been reported is the influence NS1 appears to have in reducing the level of detectable viral genomic RNA [14]. However, this prior work relied on bulk RNA detection with labeled probes, and the findings here, reliant on fundamentally different technology, do not necessarily refute or support such a notion. The data only suggests that the various NS1 amino acid substitutions may influence the 1) rate of full-length genome accumulation and 2) rate of select genome sequence accumulation, but no clear association can be made between virus growth levels and genome levels detected by these assays. These assays likely complicate the evaluation of AUTOMUTE 2.0 predictive utility by providing results lacking sufficient context.

CHAPTER 10. VIRAL PROTEIN EXPRESSION (WESTERN BLOT)

10.1 Abstract

Viral protein expression is after viral mRNA expression. Based on observed differences in virus growth and RNA expression, the expression level of select viral proteins, RSV NS1, RSV N and RSV F, were measured [216–218]. The expression levels of these protein were used to determine whether previously measured RNA expression was reflected in protein expression. The ability to detect RSV NS1 is only made by the recent development of numerous monoclonal antibodies against RSV NS1 protein. In addition, the measurements of RSV N and RSV F protein is used to assess impacts on fundamental aspects of RSV virion structure (N) as well the availability of a major immunogenic glycoprotein (F).

10.2 Introduction

The classic molecular cell biology dogma relies on a DNA to RNA to protein (function) schema. After measuring the differential effects of amino acid substitutions on expressed mRNA, it would follow to measure whether the expression level of various viral proteins was similarly affected. In this work protein expression is measured using the western blot process, which detects sequence epitope aspects of a protein (the linear or conformational region where a particular antibody will bind to). Conceptionally, the translation of RNA to protein is straight forward, but the level of stoichiometry involved may not be. Particularly when considering viral protein expression dynamic during viral infection as compared to eukaryotic cell protein expression, these conventions need

reevaluation each time. In this work we looked to measure whether viral protein expression variation may coincide with NS1 amino acid substitution(s).

10.3 Material and methods

Cell Line

A549 cells (adenocarcinoma human alveolar basal epithelial cell) (ATCC) were maintained in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 10% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). A549 cell infections were in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 2% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). A549 cells provide a relevant experimental system to examine virus growth in a fully IFN competent human lung epithelial host cell system.

Virus Infections

Triplicate experiments of A549 cells in 12 well dishes were infected with virus and samples collected at 18 and 24 hours post infection. For these virus infections, virus stocks were diluted to a multiplicity of infection of 3 (MOI = 3) pfu/cell in OptiPRO SFM (Gibco-Life Technologies), 4 mM L-glutamine (Gibco-Life Technologies) and added to cells in appropriate volumes. Infections were left on cells for 1-2 hours with periodic rocking at 37°C; 5% CO_2. Subsequently the infectious volume is removed, monolayers washed with OptiPRO SFM, 4 mM L-glutamine media and refed with F-12K, 2% FBS, 2 mM L-glutamine. Infected cells were left at 37°C; 5% CO_2 until sample collection.

Western blot analysis for the quantification of RSV viral proteins

Experimental samples were lysed with 100 µl 1x LDS sample buffer (Life Technologies) supplemented with cOmplete Protease Inhibitor Cocktail (Roche) tablet. Aliquots (8-10 µL) of lysate were denatured, reduced, and electrophoresed on 4-12% Bis-Tris SDS gels (Life Technologies) in 1X NuPAGE MOPS SDS Running Buffer (Life Technologies). Proteins were transferred onto PVDF membranes via the iBlot2 transfer system (Life Technologies). The following primary antibodies were used in the Western blot detection process. RSV NS1 was detected with an anti-NS1 mouse monoclonal antibody (mAb 1E-5-1, HHS Reference E-167-2018-0), RSV N protein was detected with an anti-N mouse mAb (Abcam, Cambridge, MA; ab94806) and RSV F protein was detected with an anti-mouse mAb (Abcam, Cambridge, MA; ab43812). Samples transferred to individual PVDF membranes and incubated separately with these primary antibodies overnight at 16°C, rocking slowly. These primary antibodies were detected using an infrared dye-conjugated goat anti-mouse immunoglobulin 800CW (Li-Cor, Lincoln, NE). Simultaneously on each PVDF membrane, GAPDH protein, used as a protein sample loading control, was detected with an anti-rabbit polyclonal antibody (pAb) (Proteintech; 10494-1-AP). This was secondarily detected with an infrared dye-conjugated goat anti-rabbit immunoglobulin 680RD (Li-Cor, Lincoln, NE). Western blot images were acquired on the Odyssey infrared scanner (Li-Cor) and analyzed with Image Studio Software (Version 5.2.5, Li-Cor). All protein band intensity values were normalized to GAPDH and reported relative to protein levels found in the RSV wt infected sample at 18 hours. These values are averaged between three repeat experiments

and shown in the bar graphs in which samples are reordered based on growth level groupings evident in multicycle A549 virus growth data.

10.4 Results

The level of NS1 protein expressed in the RSV wt sample decreases 0.6 fold from 18-24 hours. All subsequent discussions of protein fold changes will always be considered in relation to the GAPDH scaled level of expression of that protein in the RSV wt sample at 18 hours while the graphs will show the GAPDH scaled measurement. The NS1 protein expression level change in the wt-like growth group, NS1 29 and NS1 98-109 are 0.9-0.7 and 1.4-1.0 fold increased respectively during the 18 and 24 hour time points. The level of NS1 protein expression in the intermediate-low growth group is considerably reduced or absent except for NS1 47. Group members, NS1 11-13-15, NS1 58-66 and NS1 54-57-60 do express NS1, but at low levels between 18 and 24 hours (0.2-0.2, 0.3-0.3 and 0.3-0.4 fold change respectively). NS1 06-09 expresses even less NS1 protein during this time frame (0.04-0.04 fold change) while NS1 98-104st and ΔNS1 are negative for viral NS1 protein expression. Curiously, NS1 47 however is more wt-like with 18-24 hour NS1 expression level fold increase of 0.8-0.6. The remaining group of single amino acid replacement viruses express NS1 in one of three manners. NS1 06Q is similar to RSV wt with NS1 protein expression fold changes of 0.9-0.7 over the time period. NS1 06W and NS1 06R appear similar with lower NS1 protein expression ranging from 0.7-0.5 and 0.5-0.4 respectively. The most significant expression reduction is seen in the NS1 06D sample where NS1 protein expression is about one tenth the level of NS1 expression in RSV wt at both time points (0.1-0.1 fold change).

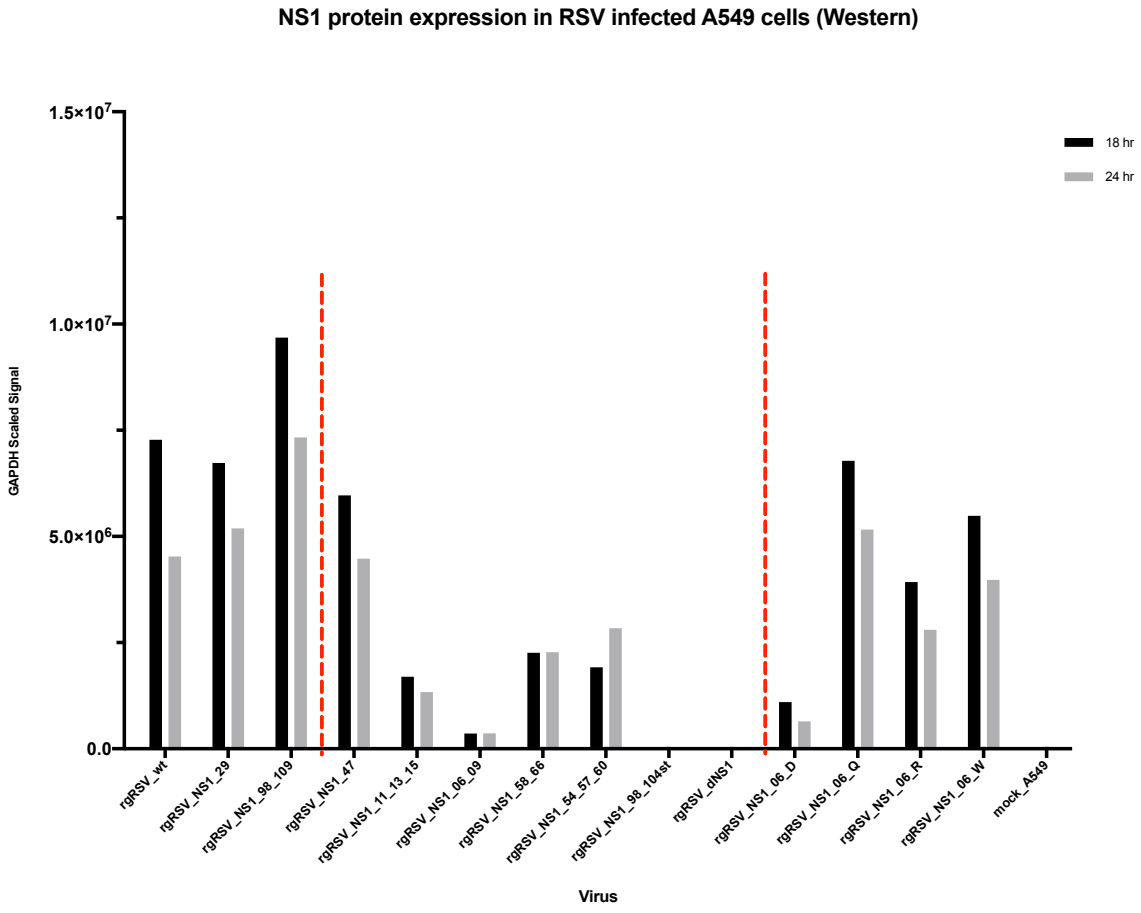


Figure 29 Viral RSV NS1 protein expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Viral protein expression levels of RSV NS1 protein were measured in total infected cell lysate samples collected at 18 and 24 hours post infection (hpi) using a newly developed mouse monoclonal antibody against RSV NS1 (1E-5-1, HHS Reference E-167-2018-0) as a primary antibody and an infrared dye-conjugated goat anti-mouse immunoglobulin 800CW (Li-Cor, Lincoln, NE) secondary antibody. GAPDH protein expression was detected using a rabbit polyclonal primary antibody 10494-1-AP (Proteintech) and an infrared dye-conjugated goat anti-rabbit immunoglobulin 680RD (Li-Cor, Lincoln, NE) secondary antibody. Each sample protein expression value was normalized to GAPDH, then calibrated to that protein expression level found in the RSV wt 18 hpi sample and reported as a GAPDH Scaled Signal. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

The level of N protein expression in RSV wt increased 1.6 fold by 24 hours. The other wt-like growth group members, NS1 29 and NS1 98-109, have fold changes in N expression like RSV wt of 1.1-1.8 and 1.0-1.4 respectively. The level of N protein expression is mostly lower in the intermediate-low growth group (NS1 47, NS1 11-13-15, NS1 06-09, NS1 58-66, NS1 54-57-60, NS1 98-104st and ΔNS1) all express N protein at levels lower than RSV wt (0.6-1.0, 0.7-0.9, 0.5-0.7, 0.6-0.9, 0.5-0.9, 0.4-0.7 and 0.9-1.0 fold increase respectively). The remaining four single amino acid replacement viruses all express N in a very similar pattern with fold changes ranging from 0.5-0.7 at 18 hours up to 0.7-1.1 by 24 hours.

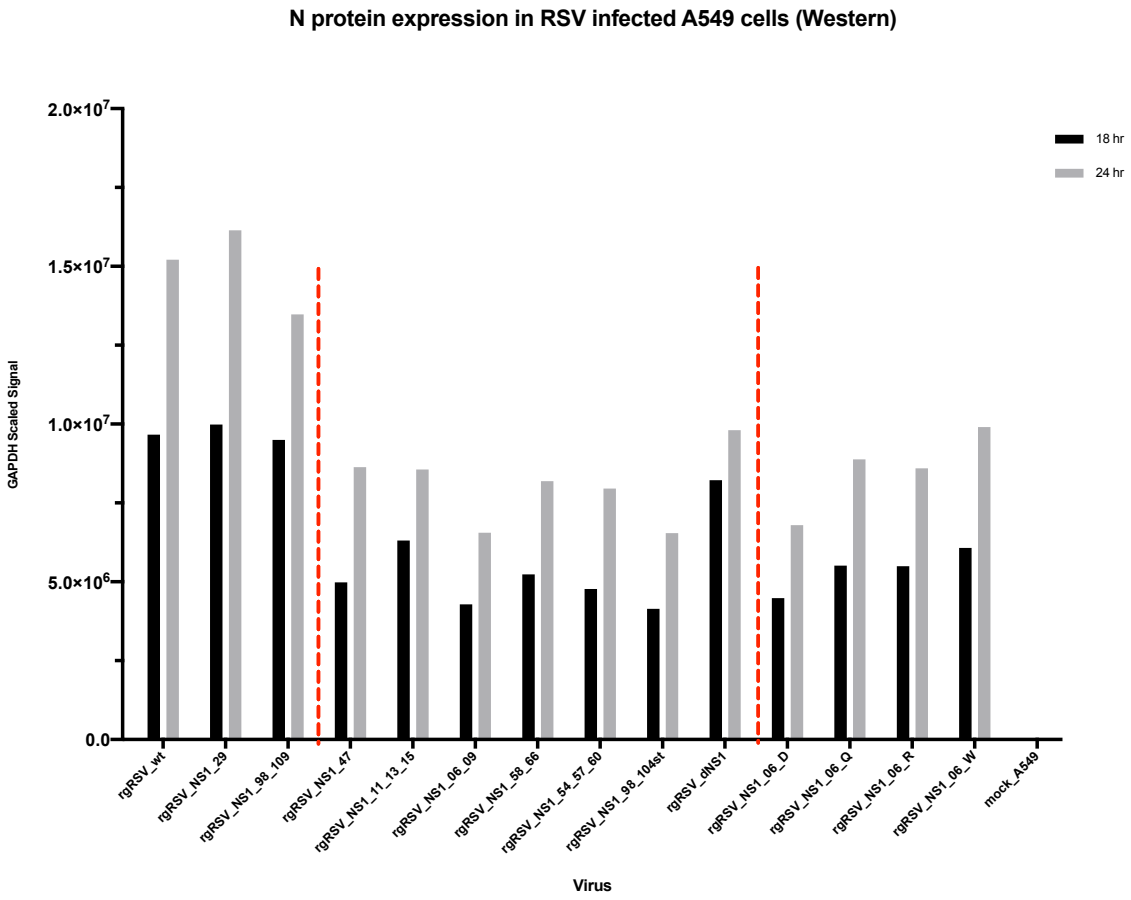


Figure 30 Viral RSV N protein expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Viral protein expression levels of RSV N protein were measured in total infected cell lysate samples collected at 18 and 24 hours post infection (hpi) using a mouse monoclonal antibody against RSV N (ab94806 Abcam) as a primary antibody and an infrared dye-conjugated goat anti-mouse immunoglobulin 800CW (Li-Cor, Lincoln, NE) secondary antibody. GAPDH protein expression was detected using a rabbit polyclonal primary antibody 10494-1-AP (Proteintech) and an infrared dye-conjugated goat anti-rabbit immunoglobulin 680RD (Li-Cor, Lincoln, NE) secondary antibody. Each sample protein expression value was normalized to GAPDH, then calibrated to that protein expression level found in the RSV wt 18 hpi sample and reported as a GAPDH Scaled Signal. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

The level of F protein expression in RSV wt increased 1.8 fold by 24 hours. Like RSV N protein expression, the wt-like growth group members NS1 29 and NS1 98-109

have fold changes in F expression similar to RSV wt of 0.9-1.6 and 0.7-1.3 respectively. The level of F protein expression is mostly lower in the intermediate-low growth group. NS1 11-13-15 and NS1 58-66 have F protein expression that is 16-30% of RSV wt F protein expression while NS1 47, NS1 06-09, NS1 54-57-60 and NS1 98-104st express F protein at 9-21% of the level of RSV wt F protein expression. The ΔNS1 infection expresses F protein at about a 50% of RSV wt expression levels of F protein. The remaining four single amino acid replacement viruses all express reduced levels of F. NS1 06Q, NS1 06R and NS1 06W range from a low of 23-28% (18 hours) up to a range of 55-65% (24 hours) compared to RSV wt. NS1 06D expresses F protein the least ranging from 14% at 18 hours up to only 28% by 24 hours of RSV wt F protein expression.

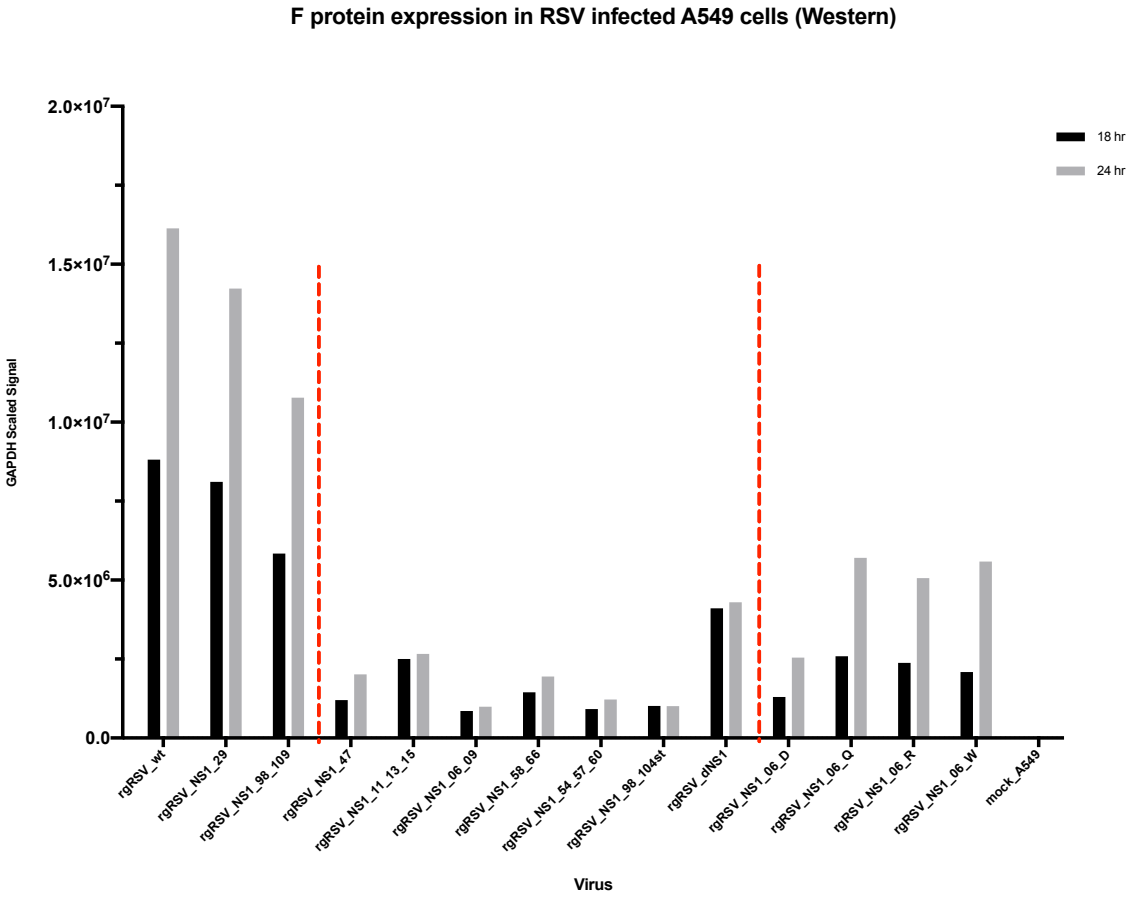


Figure 31 Viral RSV F protein expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Viral protein expression levels of RSV F protein were measured in total infected cell lysate samples collected at 18 and 24 hours post infection (hpi) using a mouse monoclonal antibody against RSV F (ab43812 Abcam) as a primary antibody and an infrared dye-conjugated goat anti-mouse immunoglobulin 800CW (Li-Cor, Lincoln, NE) secondary antibody. GAPDH protein expression was detected using a rabbit polyclonal primary antibody 10494-1-AP (Proteintech) and an infrared dye-conjugated goat anti-rabbit immunoglobulin 680RD (Li-Cor, Lincoln, NE) secondary antibody. Each sample protein expression value was normalized to GAPDH, then calibrated to that protein expression level found in the RSV wt 18 hpi sample and reported as a GAPDH Scaled Signal. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

10.5 Discussion

Several interesting findings are seen in the viral protein expression data. To begin, the amount of RSV NS1 which has so far never been examined natively with a monoclonal antibody and during an actual virus infection, reveals interesting kinetics. Overall, the level of NS1 protein expressed appears maximal at 18 hours and lessen by 24 hours. This contrasts with the previously measured mRNA expression levels seen for the same viruses that were increasing over the same time frame. With respect to the wt-like virus set, the level of RSV NS1 is quite like RSV wt infection levels. This finding helps support the notion that the predictions made for the amino acid substitution of NS1 29 and NS1 98-109 have an **Increased** prediction associated with each. Therefore, the level of protein expression not being reduced, and the wt-like virus growth levels suggest that these amino acid changes at least are not contrary to the **Increased** prediction made by AUTOMUTE 2.0. It may also suggest that these amino acid substitutions aside from not negatively impacting NS1 protein function also do not negatively impact NS1 protein stability or expression. The next set of intermediate-low growth viruses are equally interesting. Here the level of NS1 protein expression is quite varied. Unlike the wt-like virus set, the amino acid substitution predictions are mostly labeled **Decreased** rather than **Increased**. Following this generality in prediction, the expression level of NS1 in this set of viruses is considerably lower in this group. However, NS1 47 is predicted as **Decreased** yet the amount of expressed NS1 is quite similar to wt NS1 expression. In addition, NS1 58-66 (like NS1 98-109) has a mixed **Decreased - Increased** prediction, however the expression level of NS1 here is certainly lower than RSV wt NS1 protein

expression. One explanation for these contradictions may relate to the physical protein structure each amino acid is located within. For instance, NS1 98-109, amino acid 98 (associated with a Helix based region) is predicted as **Increased** and 109 (associated with a Coil based region) is predicted as **Decreased** and NS1 58-66, amino acid 58 (associated with a Sheet based region) is predicted as **Decreased** and 66 (associated with a Coil based region) is predicted as **Increased**. It seems that measured outcomes for the stability predictions related to non-Coil protein structure regions (Sheet or Helix structures) seem to have better agreement than outcomes measured for stability predictions related to Coil protein structures. Other members of the intermediate-low virus growth group (NS1 06-09, NS1 11-13-15 and NS1 54-57-60) all are predicted overall as **Decreased** and in fact have lower levels of NS1 protein expression, lower virus growth compared to wt virus and stability predictions of amino acids not found in Coil structures. Of consideration is whether the stability predictions refer to the intrinsic protein function assuming the protein is stable and present or is the prediction more generalizable and includes "lack of stability" effects in a **Decreased** prediction. These qualitative differences while interesting are beyond the scope of work presented here. Looking at the remaining set of four single amino acid substitution viruses, this question regarding stable but nonfunctioning versus unstable therefore nonfunctioning arises. All four amino acid substitutions (amino acid 6, Leucine) are predicted as **Decreased**, grow less well than RSV wt but range in NS1 protein expression as compared to wt virus. Here these viruses with a **Decreased** AUTOMUTE 2.0 prediction had approximately similar virus growth reductions yet extremely varied NS1 expression levels ranging from

wt-like (NS1 06Q) down to almost undetectable (NS1 06D). This set of mutations and the subsequent phenotypic outcomes may exemplify the previous mentioned possibilities: stable structure, complete function – NS1 06Q; slightly unstable structure, slightly reduced function – NS1 06R and W; and unstable structure, reduced function – NS1 06D.

It is equally important to see whether the variation in NS1 expression relates to the expression of other viral proteins. RSV N protein is essential for productive virus replication and infection and is the first structural protein after NS1 (and NS2) to be expressed [124]. The level of RSV N protein expression within each virus growth group is extremely consistent and reflective of the level of associated virus growth for that group. All viruses express and increase RSV N protein levels between 18 and 24 hours. The difference is the amount of RSV N protein produced and here the wt-like group is highest while the remaining intermediate-low group and four single amino acid replacement group is lower. The single understandable exception is ΔNS1. While it would appear to be an appropriate control virus for defective NS1 function, in fact it is a wholly different virus in that it makes neither NS1 mRNA or NS1 protein and is a full gene shorter in length as compared to all other viruses being tested. For these reasons the growth and other functional measures seen in this virus may not be comparatively relevant regarding NS1 functional defects possible in any of the amino acid substitution mutants. For example, ΔNS1 is not showing a reduction in the level of RSV N as might be expected relating to the dramatic reduction in A549 virus growth. This may reflect a phenomenon that is only evident when NS1 is completely removed and have nothing to

do with a functional defect in NS1, but rather be a consequence unique to deletion virus performance.

Lastly, we evaluated the level of RSV F protein expression and see the characteristics of expression levels are quite like the levels seen with RSV N. The amount of RSV F protein produced in the wt-like group is highest while the intermediate-low group is lowest. Interesting is the variation among the four single amino acid substitution viruses. As was seen with RSV NS1 or N protein expression, NS1 06Q, R and W viruses express RSV F protein expression lower than wt but higher than the intermediate-low growth virus group. Only NS1 06D has an F protein expression profile like the intermediate-low growth virus group F protein expression level. Lastly ΔNS1 seems to continue with unexpectedly higher levels of F protein expression inconsistent with its reduced growth level. It is likely only the variation of measured NS1 protein expression is directly related to the amino acid substitutions and subsequent protein stability. It is unlikely this data supports the concept that variations in the expression of the other viral proteins is primarily related to any changes in NS1 function(s). The variation in both RSV N and RSV F are more likely explained as consequences related to virus growth levels (that likely are related but not directly, to changes in NS1 functions due to amino acid substitutions). Therefore, amino acid substitutions in NS1 generally do not appear to contradict the AUTOMUTE 2.0 predictions.

CHAPTER 11. HOST CELL MRNA EXPRESSION (QRT-PCR)

11.1 Abstract

RSV NS1 is suggested to antagonize virus infected cell innate immune responses, particularly the interferon stimulated gene (ISG) response system [11–13,219]. Absent functional NS1, the interferon cytokines are available to influence both infected and adjacent host cells including host cell chemokines. Measuring the levels of host cell mRNA can help qualify whether any amino acid substitutions in NS1 affect host cell mRNA expression.

11.2 Introduction

Important cellular effectors regarding the ISG response system include Type I and Type III interferons (IFN), specifically interferon beta 1 (IFN-β – Type I) and interferon lambda 1 (IFN-λ1 – Type III) [4] as well as proinflammatory chemokines CCL5 also known as RANTES (regulated on activation, normal T cell expressed and secreted) and TNFα (tumor necrosis factor alpha) [220–222]. These effectors have a wide range of influence on cell viability but notably on influencing the host response to viral infections such as RSV. It has been reported that infection with RSV leads to the activation of proinflammatory cytokines and chemokines that may contribute to disease in RSV infected infants [223,224]. It has been shown that RSV lacking NS1 surprisingly induced lower levels of these proteins as compared to virus with NS1 [12]. Conversely, most viruses, particularly RNA viruses encode proteins that inhibit innate host defense response(s) [225,226]. The NS1 protein of RSV is known to influence virus replication

as well as antagonize the induction and signaling of Type I IFN [11–14,204,219]. Previous work indicated that removal of NS1 as well as specific mutations in NS1 are capable of demonstrating these outcomes [73,227,228]. The work here attempts to show that computationally predicted amino acid substitutions in NS1 can be evaluated in terms of the consequential influence viruses with these mutations have on host gene expression.

11.3 Material and methods

Cell Line

A549 cells (adenocarcinoma human alveolar basal epithelial cells) (ATCC) were maintained in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 10% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). A549 cell infections were in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 2% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). A549 cells provide a relevant experimental system to examine virus growth in a fully IFN competent human lung epithelial host cell system.

Virus Infections

Triplicate experiments of A549 cells in 12 well dishes were infected with virus and samples collected at 18 and 24 hours post infection. For these virus infections, virus stocks were diluted to a multiplicity of infection of 3 (MOI = 3) pfu/cell in OptiPRO SFM (Gibco-Life Technologies), 4 mM L-glutamine (Gibco-Life Technologies) and added to cells in appropriate volumes. Infections were left on cells for 1-2 hours with periodic rocking at 37°C; 5% CO_2. Subsequently the infectious volume is removed, monolayers washed with OptiPRO SFM, 4 mM L-glutamine media and refed with F-

12K, 2% FBS, 2 mM L-glutamine. Infected cells were left at 37°C; 5% CO_2 until sample collection.

Quantitative RT-PCR

Triplicate experiments of infected A549 cell cultures were collected, and total RNA was isolated with RNeasy Mini Kit (Qiagen) as recommended by the manufacturer, including on-column DNase digestion using a QIAcube sample preparation machine (Qiagen). The total RNA was eluted in 50 ul of EB solution and stored at -80°C until needed. To measure gene transcript levels, a fixed volume of total RNA was combined with a specific TaqMan Gene Expression Assay using the TaqMan RNA-to-Ct 1-Step Kit (Life Technologies) as recommended by the manufacturer. The TaqMan gene expression assay 18S rRNA (Hs99999901_s1) was used as a normalization control in conjunction with the following TaqMan assays IFNB1 (Hs01077958_s1), IFNL1 (Hs00601677_g1), CCL5 (Hs00174575_m1) and TNF (Hs00174128_m1).

Below are the reaction volumes used for a given qRT-PCR reaction.

2X TaqMan RT-PCR Mix	5 ul
40X TaqMan RT Enzyme	0.25 ul
20X TaqMan Gene Expression Assay	0.5 ul
Total RNA sample	0.05 ul
Water (H_2O)	4.2 ul

The reaction volume is 10 ul. The cycle parameters are:

Reverse transcription	48°C	15 minutes
Enzyme activation	95°C	10 minutes
Denaturation	95°C	15 seconds
Annealing/extension	60°C	1 minute (repeat 40 times)

Each RNA sample/TaqMan Gene Expression Assay is set up in triplicate and assay reactions were analyzed on the 7900HT Fast Real-Time PCR system (Applied Biosystems, Foster City, CA). The threshold cycle (Ct) for each reaction was determined by the SDS RQ Manager program (Applied Biosystems). The relative changes in each transcript level were calculated by the Comparative Ct method ($2^{-\Delta\Delta Ct}$ method) [207,208] and reported relative to an RSV wt virus sample collected 18 hours after infection. On a per sample basis, each assay Ct value is first adjusted with their 18S measured Ct value. Then on an assay basis, each "18S adjusted" Ct value is made relative to the "18S adjusted" Ct value of RSV wt sample at 18 hours. These comparisons are converted to "fold change" measures for each sample/TaqMan assay tested. These values are averaged between three repeat experiments and shown in the bar graphs in which samples are reordered based on growth level groupings evident in multicycle A549 virus growth data.

11.4 Results

The level of INFB1 mRNA in A549 cells after 24 hours of RSV wt infection is 78% of the level at 18 hours of RSV wt infection. More dramatically two other members of the wt-like growth group, NS1 29 and NS1 98-109, reduce expression of IFNB1 mRNA from 60-20% and 80-40% (respectively) between 18 and 24 hours of infection as compared to IFNB1 mRNA levels in RSV wt at 18 hours. This reduction contrasts with the level of IFNB1 mRNA expressed in the intermediate-low growth virus set. Of interest is the observation that there appears to be two types of IFNB1 mRNA expression group types/dynamics. The first group, NS1 47, NS1 11-13-15 and NS1 58-66 have peak

IFNB1 mRNA expression at 18 hours that subsequently diminishes. The second group, NS1 06-09, NS1 54-57-60, NS1 98-104st and ΔNS1, have peak IFNB1 mRNA expression at 24 hours. Lastly the four single amino acid substitution virus group have substantially reduced IFNB1 mRNA expression profiles very similar to RSV wt.

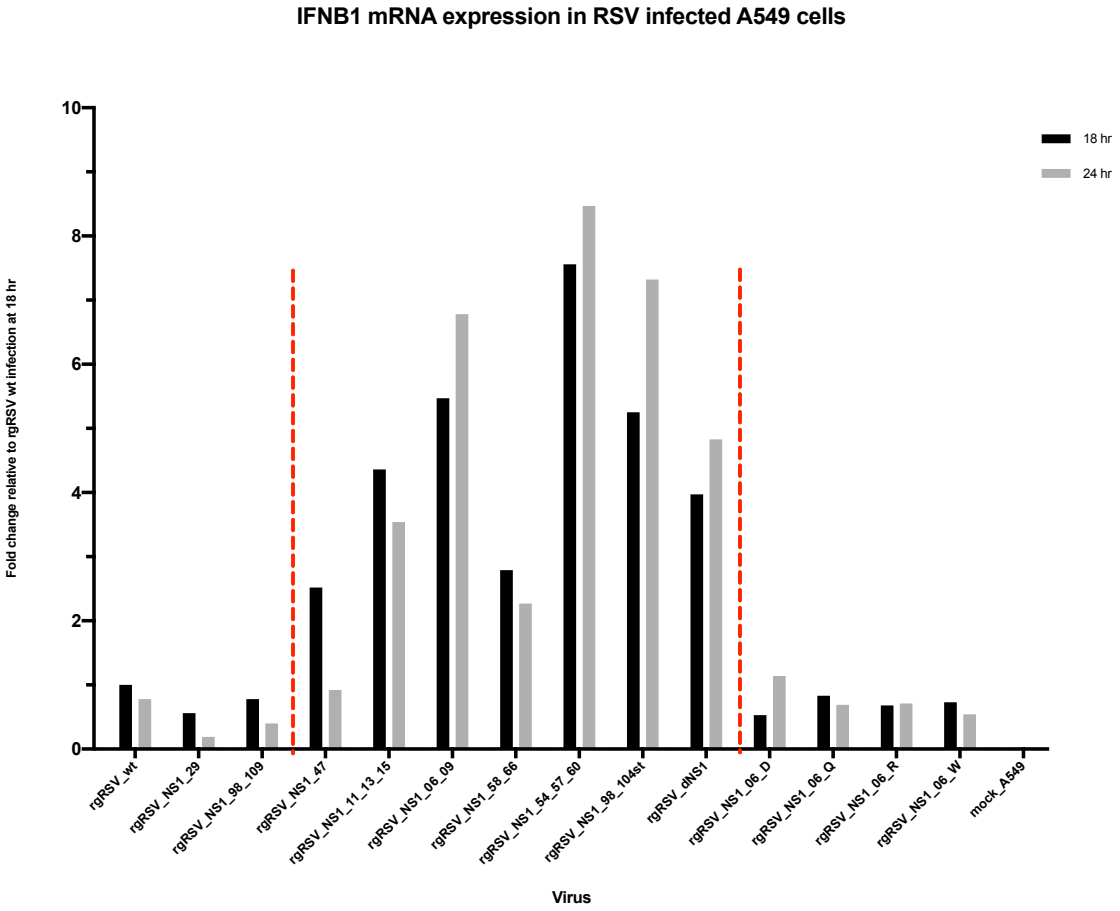


Figure 32 Host IFNB1 gene expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Host gene expression levels of IFNB1 were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a TaqMan assay IFNB1 (Hs01077958_s1). Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

Another host gene that participates in the innate immune response and whose mRNA expression is regulated during virus infection, and particularly RSV, is interferon lambda 1 (IFNL1) [13]. This cytokine (along with IFNL2 and IFNL3) are Type III interferons and distant relatives to Type I such as INFB1. During RSV wt infection of A549 cells, IFNL1 mRNA expression is 1.9 fold increased at 24 hours as compared to 18 hours. The other members of the wt-like growth group exhibit an INFL1 mRNA expression profile equivalent to RSV wt at 18 hours and fluctuate between 0.9-1.2 fold increase by 24 hours. The intermediate-low growth virus group reveal heterogeneity in IFNL1 mRNA expression restrictions. At 18 hours, NS1 47, NS1 11-13-15, NS1 58-66, NS1 98-104st and ΔNS1 seem to keep IFNL1 mRNA expression similar to RSV wt levels at 18 hours (ranging 0.7-1.3 fold increase). By 24 hour, IFNL1 mRNA expression levels among these virus infections fluctuate between 0.9 and 2.6 fold increase. However, two other members, NS1 06-09 and NS1 54-57-60, seem less able to restrict IFNL1 mRNA expression. By 24 hours IFNL1 mRNA expression is 7.0 to 8.2 fold higher than RSV wt (18 hour). The remaining four single amino acid replacement viruses all appear able to limit IFNL1 mRNA expression much like RSV wt with 24 hour fold changes ranging from 1.0-1.2 fold change.

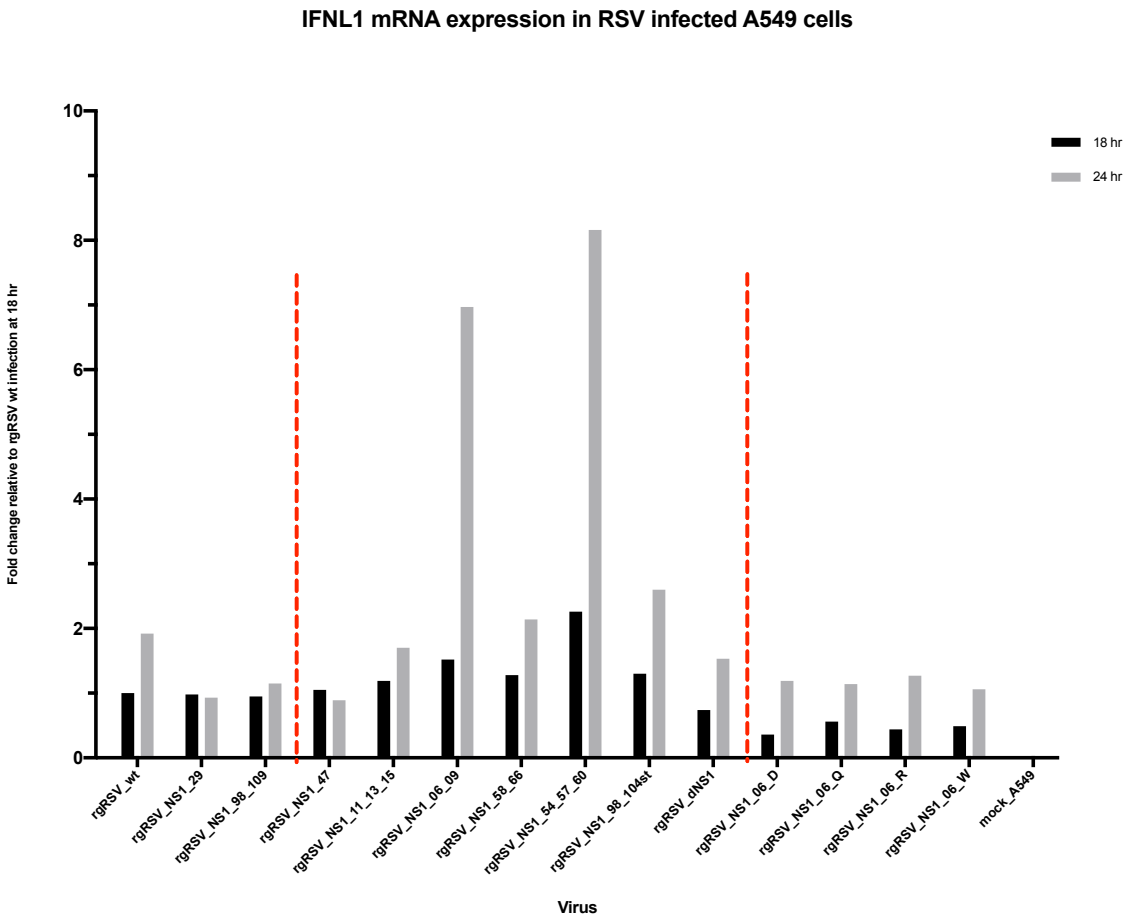


Figure 33 Host IFNL1 gene expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Host gene expression levels of IFNL1 were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a TaqMan assay IFNB1 (Hs00601677_g1). Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

The expression of C-C motif chemokine ligand 5 (CCL5) mRNA in RSV wt sample increased 5.7 fold (relative to RSV 18 hour level) by 24 hours after infection followed closely by the two other members of the wt-like growth group, NS1 29 and NS1

98-109, with peak CCL5 mRNA expression levels of 4.3 and 4.4 fold increase respectively at 24 hours. Similar to the expression dynamics seen with IFNL1 mRNA expression, the intermediate-low growth virus group revealed CCL5 mRNA expression heterogeneity after virus infection. At 18 hours, NS1 47, NS1 11-13-15, NS1 58-66, NS1 98-104st and ΔNS1 seem to keep CCL5 mRNA expression close to or below RSV wt at 18 hours (ranging 0.3-0.9 fold increase). By 24 hour, CCL5 mRNA expression levels among these virus infections fluctuated between 1.0 and 2.7 fold increase. However, two other members, NS1 06-09 and NS1 54-57-60, seem less able to restrict CCL5 mRNA expression. By 24 hours CCL5 mRNA expression is 5.8 to 6.6 fold higher than RSV wt (18 hour). The remaining four single amino acid replacement viruses with abundant CCL5 mRNA expression by 24 hours, fell into two categories. NS1 06D induces CCL5 mRNA expression virtually identical to RSV wt (1.2-6.6 fold increase). The remaining three, NS1 06Q, R and W while capable of inducing CCL5 mRNA expression at a rate similar to RSV wt, have much lower measurable expression at 18 hours. By starting at 60-70% of the amount expressed in RSV wt (18 hours), the peak expression level at 24 hours does not exceed 3.2-3.8 fold increase. Therefore, the absolute expression levels of CCL5 mRNA are different among the single amino acid substitution viruses, but the rate of change for CCL5 mRNA in this group is virtually the same. Moreover, this rate is almost identical to the expression rate change for CCL5 mRNA in wt-like virus infected samples.

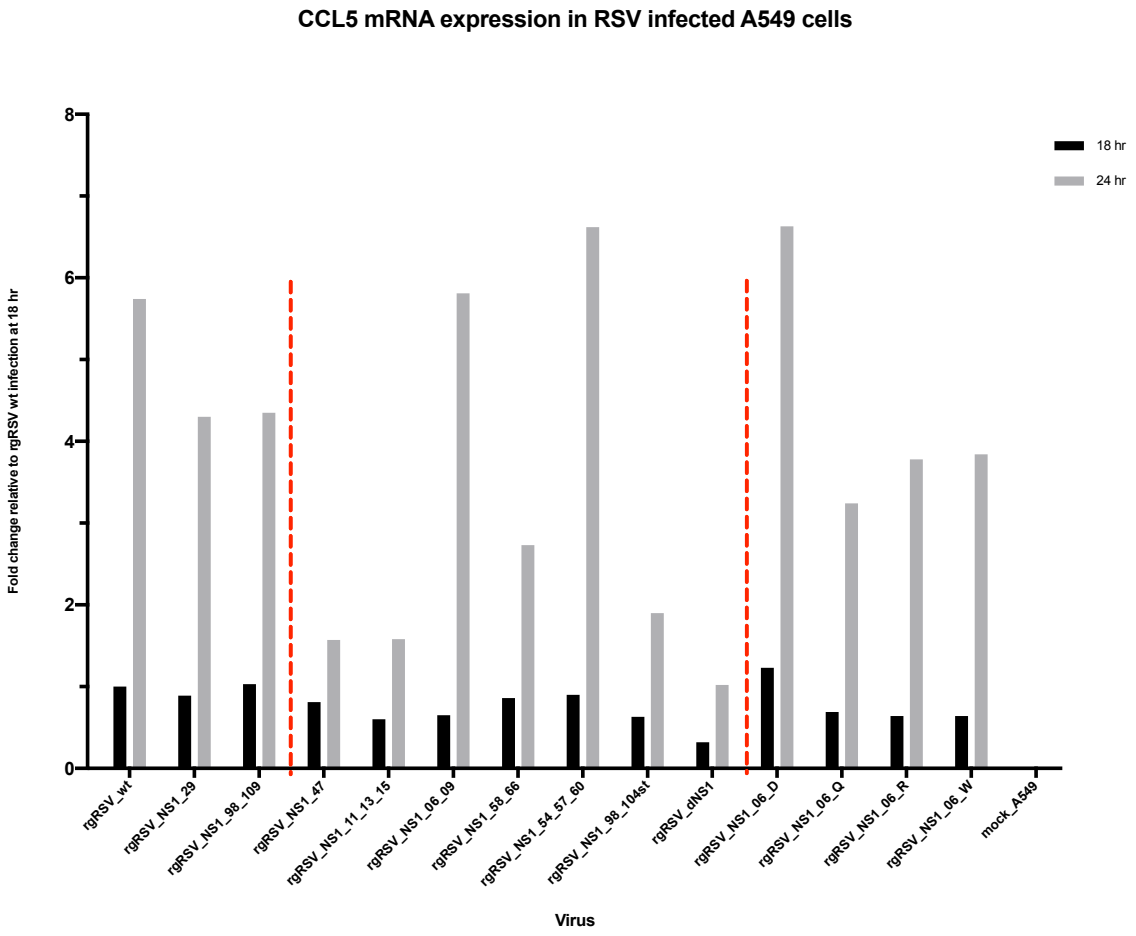


Figure 34 Host CCL5 (RANTES) gene expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Host gene expression levels of CCL5 were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a TaqMan assay CCL5 (Hs00174575_m1). Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

Tumor necrosis factor alpha (TNFα) is a proinflammatory cytokine activated upon RSV infection [223,224]. Among the complete set of viruses, the TNFα mRNA expression dynamics appear to be reciprocal to the IFNB1 levels in each growth set. The

expression of TNFα mRNA in RSV wt sample increased to 3.8 fold by 24 hours after infection. The other wt-like growth group members, NS1 29 and NS1 98-109, have TNFα mRNA levels of expression beginning at 1.2 and 0.8 fold increase respectively at 18 hours. By 24 hours, these levels peak at 1.5 and 2.4 fold increase. The level of TNFα mRNA expression in the intermediate-low growth virus group is generally lowered compared to the other virus groups. At 18 hours, NS1 11-13-15, NS1 06-09, NS1 58-66 and NS1 55-57-60 are profoundly lower with TNFα mRNA expression 0.1-0.2 fold increase. By 24 hours these levels remain low at 0.2-0.3 fold increase. NS1 47 has similarly reduced levels of TNFα mRNA expression but not as low; fluctuating around 0.7 fold increase relative to RSV wt, 18 hours. Lastly, the level of TNFα mRNA expression in NS1 98-104st and ΔNS1 present interesting dynamics. TNFα mRNA expression in NS1 98-104st are low at 18 hours and continue lower by 24 hours (0.7-0.6 fold increase). In contrast, TNFα mRNA expression in ΔNS1 is quite low at 18 hours but while still low, appears to be less so by 24 hours (0.3-0.9 fold increase). The four single amino acid replacement viruses all have limited TNFα mRNA expression compared to RSV wt at 18 hours (ranging from 0.3-0.6 fold increase). By 24 hpi, TNFα mRNA expression seems to fall into two categories. NS1 06D and R maintain limited TNFα mRNA expression (1.1 and 1.5 fold increase respectively) relative to RSV wt 18 hours. The remaining two, NS1 06Q and W, continue to increase TNFα mRNA expression similar to the expression levels seen at 24 hours in the wt-like growth group (2.5 and 4.3 fold increase respectively).

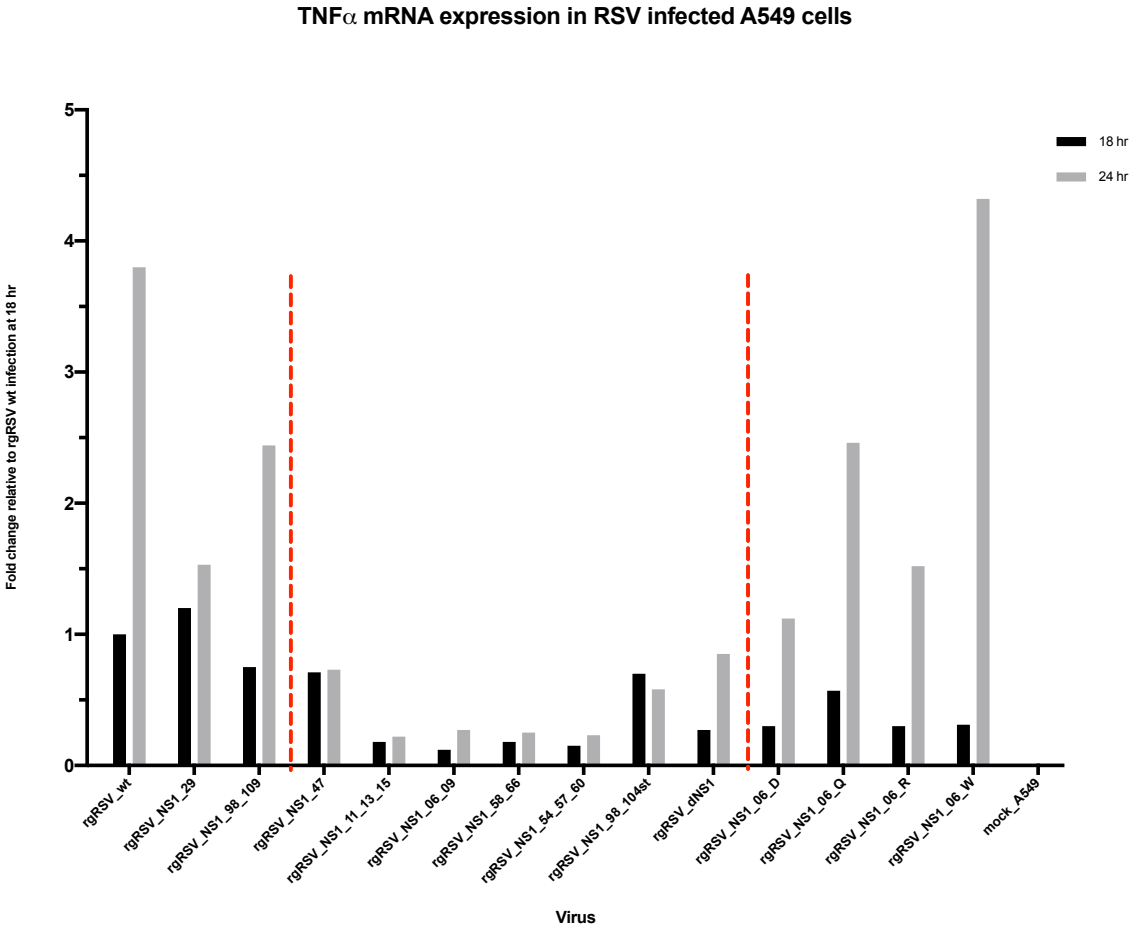


Figure 35 Host TNF alpha (TNFα) gene expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Host gene expression levels of TNFα were measured in total RNA samples collected at 18 and 24 hours post infection (hpi) using a TaqMan assay TNF (Hs00174128_m1). Measured expression levels (Ct) were normalized on a per sample basis to its 18S rRNA (Hs99999901_s1) measure then calibrated to the same 18S normalized measure in RSV wt 18 hpi sample providing a fold change value. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

11.5 Discussion

The ability of viruses to affect host gene transcription is well established [11,219,229]. The way viruses accomplish such a task is as varied as the number of viruses in existence. Looking specifically at the role the RSV NS1 gene product may play in this process is important considering the effects NS1 seems to have on viral gene regulation. RSV infection is known to activate proinflammatory cytokines and chemokines [223,224]. NS1 is suggested to antagonize infected cell innate immune response, particularly IFNB1 production [12,13,204,219]. The computationally predicted amino acid substitutions in NS1 are grouped in terms of A549 viral growth characteristics. Looking at various host gene transcript levels allow the evaluation of these amino acid substitutions with regard to the computational predictions. It has been shown that levels of IFNB1 message is differentially expressed when comparing RSV wt to ΔNS1 viral infections [13]. The use of a high MOI = 3 and an 18-24 hour time period helps ensure that measured results are related to a primary effect of NS1 but does not address the mechanism by which this may occur.

In agreement with prior work, the same discrepancy in IFNB1 mRNA expression levels is found between RSV wt and ΔNS1 in this work. The level of IFNB1 expressed mRNA is more than 20 fold higher in the intermediate-low growth group (which includes ΔNS1), compared to the wt-like growth group during this 18-24 hour time frame. [Average wt-like (18-24 hour) = 0.8-0.5 fold increase; average intermediate-low (18-24 hour) = 4.6-4.9 fold increase]. The single amino acid substitution virus set has IFNB1 expression levels more like wt-like virus group. This similarity is may reflect the NS1

substitutions 1) have maintained “wt” NS1 function relating to affecting host IFNB1 expression during the 18-24 hour time frame or 2) allow close to wt like virus growth levels/NS1 protein production leading to wt-like reduction of host IFNB1 expression.

Levels of IFNL1 have also been shown to have an expression differential when comparing RSV wt to ΔNS1 viral infections [13]. This discrepancy is not seen in these data regarding IFNL1 mRNA expression. The level of IFNL1 expressed mRNA is very similar at 18 hours between wt-like and intermediate-low growth virus sets. Even at 24 hours the expression levels continue to be similar with the exclusion of NS1 06-09 and NS1 54-57-60 samples. The increase in IFNL1 mRNA expression in these two samples is curious and without explanation. It is worth following up to see if these specific amino acid changes highlight regions of discrete function in NS1. Lastly the level of IFNL1 mRNA expressed by the single amino acid substituted virus group are unremarkable and minimally increased at both 18 and 24 hours compared to (up about 1.5 fold increased from 18-24 hours).

In addition to these classic Type I and III IFN proteins, RSV is also known to affect inflammatory chemokines and cytokines [12,230,231]. Unlike the Type I and III IFN, transcription of CCL5 (a proinflammatory chemokine) mRNA is increased during RSV infection [12,230]. RSV, as well as NS1 29 and NS1 98-109 viruses all show increased levels compared to the intermediate-low growth group. The expression levels of CCL5 mRNA among the wt-like growth group members at 18 hours are virtually identical to RSV wt and continue to increase almost 5 fold by 24 hours (5.7, 4.3 and 4.4 fold increase respectively). The level of CCL5 expressed mRNA is similar at 18 hours

between wt-like and intermediate-low growth virus sets. By 24 hours however, the expression levels are no longer similar with the exclusion of NS1_06-09 and NS1_54-57-60 samples. The increase in CCL5 mRNA expression in these two samples is reminiscent of the unusual increase seen in IFNL1 mRNA. The level of CCL5 mRNA expressed by the single amino acid substituted virus group appear unique. NS1 06D exemplifies wt-like expression levels between 18-24 hours (1.2-6.6 fold increase). NS1 06Q, R and W seem to have a lower 24 hour peak level of CCL5 mRNA expression.

Lastly, TNFα, another proinflammatory cytokine, is also upregulated during RSV infection [12]. RSV, as well as NS1 29 and NS1 98-109 viruses all show increased but varied levels of TNFα mRNA expression. The expression levels of TNFα mRNA among the wt-like growth group members at 18 hours are very similar to RSV wt. At 24 hours however the expression patterns are more distinct but increased: RSV wt; 3.8 fold increase, NS1 98-109; 2.4 fold increase and NS1 29; 1.5 fold increase. The level of TNFα expressed mRNA in the intermediate-low growth virus set is generally much lower but varied at both 18 and 24 hours compared to the TNFα mRNA expression levels in the wt-like set. NS1 11-13-15, NS1 06-09, NS1 58-66 and NS1 54-57-60 have the lowest expression levels at 18 and 24 hours (0.2-0.2, 0.1-0.3, 0.2-0.3 and 0.2-0.2 fold increase respectively). The remaining members, NS1 47, NS1 98-104st and ΔNS1 express TNFα mRNA in a more varied fashion (0.7-0.7, 0.7-0.6 and 0.3-0.9 fold respectively).

The level of TNFα mRNA expressed by the single amino acid substituted virus group appear unique and unlike the pattern seen in CCL5 expression. NS1 06D and R

TNFα mRNA expression levels between 18-24 hours (0.3-1.1 and 0.3-1.5 fold increase). NS1 06Q is 0.6-2.5 fold increase and NS1 06W is 0.3-4.3 fold increase. It is unclear what these varied expression levels mean or whether they have any relationship to the predictions. The overall picture suggests that there is a general association among the host gene expression profiles and the AUTOMUTE 2.0 predictions. The direction/magnitude of influence on a given host gene expression however is not well associated.

CHAPTER 12. HOST CELL PROTEIN EXPRESSION (ELISA)

12.1 Abstract

While certain amino acid substitutions in NS1 appear to affect host mRNA expression in response to RSV infection, it is anticipated this antagonism would evidently affect host protein expression. Since host innate immune response pathways include activation of and mediation by two unique cytokine proteins, IFN-β and IFN-λ1 whose protective activities are countered by NS1, enzyme-linked immunosorbent assay (ELISA) was used to measure the level of IFN-β [13] and IFN-λ1 proteins found in virus infected A549 cell growth media. Measuring the levels of these interferon proteins provides insights into whether NS1 stability predictions influence the expression of interferon proteins in infected cells.

12.2 Introduction

The antagonism of host defense systems is known to not only influence host gene expression, but related host protein translation/production. The influence RSV NS1 may have on infected cells is suggested by the differences seen in host gene expression levels among NS1 mutants and RSV wt but the extent these changes in host gene expression are reflected in translated protein levels must be measured. In this work the amount of IFN-β and IFN-λ1 protein present in cell growth media samples from infected A549 cells provide an additional measure regarding any impact NS1 substitutions may have on antiviral host responses. The biologic activity of these classic Type I and III IFN's are not measured, rather just the output levels of each in the environment.

12.3 Material and methods

Cell Line

A549 cells (adenocarcinoma human alveolar basal epithelial cells) (ATCC) were maintained in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 10% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). A549 cell infections were in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 2% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). A549 cells provide a relevant experimental system to examine virus growth in a fully IFN competent human lung epithelial host cell system.

Virus Infections

Triplicate experiments of A549 cells in 12 well dishes were infected with virus and media collected at 18, 24 and 36 hours post infection. For these virus infections, virus stocks were diluted to a multiplicity of infection of 3 (MOI = 3) pfu/cell in OptiPRO SFM (Gibco-Life Technologies), 4 mM L-glutamine (Gibco-Life Technologies) and added to cells in appropriate volumes. Infections were left on cells for 1-2 hours with periodic rocking at 37°C; 5% CO_2. Subsequently the infectious volume is removed, monolayers washed with OptiPRO SFM, 4 mM L-glutamine media and refed with F-12K, 2% FBS, 2 mM L-glutamine. Infected cells were left at 37°C; 5% CO_2 until sample collection.

Cytokine Assays (ELISA)

Culture supernatant from triplicate experiments of infected A549 cells were collected at 18, 24 and 36 hours. The concentration of IFN-β was measured using the

VeriKine Human IFN Beta ELISA Kit (PBL) and IFN-λ1 was measured using the Human IL-29 ELISA Kit (Invitrogen) as recommended by the manufacturers. Media from infected samples were diluted fivefold for IFN-β and one hundred-fold for IFN-λ1 in OptiPRO-SFM media. All samples were measured in triplicate; all standards were measured in duplicate. ELISA plates measurements were recorded as amount of absorbance at 450 nm using a Synergy 2 Multi Mode Microplate Reader (BioTek) as recommended by the manufacturer. Standard curves were generated using a 4-parameter fit in the Gen5 software (Version 3.04, BioTek) and used to convert absorbance measurements into absolute concentrations of each cytokine in pg/ml units for samples in each experiment. These values are averaged between three repeat experiments and shown in the bar graphs in which samples are reordered based on growth level groupings evident in multicycle A549 virus growth data.

12.4 Results

The levels of detected IFN-β protein among all virus groups does not vary by more than 4 fold. It appears that RSV wt IFN-β protein levels peak at 24 hours (6290 pg/ml) and recede by 36 hours (4280 pg/ml). During the same 24-36 hour time frame, a similar IFN-β expression pattern is seen in RSV 29 (3698-2933 pg/ml) and RSV 98-109 (5754-4326 pg/ml). The intermediate-low virus group present a more mixed phenotype. NS1 47, NS1 06-09 and NS1 58-66 seem better than RSV wt at keeping IFN-β protein expression reduced (18-36 hour range of expression; 1509-3152 pg/ml). Other members, NS1 11-13-15 and NS1 54-57-60 regulate IFN-β expression like RSV wt (18-36 hour range of expression; 3067-4943 pg/ml). The remaining members of the intermediate-low

growth group exhibit distinctly different IFN-β dynamics. NS1 98-104st allows some of the highest levels of IFN-β protein expression (18, 24 and 36 hours; 5246, 6836 and 6244 pg/ml) while ΔNS1 infected samples reveal IFN-β protein expression delayed at 18 hours but achieve wt levels by 36 hours (18, 24 and 36 hours; 2226, 4922 and 5814 pg/ml). Three of the four single amino acid substituted viruses, NS1 06Q, R and W, present an IFN-β expression profile pattern quite similar (average 18 hour, 1961 pg/ml; average 24 hour, 5103 pg/ml and average 36 hour, 4960 pg/ml) but lower than NS1 06D across the same time scale (8304, 7976 and 6452 pg/ml).

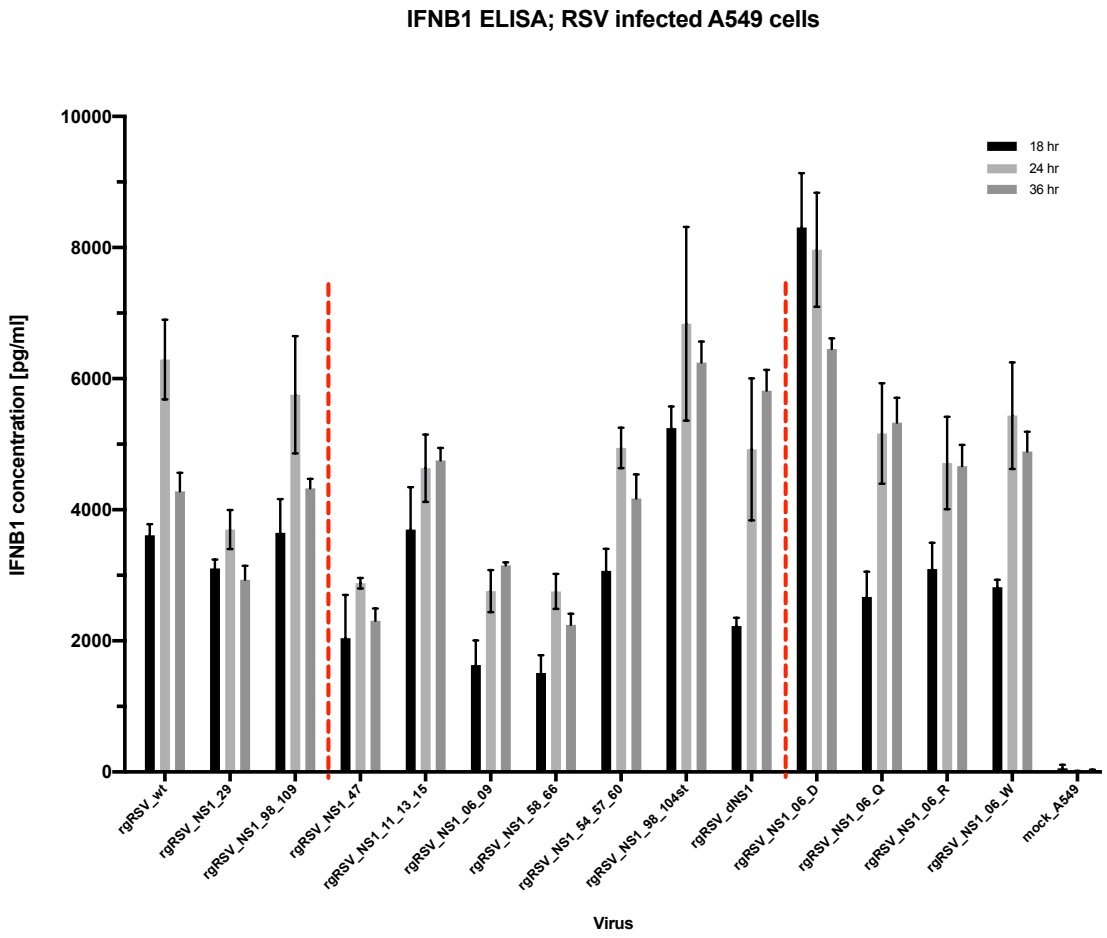


Figure 36 Host IFNB1 protein expression in virus infected A549 cell culture media
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Host protein expression levels of secreted IFNB1 were measured in cell growth medium collected at 18, 24 and 36 hours post infection (hpi) using an VeriKine Human IFN Beta ELISA Kit (PBL). Sample concentrations [pg/ml] were quantified based on a linear curve generated from standards. Average and standard deviations of the concentrations were calculated from triplicate experiment values. Vertical bars (Black-18 hpi, Light Grey-24 hpi or Dark Grey-36 hpi) represent the average concentration of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

The amount of IFN-λ1 protein detected presented a more consistent pattern within and between the virus growth groups. RSV wt, as well as NS1 29 and NS1 98-109, exhibit increasing IFN-λ1 level between 18-24 hours and peaking between 24-36 hours.

RSV wt had peak IFN-λ1 expression level (283767 pg/ml) at 24 hours while NS1 29 and NS1 98-109 were similarly elevated (250000 and 228700 pg/ml respectively) at 36 hours. A remarkable homogeneity of decreased IFN-λ1 protein expression is seen among the intermediate-low virus growth group, with peak levels at 36 hours essentially 30% the level of RSV wt. Among the four single amino acid substitution viruses, NS1 06D expresses wt-like levels of IFN-λ1 (18, 24 and 36 hours; 125433, 163133 and 170633 pg/ml) compared to NS1 06Q, R and W which were delayed only reaching peak IFN-λ1 expression levels (211733, 179167 and 172333 pg/ml respectively) similar to RSV wt at 36 hours.

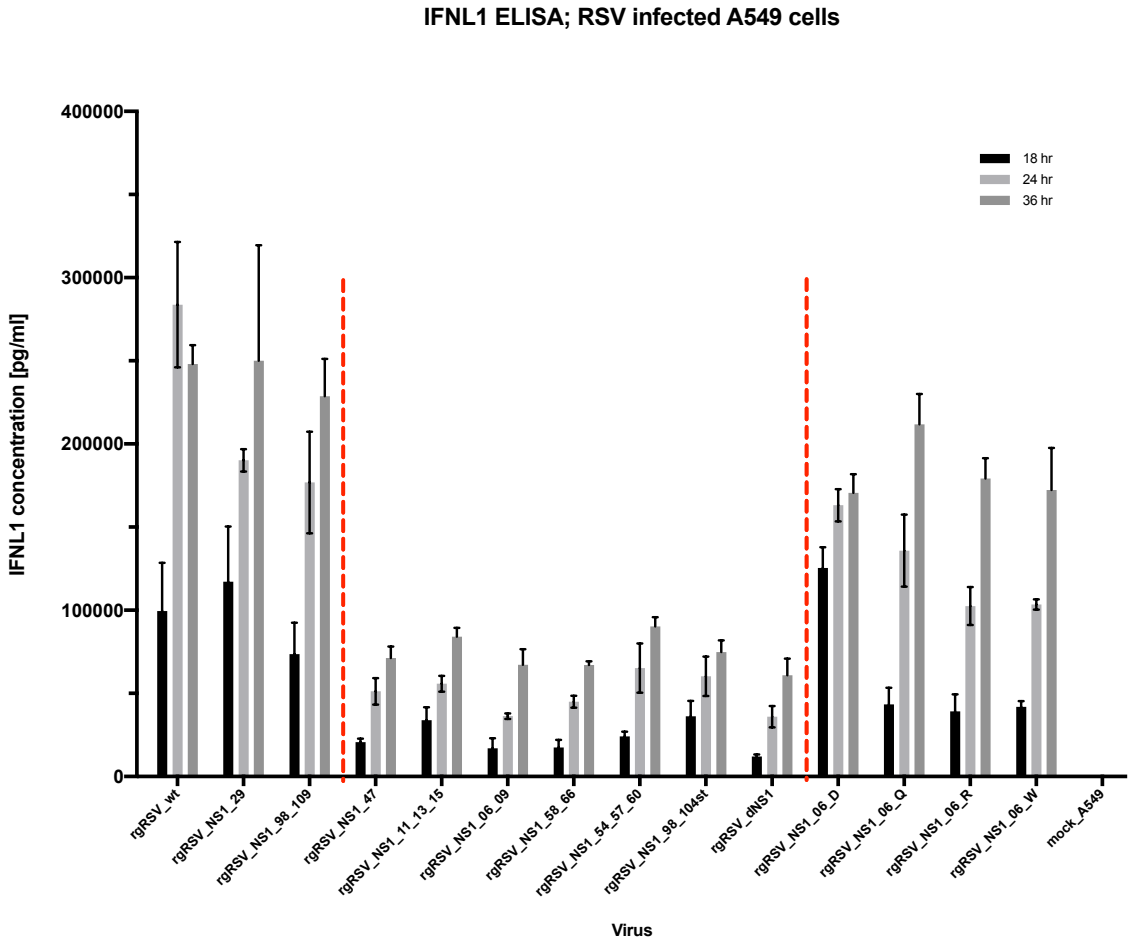


Figure 37 Host IFNL1 protein expression in virus infected A549 cell culture media
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Host protein expression levels of secreted IFNL1 were measured in cell growth medium collected at 18, 24 and 36 hours post infection (hpi) using an Human IL-29 ELISA Kit (Invitrogen). Sample concentrations [pg/ml] were quantified based on a linear curve generated from standards. Average and standard deviations of the concentrations were calculated from triplicate experiment values. Vertical bars (Black-18 hpi, Light Grey-24 hpi or Dark Grey-36 hpi) represent the average concentration of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

12.5 Discussion

It is believed that RSV NS1 protein can antagonize the innate immune response in infected cells and further suggested by the differences seen in host gene transcription among NS1 mutants and RSV wt [12,13,230]. While a previous finding that an RSV virus lacking the NS1 gene was able to produce more than twice the amount of IFN-β as compared to RSV wt [13], this work finds the 2.2 fold increase of IFN-β protein in ΔNS1 virus infected cells closely followed by a 1.7 fold increase found in the RSV wt infected sample at 24 hours. The other wt-like viruses tested had levels of detectable IFN-β proteins similar to levels seen with RSV wt infection.

A more mixed phenotype among the intermediate-low viruses indicates some viruses group (NS1 47, NS1 06-09 and NS1 58-66) seem better than RSV wt at keeping IFN-β protein expression reduced while others (NS1 11-13-15 and NS1 54-57-60) reduce IFN-β expression very similar to RSV wt. Interestingly, NS1 98-104st with a premature stop codon in the NS1 ORF, exhibits the highest levels of IFN-β protein expression and may be a more appropriate control sample for which there is no NS1 protein to interfere with host IFN-β protein expression. This virus appears less able to restrict/reduce the level of IFN-β protein expression regardless of the timepoint.

The four single amino acid substitution viruses suggest that there may be distinct regions of NS1 associated with influencing host innate immune response. NS1 06D, expressing the high levels of IFN-β is only a one or two amino acid different than its other virus relatives (NS1 06Q, R, W and NS1 06-09) which exhibit a wide range of IFN-

β expression during infection. This adds further interest to the "tunable vaccine" concept in terms of adjusting levels of broadly acting cell mediators of immunity.

A virus infected host cell response tool box also includes Type III interferon, IFN-λ1 [13]. The remarkable homogeneity of IFN-λ1 protein expression within each virus growth group has not been described for RSV infection beyond 24 hours. A consistent pattern of increased IFN-λ1 expression is seen with RSV wt, as well as NS1 29 and NS1 98-109. Conversely a consistent decreased pattern of IFN-λ1 expression was seen in the intermediate-low growth virus group. Lastly and similar to the phenomena described regarding IFN-β, the four single amino acid substitution viruses continue the suggestion that distinct regions of NS1 associate with influencing host innate immune response. NS1 06D, is limited in repressing IFN-λ1 expression at any timepoint during infection, in contrast to the other single amino acid different counterparts (NS1 06Q, R and W) which initially delay IFN-λ1 expression. It must be mentioned however that the level of IFN-λ1 protein expression seems related to some level of virus replication (which may or may not relate to proper NS1 function). In general, the level of IFN-λ1 expressed is correlated well with the virus growth grouping. The wt-like and single amino acid substitution viruses all grew well enough in A549 cells and as a result these infected cells can produce IFN-λ1 protein. The intermediate-low growth virus set are less able to express IFN-λ1 protein under the same infection conditions. To date there are no data comparing RSV wt and ΔNS1 in terms of IFN-λ1 expression level over the 18-36 hour time frame. As there is little functional work regarding the timing and consequence of IFN-λ1 protein

expression during RSV infection, these viruses help establish aspects of baseline measures additional experiments can extend.

The suggestion that RSV NS1 (and NS2) reduce Type I and III IFN (IFN-β and IFN-λ1 respectively) protein levels in infected cells [13,18,232] may be true over multicycle infections and relate to concentration dependent timing events, however during a high MOI infections and short time frames looking at primary host response events, the data is less clear. In this experiment, the suggestion becomes that without an intact NS1 protein, the level of IFN-λ1 is reduced. However previous data whereby retinoic acid-inducible gene I (RIG-I) knocked down with siRNA led to IFN-λ1 protein expression loss in A549 cells [232] and the report of NS1 interfering with RIG-I present some inconsistencies. In the later, NS1 interferes with RIG-I association to Mitochondrial antiviral-signaling (MAVS) and this prevents Type I (and III) IFN induction. But siRNA removal of RIG-I (which would be anticipated to also prevent RIG-I to associate with MAVS leads to increase in Type III IFN. Therefore, NS1 function is likely considerably more complex/dynamic and context dependent [18]. This work indicates that there could be a primary effect of NS1 on IFN-λ1 protein expression and there could be a primary effect of NS1 on IFNB1 mRNA expression.

The relationship between RSV infection and Type I and III IFN production is developing. Moreover, the expression relationship these proteins have with NS1 protein is also unclear likely due to the context in which these relationships are being evaluated. The reliability of IFN-β and IFN-λ1 production as surrogates of NS1 protein function may need reconsideration however the use of these host protein measurements in order to

see the effects of amino acid substitutions in NS1 on measurable interferon cytokine production indicate a relationship with the stability predictions made by the AUTOMUTE 2.0 program. In this work it appears that the phenotype outcomes may be incomplete reflections of NS1 effects.

CHAPTER 13. HOST CELL PROTEIN EXPRESSION (WESTERN BLOT)

13.1 Abstract

The level of signal transducer and activator of transcription 2 (STAT2) is believed to be reduced during RSV infection in order to interfere with various aspects of Janus kinase (Jak/STAT) signal transduction [83,204]. In particular it is reported that RSV NS1 plays a role in STAT2 reduction [227,228,233]. Consequently, the formation of interferon-stimulated gene factor 3 (ISGF-3) transcriptional activator complex (phosphorylated STAT2 and STAT1 associate with interferon regulatory factor 9 (IFR-9)) is unable to form and the IFN-stimulated genes (ISG) that containing IFN-stimulated response elements (ISRE) sequences are not upregulated. In this work we look to see whether any AUTOMUTE 2.0 stability predictions of NS1 amino acid substitutions influence STAT2 protein expression.

13.2 Introduction

The innate immune system is a collection of varied but generally focused series of evolved barriers cells create to interfere with successful virus production. These range from mild (restricted cell resource access) to extreme (programed cell death). Viruses in turn have evolved an equivalently varied set of responses to counter these barriers. A plethora of equilibriums are being disrupted and exploited during host-virus interactions; as virus exploits various cell processes for viral replication, the host attempts to re-equilibrate these very same aspects [214]. The cellular bureaucracy, initially, is no match for the hyperactive focus of a virus but will nevertheless attempt a holistic resolution. In

this work we look at whether NS1 substitution mutants influence STAT2 protein levels, an important mediator of multiple host defenses and whether there is an association between prediction mutant properties and STAT2 levels. There are reports that RSV NS1 plays a role in the reduction of STAT2 [227,228,233]. When the STAT2 protein is reduced, the formation of interferon-stimulated gene factor 3 (ISGF-3) transcriptional activator complex is unable to form and the IFN-stimulated genes (ISG) that containing IFN-stimulated response elements (ISRE) sequences are not upregulated. When this complex is not formed or reduced, less can translocate to the nucleus resulting in reduced ISRE gene transcription that both diminishes host antiviral immunity and perpetuates a cellular environment favoring virus replication.

13.3 Material and methods

Cell Line

A549 cells (adenocarcinoma human alveolar basal epithelial cell) (ATCC) were maintained in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 10% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). A549 cell infections were in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 2% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). A549 cells provide a relevant experimental system to examine virus growth in a fully IFN competent human lung epithelial host cell system.

Virus Infections

Triplicate experiments of A549 cells in 12 well dishes were infected with virus and samples collected at 18 and 24 hours post infection. For these virus infections, virus

stocks were diluted to a multiplicity of infection of 3 (MOI = 3) pfu/cell in OptiPRO SFM (Gibco-Life Technologies), 4 mM L-glutamine (Gibco-Life Technologies) and added to cells in appropriate volumes. Infections were left on cells for 1-2 hours with periodic rocking at 37°C; 5% CO_2. Subsequently the infectious volume is removed, monolayers washed with OptiPRO SFM, 4 mM L-glutamine media and refed with F-12K, 2% FBS, 2 mM L-glutamine. Infected cells were left at 37°C; 5% CO_2 until sample collection.

Western blot analysis for the quantification of host cell proteins

Triplicate experiments of A549 cells in 12-well were lysed with 100 µl 1x LDS sample buffer (Life Technologies). Fifteen (15) µL aliquots of lysate were denatured, reduced and electrophoresed on 4-12% Bis-Tris SDS gels (Life Technologies). Proteins were transferred onto PVDF membranes via the iBlot2 transfer system (Life Technologies). An anti-Stat2 (A-7) mouse mAb (Santa Cruz; sc-1668) was used in a Western blot to detect STAT2 protein in samples. Samples transferred to PVDF membranes were incubated with the primary antibody overnight at 16°C, rocking slowly. The primary antibody was subsequently detected using the corresponding infrared dye-conjugated goat anti-mouse immunoglobulin 800CW (Li-Cor, Lincoln, NE). Simultaneously, GAPDH protein, used as a protein sample loading control, was detected with an anti-rabbit polyclonal antibody (pAb) (Proteintech; 10494-1-AP) followed by detection with an infrared dye-conjugated goat anti-rabbit immunoglobulin 680RD (Li-Cor, Lincoln, NE). Western blot images were acquired on the Odyssey infrared scanner (Li-Cor) and analyzed with Image Studio Software (Version 5.2.5, Li-Cor). All protein

band intensity values were normalized to GAPDH and reported relative to protein levels found in the RSV wt infected sample at 18 hours. These values are averaged between three repeat experiments and shown in the bar graphs in which samples are reordered based on growth level groupings evident in multicycle A549 virus growth data.

13.4 Results

The wt-like group members (RSV wt, NS1 29 and NS1 98-109) all have substantially reduced STAT2 protein levels at both 18 and 24 hours after infection (14.5-29.8, 9.6-8.1 and 4.4-15.6 fold decrease relative to mock treated A549 cells at 18 hours respectively). Similarly, the intermediate-low growth group (NS1 47, NS1 11-13-15, NS1 06-09, NS1 58-66, NS1 54-47-60, NS1 98-104st and ΔNS1) show reduced STAT2 protein expression at 18 hours that continue to decline further by 24 hours (average = 3.3-6.3 fold decrease compared with mock treated A549 cells at 18 hours). Unexpectedly the ΔNS1 infection appears to have STAT2 protein levels almost as low as RSV wt. The NS1_98-104st virus which also does not produce NS1 proteins, but unlike ΔNS1 has the same full-length genome size as RSV wt, appears less able to reduce STAT2 protein levels (2-3 fold decrease compared to mock A549 18 hour). The single amino acid replacement set all show almost no loss of STAT2 at 18 hours (compare with STAT2 protein level in mock treated A549 cells at 18 hours), but by 24 hours, STAT2 protein expression is substantially lower (average = 6.5 fold decrease relative to STAT2 protein level in mock treated A549 cells at 18 hours).

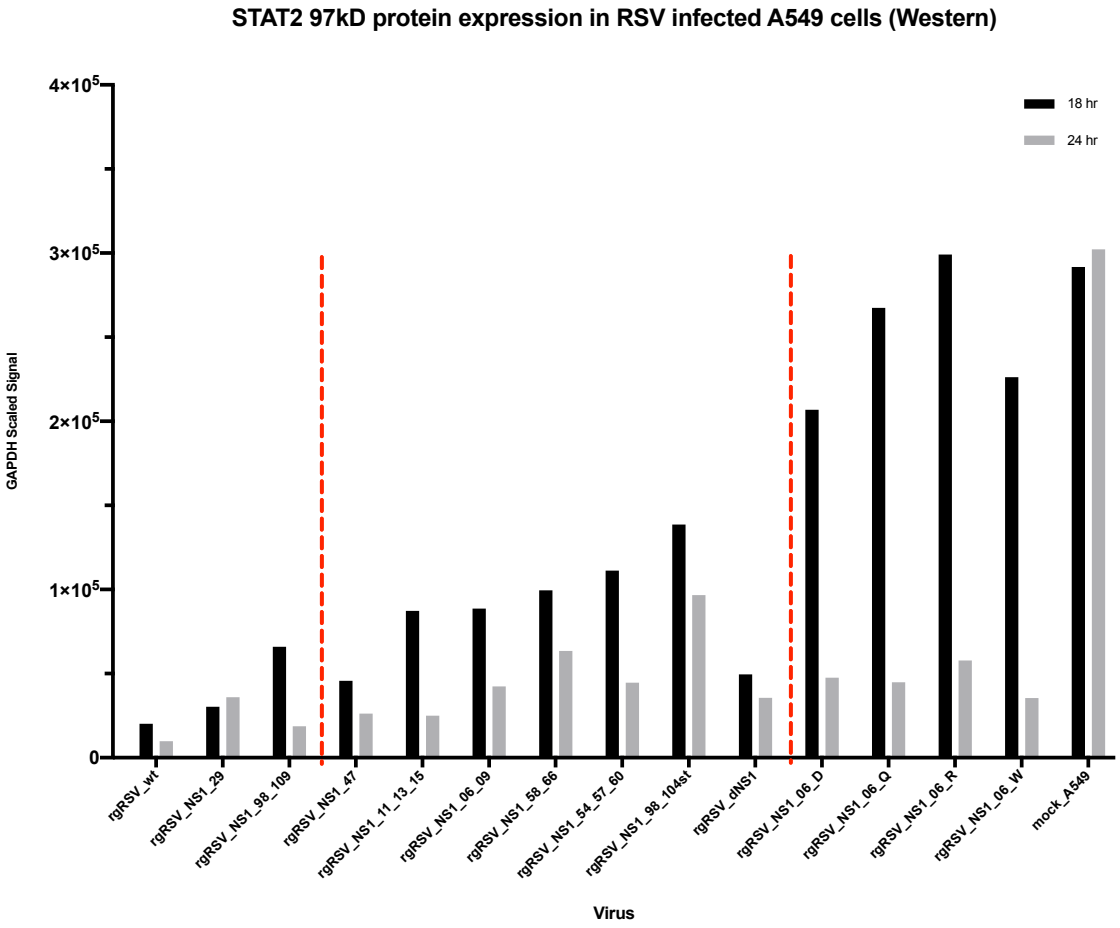


Figure 38 Host STAT2 protein expression in virus infected A549 cells
Monolayer cultures of A549 cells were infected in triplicate with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Host protein expression levels of STAT2 protein were measured in total infected cell lysate samples collected at 18 and 24 hours post infection (hpi) using a mouse monoclonal antibody against human STAT2 (Santa Cruz; sc-1668) as a primary antibody and an infrared dye-conjugated goat anti-mouse immunoglobulin 800CW (Li-Cor, Lincoln, NE) secondary antibody. GAPDH protein expression was detected using a rabbit polyclonal primary antibody 10494-1-AP (Proteintech) and an infrared dye-conjugated goat anti-rabbit immunoglobulin 680RD (Li-Cor, Lincoln, NE) secondary antibody. Each sample protein expression value was normalized to GAPDH, then calibrated to that protein expression level found in the RSV wt 18 hpi sample and reported as a GAPDH Scaled Signal. Vertical bars (Black-18 hpi or Grey-24 hpi) represent the average fold change of the triplicate experiment measures. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

13.5 Discussion

Unlike the Type I and III IFN's described previously, STAT2 is a host protein ubiquitously expressed at normal levels in A549 cells. During RSV infection, STAT2 protein levels are altered [227,228,233]. There appears to be a correlation between virus growth group and STAT2 protein level reduction. The viruses that grow like wt RSV as well as RSV wt itself appear quite capable of reducing detectable STAT2 protein as early as 18 hours and driving it down further up to 24 hours. While this is in agreement with reports [227,228,233] showing RSV infection of A549 cells leads to decreased STAT2 protein levels, it is not known whether specific regions of NS1 are responsible. Furthermore, variable amounts of STAT2 detected in two different samples, both of which lack NS1, raise questions regarding the NS1 protein associated STAT2 loss. These results suggest that the AUTOMUTE 2.0 stability predictions (primarily relating to virus growth groupings) appear to associate with levels/rates of STAT2 protein loss. Therefore, this data supports the idea that STAT2 loss is improved by fully functional NS1 (as seen in wt-like growth group), while changes in NS1 form (predicted amino acid substitutions) that alter NS1 function appear less able to reduce STAT2 protein levels (as seen in the intermediate-low growth group). The data from the four single amino acid substitutions further suggest that any of these specific changes will lead to reductions in STAT2 levels, but in a delayed fashion. In addition, it is possible that none of these amino acid changes completely defect the STAT2 degradation process, but rather suggest that 1) NS1 is a participant in the degradation process and 2) there may be regional features in NS1 that modulate the rate of STAT2 degradation which together provide a

less clear association between specific AUTOMUTE 2.0 predictions and STAT2 degradation levels.

CHAPTER 14. HOST CELL APOPTOSIS (LUMINESCENCE)

14.1 Abstract

Apoptosis is a cellular response of last resort that is delayed during RSV infection with NS1 protein reported as an instrumental antagonist [73,205,214,234]. This work will examine whether there is any agreement between amino acid substitutions and consequential apoptotic activity recognized by increased activation of the apoptotic cascade reliant upon Caspase 3 and 7 enzymes. Using a plate-based detection system, the timing and levels of Caspase 3/7 activation can infer antiviral apoptotic activity that occurs during viral infection of tissue culture cells.

14.2 Introduction

As mentioned previously, cells have evolved numerous barriers to infection and viruses have figured out how to circumvent these barriers. The most extreme of these barriers, apoptosis, leads to programmed cell death [235,236] and RSV (like many viruses) has developed temporal solutions to this problem [214,234] including the NS1 protein of RSV, described as affecting this host cell process during infection [73]. Apoptosis is a very comprehensive set of processes that can be started under an equally diverse set of conditions but leading to an irrecoverable end result. It is of interest to see whether the groups of amino acid substitutions (relating to the AUTOMUTE 2.0 predictions) result in group like outcomes relating to apoptosis as measured through the activation of Caspase 3 and/or 7 [237]. The increased level of these enzymes is indicative of eventual apoptotic consequences. Specifically, this work looks to identify

not only if apoptosis is occurring, but the timing and kinetics of this occurrence with relation to the AUTOMUTE 2.0 stability predictions.

14.3 Material and methods

Cell Line

A549 cells (adenocarcinoma human alveolar basal epithelial cell) (ATCC) were maintained in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 10% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). A549 cell infections were in F-12 Kaighn's Modification of Hams F12 with L-Glutamine (ATCC), supplemented with 2% FBS (HyClone), 2 mM L-glutamine (Gibco-Life Technologies). A549 cells provide a relevant experimental system to examine virus growth in a fully IFN competent human lung epithelial host cell system.

Virus Infections

Duplicate Cell Culture Microplate, 96 well PS F-Bottom – 655090 (Greiner Bio-One) were seeded with 40,000 A549 cells per well in 100 ul volumes. Quadruplicate wells were infected with virus and plates processed at 1, 2, 4, 6, 12, 18, 21, 24, 30, 36 and 48 hours post infection. For these virus infections, virus stocks were diluted to a multiplicity of infection of 3 (MOI = 3) pfu/cell in OptiPRO SFM (Gibco-Life Technologies), 4 mM L-glutamine (Gibco-Life Technologies) and 10 ul was added to cells. Infections were left on cells at 37°C; 5% CO_2 until sample collection.

Luminescence detection of Caspase 3/7 protein activities

Wells of infected A549 cells in 96 well PS F-Bottom Cell Culture Microplate – 655090 (Greiner Bio-One) were tested for caspase activity at 1, 2, 4, 6, 12, 18, 21, 24, 30,

36 and 48 hours post infection. The caspase activity was measured using Caspase-Glo 3/7 Assay (Promega) as recommended by the manufacturer. Quadruplicate Caspase activity measurements were recorded as the amount of luminescence produced (integrated signal for 1 second) using a Synergy 2 Multi Mode Microplate Reader (BioTek) as recommended by the manufacturer. The luminescence unit values were averaged among the quadruplicate wells and shown in the bar graphs in which samples are reordered based on growth level groupings evident in multicycle A549 virus growth data.

14.4 Results

Measuring the amount of Caspase 3/7 activity in cells is used to infer the level of eventual apoptosis that would be expected to occur in these same cells. Caspase 3/7 activity is directly related to the amount of a luminogenic compound cleaved into a substrate that can be used by luciferase to generate a luminescent signal [237]. The wt-like growth group of viruses, including RSV wt, all have similarly low levels of luminescence from 12 through 48 hours. In contrast the intermediate-low growth group of viruses have earlier and higher levels of luminescence (more than 2 fold increase) within the same time frame. The wt-like group have peak levels of luminescence between 24 and 36 hours after infection, the intermediate-low group appears to have a peak levels of luminescence activity at 21-24 hours. Finally, the 4 single amino acid substituted viruses seem to exemplify the measurable difference in apoptosis related caspase activity. The level of caspase activity is earlier and higher during NS1 06D virus

infection (peak luminescence at 21-24 hours) compared to the other members, NS1 06Q, R or W (peak luminescence at 36-48 hours).

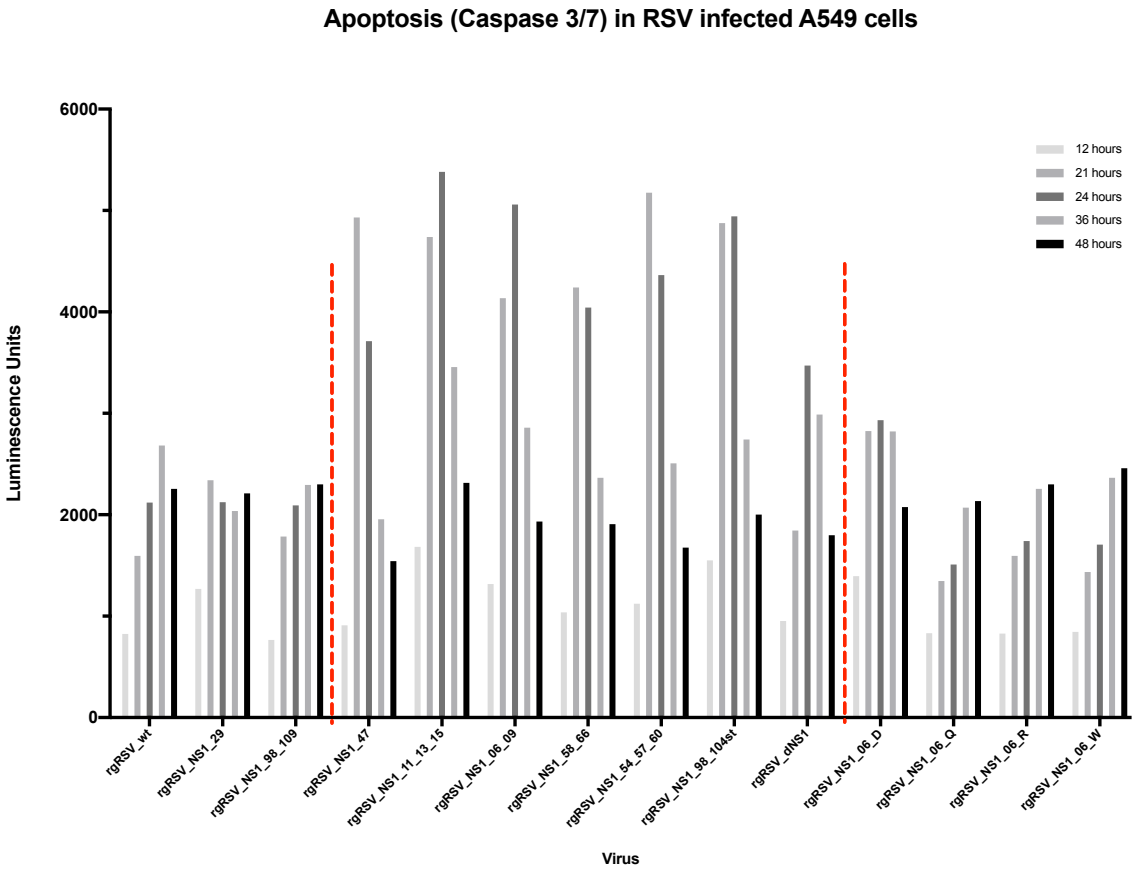


Figure 39 Host Caspase 3/7 activity (apoptosis) in virus infected A549 cells
Monolayer cultures of A549 cells in duplicate 96 well plates were infected with the Group 1, Group 2 NS1 substitution viruses, wt RSV-GFP, and RSV-GFP/ΔNS1 at a MOI of 3 PFU/cell and incubated at 37ºC. Caspase 3/7 activity was measured as a luminescence signal in averaged quadruplicate sample wells per plate at 1, 2, 4, 6, 12, 18, 21, 24, 30, 36 and 48 hours post infection (hpi) using the Caspase-Glo 3/7 Assay (Promega) kit. Vertical bars increasing in darkness from Light Grey on left (12 hpi) through Black on right (36 hpi) represent the average sample measurements from duplicate experiments taken from 12-36 hpi. Sample order was determined by decreasing multicycle growth levels in A549 cells. Red vertical dashed line separate RSV wt-like growth and intermediate-low growth (right) and single substitution plus mock (left).

14.5 Discussion

If NS1 protein is responsible for the reduced or delayed level of apoptosis in virus infected cells, then the data suggests that the wt-like virus set encode an NS1 protein that retains this functionality [73,205,234]. Conversely, the intermediate-low growth virus set data suggests that these viruses are less able to reduce and delay apoptosis. Furthermore, although the single amino acid replacement virus set replicate to quite similar levels in A549 cells, we see that D replacement seems to be more deleterious than the other (Q, R or W) replacements as indicated by the increase in apoptosis both level and timing. Together this data agrees with the notion that functional NS1 will help limit/delay apoptosis in infected cells [73,234]. Alternatively, the data also supports the suggestion that changes to the function of NS1 (amino acid changes) could eliminate this limit/delay property and therefore apoptosis may occur sooner and/or to a higher level in cells infected with defective NS1 containing viruses. A concern regarding overall viral replication levels influencing apoptosis is addressed with the single amino acid substitution virus set. Here the viruses all grow equivalently, yet there is a noticeable difference between amino acid 6 replaced with D (proapoptotic) and the other replacements Q, R and W (wt like; antiapoptotic). Taken together, these finding suggest that there is a general agreement between the stability predictions of amino acid replacements in NS1 and reduced function of NS1 in terms of antagonizing host cell apoptosis during infection.

CHAPTER 15. FUTURE DIRECTIONS

15.1 Abstract

The need to improve virus vaccine candidate design is indisputable with RSV itself having more than 100 different vaccine candidates [144–150] and development being hampered by the previous formalin-inactivated vaccine trial failure in the 1960's [20]. Computational biology is poised to bridge design gaps if not lead to a complete redesign of the process. In this work, the virus phenotypes are generally well matched with the algorithm predictions made by AUTOMUTE 2.0. While derived from a relatively small set of viruses, implementation of this type of objective pipeline could reduce or eliminate prior vaccine design failures while improve identification of attenuating changes in emerging virus vaccination efforts.

15.2 RSV F and G attenuation identification

The F and G proteins of RSV remain priority antigenic targets in current vaccine efforts. The identification and recognition of pre fusion (preF) protein structure as a vaccine target has energized the RSV vaccine field most recently. The use of available crystal structures for both RSV F (3RRR, 5ITB, 5UDC, etc.) and RSV G (6BLI, 5WN9, 5WNA, etc.), could provide a starting point in determining what other attenuation mutation(s) are possible in either of these two very important glycoproteins. The combination of computational mutation selection with reverse genetic system ability to create and test all variations will likely provide both functional insight to these proteins as well as refine the essentials of "best in class" vaccines.

15.3 Expansion of NS1 and NS2 attenuation identification

While this work identified useful changes in NS1 structure, it would be essential to further evaluate these in suitable animal (or human) vaccination studies. There continues to be the possibility that any combination of the identified NS1 changes could be added to existing vaccine candidates in an effort to "tune" or improve these vaccines. In addition, the same modeling/substitution selection process could be reused in order to explore NS2 changes that have never been tested.

15.4 Identification of improved non-viral protein structures

Finally, the concept of altering a proteins structure is not limited to viral proteins expressed during an infection. For instance, any class of antibody-based antivirals could be further refined through computationally identified changes to a particular antigenic protein that consequentially improves antibody protection.

REFERENCES

1. Hall, C.B.; Weinberg, G.A.; Iwane, M.K.; Blumkin, A.K.; Edwards, K.M.; Staat, M.A.; Auinger, P.; Griffin, M.R.; Poehling, K.A.; Erdman, D.; et al. The Burden of Respiratory Syncytial Virus Infection in Young Children. *N. Engl. J. Med.* **2009**, *360*, 588–598, doi:10.1056/NEJMoa0804877.
2. Carvajal, J.J.; Avellaneda, A.M.; Salazar-Ardiles, C.; Maya, J.E.; Kalergis, A.M.; Lay, M.K. Host Components Contributing to Respiratory Syncytial Virus Pathogenesis. *Front. Immunol.* **2019**, *10*, doi:10.3389/fimmu.2019.02152.
3. Mazur, N.I.; Higgins, D.; Nunes, M.C.; Melero, J.A.; Langedijk, A.C.; Horsley, N.; Buchholz, U.J.; Openshaw, P.J.; McLellan, J.S.; Englund, J.A.; et al. The Respiratory Syncytial Virus Vaccine Landscape: Lessons from the Graveyard and Promising Candidates. *Lancet Infect. Dis.* **2018**, *18*, e295–e311, doi:10.1016/S1473-3099(18)30292-5.
4. Collins, P.L.; Fearns, R.; Graham, B.S. Respiratory Syncytial Virus: Virology, Reverse Genetics, and Pathogenesis of Disease. In *Challenges and Opportunities for Respiratory Syncytial Virus Vaccines*; Anderson, L.J., Graham, B.S., Eds.; Current Topics in Microbiology and Immunology; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; pp. 3–38 ISBN 978-3-642-38919-1.
5. Correia, B.E.; Bates, J.T.; Loomis, R.J.; Baneyx, G.; Carrico, C.; Jardine, J.G.; Rupert, P.; Correnti, C.; Kalyuzhniy, O.; Vittal, V.; et al. Proof of Principle for Epitope-Focused Vaccine Design. *Nature* **2014**, *507*, 201–206, doi:10.1038/nature12966.

6. Collins, P.L.; Huang, Y.T.; Wertz, G.W. Identification of a Tenth MRNA of Respiratory Syncytial Virus and Assignment of Polypeptides to the 10 Viral Genes. *J Virol* **1984**, *49*, 572–8.

7. Collins, P.L.; Wertz, G.W. CDNA Cloning and Transcriptional Mapping of Nine Polyadenylylated RNAs Encoded by the Genome of Human Respiratory Syncytial Virus. *Proc. Natl. Acad. Sci.* **1983**, *80*, 3208–3212.

8. Sedeyn, K.; Schepens, B.; Saelens, X. Respiratory Syncytial Virus Nonstructural Proteins 1 and 2: Exceptional Disrupters of Innate Immune Responses. *PLoS Pathog.* **2019**, *15*, doi:10.1371/journal.ppat.1007984.

9. Byrd, L.G.; Prince, G.A. Animal Models of Respiratory Syncytial Virus Infection. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* **1997**, *25*, 1363–1368, doi:10.1086/516152.

10. Cherrie, A.H.; Anderson, K.; Wertz, G.W.; Openshaw, P.J. Human Cytotoxic T Cells Stimulated by Antigen on Dendritic Cells Recognize the N, SH, F, M, 22K, and 1b Proteins of Respiratory Syncytial Virus. *J. Virol.* **1992**, *66*, 2102–2110.

11. Bitko, V.; Shulyayeva, O.; Mazumder, B.; Musiyenko, A.; Ramaswamy, M.; Look, D.C.; Barik, S. Nonstructural Proteins of Respiratory Syncytial Virus Suppress Premature Apoptosis by an NF- B-Dependent, Interferon-Independent Mechanism and Facilitate Virus Growth. *J. Virol.* **2007**, *81*, 1786–1795, doi:10.1128/JVI.01420-06.

12. Spann, K.M.; Tran, K.C.; Collins, P.L. Effects of Nonstructural Proteins NS1 and NS2 of Human Respiratory Syncytial Virus on Interferon Regulatory Factor 3, NF-

KB, and Proinflammatory Cytokines. *J. Virol.* **2005**, *79*, 5353–5362, doi:10.1128/JVI.79.9.5353-5362.2005.

13. Spann, K.M.; Tran, K.-C.; Chi, B.; Rabin, R.L.; Collins, P.L. Suppression of the Induction of Alpha, Beta, and Lambda Interferons by the NS1 and NS2 Proteins of Human Respiratory Syncytial Virus in Human Epithelial Cells and Macrophages. *J. Virol.* **2004**, *78*, 4363–4369, doi:10.1128/JVI.78.8.4363-4369.2004.
14. Atreya, P.L.; Peeples, M.E.; Collins, P.L. The NS1 Protein of Human Respiratory Syncytial Virus Is a Potent Inhibitor of Minigenome Transcription and RNA Replication. *J VIROL* **1998**, *72*, 10.
15. Chatterjee, S.; Luthra, P.; Esaulova, E.; Agapov, E.; Yen, B.C.; Borek, D.M.; Edwards, M.R.; Mittal, A.; Jordan, D.S.; Ramanan, P.; et al. Structural Basis for Human Respiratory Syncytial Virus NS1-Mediated Modulation of Host Responses. *Nat. Microbiol.* **2017**, *2*, 1–8, doi:10.1038/nmicrobiol.2017.101.
16. Teng, M.N.; Whitehead, S.S.; Bermingham, A.; Claire, M.S.; Elkins, W.R.; Murphy, B.R.; Collins, P.L. Recombinant Respiratory Syncytial Virus That Does Not Express the NS1 or M2-2 Protein Is Highly Attenuated and Immunogenic in Chimpanzees. *J. Virol.* **2000**, *74*, 9317–9321, doi:10.1128/JVI.74.19.9317-9321.2000.
17. Evaluating the Infectivity, Safety and Immunogenicity of Respiratory Syncytial Virus Vaccines, RSV 6120/ΔNS1 and RSV 6120/F1/G2/ΔNS1, in RSV-Seropositive Children and RSV-Seronegative Infants and Children - Full Text View -

ClinicalTrials.Gov Available online: https://clinicaltrials.gov/ct2/show/NCT03596801 (accessed on 23 January 2020).

18. Collins, P.L.; Melero, J.A. Progress in Understanding and Controlling Respiratory Syncytial Virus: Still Crazy after All These Years. *Virus Res* **2011**, *162*, 80–99, doi:10.1016/j.virusres.2011.09.020.
19. Nair, H.; Nokes, D.J.; Gessner, B.D.; Dherani, M.; Madhi, S.A.; Singleton, R.J.; O'Brien, K.L.; Roca, A.; Wright, P.F.; Bruce, N.; et al. Global Burden of Acute Lower Respiratory Infections Due to Respiratory Syncytial Virus in Young Children: A Systematic Review and Meta-Analysis. *The Lancet* **2010**, *375*, 1545–1555, doi:10.1016/S0140-6736(10)60206-1.
20. Kim, H.W.; Canchola, J.G.; Brandt, C.D.; Pyles, G.; Chanock, R.M.; Jensen, K.; Parrott, R.H. RESPIRATORY SYNCYTIAL VIRUS DISEASE IN INFANTS DESPITE PRIOR ADMINISTRATION OF ANTIGENIC INACTIVATED VACCINE. *Am. J. Epidemiol.* **1969**, *89*, 422–434, doi:10.1093/oxfordjournals.aje.a120955.
21. Murphy, B.R.; Prince, G.A.; Collins, P.L.; Van Wyke Coelingh, K.; Olmsted, R.A.; Spriggs, M.K.; Parrott, R.H.; Kim, H.-W.; Brandt, C.D.; Chanock, R.M. Current Approaches to the Development of Vaccines Effective against Parainfluenza and Respiratory Syncytial Viruses. *Virus Res.* **1988**, *11*, 1–15, doi:10.1016/0168-1702(88)90063-9.
22. Connors, M.; Collins, P.L.; Firestone, C.-Y.; Sotnikov, A.V.; Waitze, A.; Davis, A.R.; Hung, P.P.; Chanock, R.M.; Murphy, B.R. Cotton Rats Previously

Immunized with a Chimeric RSV FG Glycoprotein Develop Enhanced Pulmonary Pathology When Infected with RSV, a Phenomenon Not Encountered Following Immunization with Vaccinia—RSV Recombinants or RSV. *Vaccine* **1992**, *10*, 475–484, doi:10.1016/0264-410X(92)90397-3.

23. Collins, P.L.; Whitehead, S.S.; Bukreyev, A.; Fearns, R.; Teng, M.N.; Juhasz, K.; Chanock, R.M.; Murphy, B.R. Rational Design of Live-Attenuated Recombinant Vaccine Virus for Human Respiratory Syncytial Virus by Reverse Genetics. *Adv. Virus Res.* **1999**, *54*, 423–451.
24. Wright, P.F.; Karron, R.A.; Belshe, R.B.; Thompson, J.; Crowe, J.E.; Boyce, T.G.; Halburnt, L.L.; Reed, G.W.; Whitehead, S.S.; Anderson, E.L.; et al. Evaluation of a Live, Cold-Passaged, Temperature-Sensitive, Respiratory Syncytial Virus Vaccine Candidate in Infancy. *J. Infect. Dis.* **2000**, *182*, 1331–1342, doi:10.1086/315859.
25. Afonso, C.L.; Amarasinghe, G.K.; Bányai, K.; Bào, Y.; Basler, C.F.; Bavari, S.; Bejerman, N.; Blasdell, K.R.; Briand, F.-X.; Briese, T.; et al. Taxonomy of the Order Mononegavirales: Update 2016. *Arch. Virol.* **2016**, *161*, 2351–2360, doi:10.1007/s00705-016-2880-1.
26. Pandya, M.C.; Callahan, S.M.; Savchenko, K.G.; Stobart, C.C. A Contemporary View of Respiratory Syncytial Virus (RSV) Biology and Strain-Specific Differences. *Pathog. Basel Switz.* **2019**, *8*, doi:10.3390/pathogens8020067.

27. Rima, B.; Collins, P.; Easton, A.; Fouchier, R.; Kurath, G.; Lamb, R.A.; Lee, B.; Maisner, A.; Rota, P.; Wang, L.; et al. ICTV Virus Taxonomy Profile: Pneumoviridae. *J. Gen. Virol.* **2017**, *98*, 2912–2913, doi:10.1099/jgv.0.000959.

28. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874, doi:10.1093/molbev/msw054.

29. Blakely, K.K.; Suttle, R.; Wood, T.; Stallworth, K.; Baker, N. Measles—What's Old Is New Again. *Nurs. Womens Health* **2020**, doi:10.1016/j.nwh.2019.11.005.

30. Bockelman, C.; Frawley, T.C.; Long, B.; Koyfman, A. Mumps: An Emergency Medicine-Focused Update. *J. Emerg. Med.* **2018**, *54*, 207–214, doi:10.1016/j.jemermed.2017.08.037.

31. Ustun, C.; Slabý, J.; Shanley, R.M.; Vydra, J.; Smith, A.R.; Wagner, J.E.; Weisdorf, D.J.; Young, J.-A.H. Human Parainfluenza Virus Infection after Hematopoietic Stem Cell Transplantation: Risk Factors, Management, Mortality, and Changes over Time. *Biol. Blood Marrow Transplant.* **2012**, *18*, 1580–1588, doi:10.1016/j.bbmt.2012.04.012.

32. Nichols, W.G.; Erdman, D.D.; Han, A.; Zukerman, C.; Corey, L.; Boeckh, M. Prolonged Outbreak of Human Parainfluenza Virus 3 Infection in a Stem Cell Transplant Outpatient Department: Insights from Molecular Epidemiologic Analysis. *Biol. Blood Marrow Transplant.* **2004**, *10*, 58–64, doi:10.1016/j.bbmt.2003.09.010.

33. Jalal, H.; Bibby, D.F.; Bennett, J.; Sampson, R.E.; Brink, N.S.; MacKinnon, S.; Tedder, R.S.; Ward, K.N. Molecular Investigations of an Outbreak of Parainfluenza Virus Type 3 and Respiratory Syncytial Virus Infections in a Hematology Unit. *J. Clin. Microbiol.* **2007**, *45*, 1690–1696, doi:10.1128/JCM.01912-06.

34. Chu, H.Y.; Englund, J.A.; Podczervinski, S.; Kuypers, J.; Campbell, A.P.; Boeckh, M.; Pergam, S.A.; Casper, C. Nosocomial Transmission of Respiratory Syncytial Virus in an Outpatient Cancer Center. *Biol. Blood Marrow Transplant.* **2014**, *20*, 844–851, doi:10.1016/j.bbmt.2014.02.024.

35. Middleton, D. Hendra Virus. *Vet. Clin. North Am. Equine Pract.* **2014**, *30*, 579–589, doi:10.1016/j.cveq.2014.08.004.

36. Ang, B.S.P.; Lim, T.C.C.; Wang, L. Nipah Virus Infection. *J. Clin. Microbiol.* **2018**, *56*, doi:10.1128/JCM.01875-17.

37. Shi, T.; McAllister, D.A.; O'Brien, K.L.; Simoes, E.A.F.; Madhi, S.A.; Gessner, B.D.; Polack, F.P.; Balsells, E.; Acacio, S.; Aguayo, C.; et al. Global, Regional, and National Disease Burden Estimates of Acute Lower Respiratory Infections Due to Respiratory Syncytial Virus in Young Children in 2015: A Systematic Review and Modelling Study. *Lancet Lond. Engl.* **2017**, *390*, 946–958, doi:10.1016/S0140-6736(17)30938-8.

38. Thorburn, K. Pre-Existing Disease Is Associated with a Significantly Higher Risk of Death in Severe Respiratory Syncytial Virus Infection. *Arch. Dis. Child.* **2009**, *94*, 99–103, doi:10.1136/adc.2008.139188.

39. Scheltema, N.M.; Gentile, A.; Lucion, F.; Nokes, D.J.; Munywoki, P.K.; Madhi, S.A.; Groome, M.J.; Cohen, C.; Moyes, J.; Thorburn, K.; et al. Global Respiratory Syncytial Virus-Associated Mortality in Young Children (RSV GOLD): A Retrospective Case Series. *Lancet Glob. Health* **2017**, *5*, e984–e991, doi:10.1016/S2214-109X(17)30344-3.

40. Hall, C.B.; Walsh, E.E.; Long, C.E.; Schnabel, K.C. Immunity to and Frequency of Reinfection with Respiratory Syncytial Virus. *J. Infect. Dis.* **1991**, *163*, 693–698, doi:10.1093/infdis/163.4.693.

41. Falsey, A.R.; Walsh, E.E. Respiratory Syncytial Virus Infection in Elderly Adults. *Drugs Aging* **2005**, *22*, 577–87.

42. Falsey, A.R.; Hennessey, P.A.; Formica, M.A.; Cox, C.; Walsh, E.E. Respiratory Syncytial Virus Infection in Elderly and High-Risk Adults. *N Engl J Med* **2005**, *352*, 1749–59, doi:10.1056/NEJMoa043951.

43. Thompson, W.W.; Shay, D.K.; Weintraub, E.; Brammer, L.; Cox, N.; Anderson, L.J.; Fukuda, K. Mortality Associated With Influenza and Respiratory Syncytial Virus in the United States. *JAMA* **2003**, *289*, 179–186, doi:10.1001/jama.289.2.179.

44. Glezen, W.P.; Denny, F.W. Epidemiology of Acute Lower Respiratory Disease in Children. *N. Engl. J. Med.* **1973**, *288*, 498–505, doi:10.1056/NEJM197303082881005.

45. Shay, D.K.; Holman, R.C.; Newman, R.D.; Liu, L.L.; Stout, J.W.; Anderson, L.J. Bronchiolitis-Associated Hospitalizations Among US Children, 1980-1996. *JAMA* **1999**, *282*, 1440–1446, doi:10.1001/jama.282.15.1440.

46. Graham, B.S. Vaccines against Respiratory Syncytial Virus: The Time Has Finally Come. *Vaccine* **2016**, *34*, 3535–3541, doi:10.1016/j.vaccine.2016.04.083.

47. Ke, Z.; Dillard, R.; Chirkova, T.; Leon, F.; Stobart, C.; Hampton, C.; Strauss, J.; Rajan, D.; Rostad, C.; Taylor, J.; et al. The Morphology and Assembly of Respiratory Syncytial Virus Revealed by Cryo-Electron Tomography. *Viruses* **2018**, *10*, 446, doi:10.3390/v10080446.

48. Kiss, G.; Holl, J.M.; Williams, G.M.; Alonas, E.; Vanover, D.; Lifland, A.W.; Gudheti, M.; Guerrero-Ferreira, R.C.; Nair, V.; Yi, H.; et al. Structural Analysis of Respiratory Syncytial Virus Reveals the Position of M2-1 between the Matrix Protein and the Ribonucleoprotein Complex. *J. Virol.* **2014**, *88*, 7602–7617, doi:10.1128/JVI.00256-14.

49. Kuo, L.; Fearns, R.; Collins, P.L. Analysis of the Gene Start and Gene End Signals of Human Respiratory Syncytial Virus: Quasi-Templated Initiation at Position 1 of the Encoded MRNA. *J. Virol.* **1997**, *71*, 4944–4953.

50. Johnson, P.R.; Collins, P.L. The A and B Subgroups of Human Respiratory Syncytial Virus: Comparison of Intergenic and Gene-Overlap Sequences. *J. Gen. Virol.* **1988**, *69*, 2901–2906, doi:10.1099/0022-1317-69-11-2901.

51. Collins, P.L.; Graham, B.S. Viral and Host Factors in Human Respiratory Syncytial Virus Pathogenesis. *J. Virol.* **2008**, *82*, 2040–2055, doi:10.1128/JVI.01625-07.

52. Knipe, D.M.; Howley, P.M. *Chapter 38 Respiratory Syncytial Virus and Metapneumovirus*; Fields Virology; 6th ed.; Wolters Kluwer/Lippincott Williams & Wilkins Health: Philadelphia, PA, 2013; Vol. 1; ISBN 9781451105636.

53. Fearns, R.; Peeples, M.E.; Collins, P.L. Mapping the Transcription and Replication Promoters of Respiratory Syncytial Virus. *J Virol* **2002**, *76*, 1663–72.

54. McGivern, D.R.; Collins, P.L.; Fearns, R. Identification of Internal Sequences in the 3′ Leader Region of Human Respiratory Syncytial Virus That Enhance Transcription and Confer Replication Processivity. *J. Virol.* **2005**, *79*, 2449–2460, doi:10.1128/JVI.79.4.2449-2460.2005.

55. Krzyzaniak, M.A.; Zumstein, M.T.; Gerez, J.A.; Picotti, P.; Helenius, A. Host Cell Entry of Respiratory Syncytial Virus Involves Macropinocytosis Followed by Proteolytic Activation of the F Protein. *PLoS Pathog.* **2013**, *9*, e1003309, doi:10.1371/journal.ppat.1003309.

56. Lingemann, M.; McCarty, T.; Liu, X.; Buchholz, U.J.; Surman, S.; Martin, S.E.; Collins, P.L.; Munir, S. The Alpha-1 Subunit of the Na+,K+-ATPase (ATP1A1) Is Required for Macropinocytic Entry of Respiratory Syncytial Virus (RSV) in Human Respiratory Epithelial Cells. *PLoS Pathog.* **2019**, *15*, e1007963, doi:10.1371/journal.ppat.1007963.

57. Kolokoltsov, A.A.; Deniger, D.; Fleming, E.H.; Roberts, N.J.; Karpilow, J.M.; Davey, R.A. Small Interfering RNA Profiling Reveals Key Role of Clathrin-Mediated Endocytosis and Early Endosome Formation for Infection by Respiratory Syncytial Virus. *J. Virol.* **2007**, *81*, 7786–7800, doi:10.1128/JVI.02780-06.

58. Chang, A.; Dutch, R.E. Paramyxovirus Fusion and Entry: Multiple Paths to a Common End. *Viruses* **2012**, *4*, 613–636, doi:10.3390/v4040613.

59. Lamb, R.A.; Paterson, R.G.; Jardetzky, T.S. Paramyxovirus Membrane Fusion: Lessons from the F and HN Atomic Structures. *Virology* **2006**, *344*, 30–37, doi:10.1016/j.virol.2005.09.007.

60. Kuo, L.; Grosfeld, H.; Cristina, J.; Hill, M.G.; Collins, P.L. Effects of Mutations in the Gene-Start and Gene-End Sequence Motifs on Transcription of Monocistronic and Dicistronic Minigenomes of Respiratory Syncytial Virus. *J. Virol.* **1996**, *70*, 6892–6901.

61. Dickens, L.E.; Collins, P.L.; Wertz, G.W. Transcriptional Mapping of Human Respiratory Syncytial Virus. *J Virol* **1984**, *52*, 364–9.

62. Noton, S.L.; Cowton, V.M.; Zack, C.R.; McGivern, D.R.; Fearns, R. Evidence That the Polymerase of Respiratory Syncytial Virus Initiates RNA Replication in a Nontemplated Fashion. *Proc. Natl. Acad. Sci.* **2010**, *107*, 10226–10231, doi:10.1073/pnas.0913065107.

63. Noton, S.L.; Fearns, R. The First Two Nucleotides of the Respiratory Syncytial Virus Antigenome RNA Replication Product Can Be Selected Independently of the Promoter Terminus. *RNA* **2011**, *17*, 1895–1906, doi:10.1261/rna.2813411.

64. Fearns, R.; Collins, P.L.; Peeples, M.E. Functional Analysis of the Genomic and Antigenomic Promoters of Human Respiratory Syncytial Virus. *J. Virol.* **2000**, *74*, 6006–6014, doi:10.1128/JVI.74.13.6006-6014.2000.

65. Yu, Q.; Hardy, R.W.; Wertz, G.W. Functional CDNA Clones of the Human Respiratory Syncytial (RS) Virus N, P, and L Proteins Support Replication of RS

Virus Genomic RNA Analogs and Define Minimal Trans-Acting Requirements for RNA Replication. *J. Virol.* **1995**, *69*, 2412–2419.

66. Grosfeld, H.; Hill, M.G.; Collins, P.L. RNA Replication by Respiratory Syncytial Virus (RSV) Is Directed by the N, P, and L Proteins; Transcription Also Occurs under These Conditions but Requires RSV Superinfection for Efficient Synthesis of Full-Length MRNA. *J Virol* **1995**, *69*, 5677–86.
67. Teng, M.N.; Collins, P.L. Identification of the Respiratory Syncytial Virus Proteins Required for Formation and Passage of Helper-Dependent Infectious Particles. *J. Virol.* **1998**, *72*, 5707–5716.
68. Zhang, L.; Peeples, M.E.; Boucher, R.C.; Collins, P.L.; Pickles, R.J. Respiratory Syncytial Virus Infection of Human Airway Epithelial Cells Is Polarized, Specific to Ciliated Cells, and without Obvious Cytopathology. *J. Virol.* **2002**, *76*, 5654–5666, doi:10.1128/JVI.76.11.5654-5666.2002.
69. Roberts, S.R.; Compans, R.W.; Wertz, G.W. Respiratory Syncytial Virus Matures at the Apical Surfaces of Polarized Epithelial Cells. *J. Virol.* **1995**, *69*, 2667–2673.
70. Bossert, B.; Marozin, S.; Conzelmann, K.K. Nonstructural Proteins NS1 and NS2 of Bovine Respiratory Syncytial Virus Block Activation of Interferon Regulatory Factor 3. *J Virol* **2003**, *77*, 8661–8.
71. Ren, J.; Liu, T.; Pang, L.; Li, K.; Garofalo, R.P.; Casola, A.; Bao, X. A Novel Mechanism for the Inhibition of Interferon Regulatory Factor-3-Dependent Gene Expression by Human Respiratory Syncytial Virus NS1 Protein. *J. Gen. Virol.* **2011**, *92*, 2153–2159, doi:10.1099/vir.0.032987-0.

72. Tran, K.C.; He, B.; Teng, M.N. Replacement of the Respiratory Syncytial Virus Nonstructural Proteins NS1 and NS2 by the V Protein of Parainfluenza Virus 5. *Virology* **2007**, *368*, 73–82, doi:10.1016/j.virol.2007.06.017.

73. Munir, S.; Nouen, C.L.; Luongo, C.; Buchholz, U.J.; Collins, P.L.; Bukreyev, A. Nonstructural Proteins 1 and 2 of Respiratory Syncytial Virus Suppress Maturation of Human Dendritic Cells. *J. Virol.* **2008**, *82*, 8780–8796, doi:10.1128/JVI.00630-08.

74. Evans, J.E.; Cane, P.A.; Pringle, C.R. Expression and Characterisation of the NS1 and NS2 Proteins of Respiratory Syncytial Virus. *Virus Res* **1996**, *43*, 155–61.

75. Pretel, E.; Camporeale, G.; de Prat-Gay, G. The Non-Structural NS1 Protein Unique to Respiratory Syncytial Virus: A Two-State Folding Monomer in Quasi-Equilibrium with a Stable Spherical Oligomer. *PLoS ONE* **2013**, *8*, e74338, doi:10.1371/journal.pone.0074338.

76. Hengst, U.; Kiefer, P. Domains of Human Respiratory Syncytial Virus P Protein Essential for Homodimerization and for Binding to N and NS1 Protein. *Virus Genes* **2000**, *20*, 221–225, doi:10.1023/A:1008188527858.

77. Hastie, M.L.; Headlam, M.J.; Patel, N.B.; Bukreyev, A.A.; Buchholz, U.J.; Dave, K.A.; Norris, E.L.; Wright, C.L.; Spann, K.M.; Collins, P.L.; et al. The Human Respiratory Syncytial Virus Nonstructural Protein 1 Regulates Type I and Type II Interferon Pathways. *Mol. Cell. Proteomics MCP* **2012**, *11*, 108–127, doi:10.1074/mcp.M111.015909.

78. Boyapalle, S.; Wong, T.; Garay, J.; Teng, M.; San Juan-Vergara, H.; Mohapatra, S.; Mohapatra, S. Respiratory Syncytial Virus NS1 Protein Colocalizes with Mitochondrial Antiviral Signaling Protein MAVS Following Infection. *PloS One* **2012**, *7*, e29386, doi:10.1371/journal.pone.0029386.

79. Li, X.-D.; Sun, L.; Seth, R.B.; Pineda, G.; Chen, Z.J. Hepatitis C Virus Protease NS3/4A Cleaves Mitochondrial Antiviral Signaling Protein off the Mitochondria to Evade Innate Immunity. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 17717–17722, doi:10.1073/pnas.0508531102.

80. Yang, Y.; Liang, Y.; Qu, L.; Chen, Z.; Yi, M.; Li, K.; Lemon, S.M. Disruption of Innate Immunity Due to Mitochondrial Targeting of a Picornaviral Protease Precursor. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 7253–7258, doi:10.1073/pnas.0611506104.

81. Luthra, P.; Ramanan, P.; Mire, C.E.; Weisend, C.; Tsuda, Y.; Yen, B.; Liu, G.; Leung, D.W.; Geisbert, T.W.; Ebihara, H.; et al. Mutual Antagonism between the Ebola Virus VP35 Protein and the RIG-I Activator PACT Determines Infection Outcome. *Cell Host Microbe* **2013**, *14*, doi:10.1016/j.chom.2013.06.010.

82. Gack, M.U.; Albrecht, R.A.; Urano, T.; Inn, K.-S.; Huang, I.-C.; Carnero, E.; Farzan, M.; Inoue, S.; Jung, J.U.; García-Sastre, A. Influenza A Virus NS1 Targets the Ubiquitin Ligase TRIM25 to Evade Recognition by RIG-I. *Cell Host Microbe* **2009**, *5*, 439–449, doi:10.1016/j.chom.2009.04.006.

83. Ling, Z.; Tran, K.C.; Teng, M.N. Human Respiratory Syncytial Virus Nonstructural Protein NS2 Antagonizes the Activation of Beta Interferon Transcription by Interacting with RIG-I. *J. Virol.* **2009**, *83*, 3734–3742, doi:10.1128/JVI.02434-08.

84. Yoboua, F.; Martel, A.; Duval, A.; Mukawera, E.; Grandvaux, N. Respiratory Syncytial Virus-Mediated NF-KB P65 Phosphorylation at Serine 536 Is Dependent on RIG-I, TRAF6, and IKKβ. *J. Virol.* **2010**, *84*, 7267–7277, doi:10.1128/JVI.00142-10.

85. Roberts, N.J.; Prill, A.H.; Mann, T.N. Interleukin 1 and Interleukin 1 Inhibitor Production by Human Macrophages Exposed to Influenza Virus or Respiratory Syncytial Virus. Respiratory Syncytial Virus Is a Potent Inducer of Inhibitor Activity. *J. Exp. Med.* **1986**, *163*, 511–519, doi:10.1084/jem.163.3.511.

86. National Institute of Allergy and Infectious Diseases (NIAID) *Phase I Placebo-Controlled Study of the Infectivity, Safety and Immunogenicity of a Single Dose of a Recombinant Live-Attenuated Respiratory Syncytial Virus Vaccine, RSV 6120/ΔNS2/1030s, Lot RSV#012A, Delivered as Nose Drops to RSV-Seropositive Children 12 to 59 Months of Age and RSV-Seronegative Infants and Children 6 to 24 Months of Age*; clinicaltrials.gov, 2020;

87. Connors, M.; Collins, P.L.; Firestone, C.Y.; Murphy, B.R. Respiratory Syncytial Virus (RSV) F, G, M2 (22K), and N Proteins Each Induce Resistance to RSV Challenge, but Resistance Induced by M2 and N Proteins Is Relatively Short-Lived. *J Virol* **1991**, *65*, 1634–7.

88. Stott, E.J.; Taylor, G.; Ball, L.A.; Anderson, K.; Young, K.K.; King, A.M.; Wertz, G.W. Immune and Histopathological Responses in Animals Vaccinated with Recombinant Vaccinia Viruses That Express Individual Genes of Human Respiratory Syncytial Virus. *J. Virol.* **1987**, *61*, 3855–3861.

89. Graham, B.S.; Modjarrad, K.; McLellan, J.S. Novel Antigens for RSV Vaccines. *Curr. Opin. Immunol.* **2015**, *35*, 30–38, doi:10.1016/j.coi.2015.04.005.

90. Collins, P.L.; Huang, Y.T.; Wertz, G.W. Nucleotide Sequence of the Gene Encoding the Fusion (F) Glycoprotein of Human Respiratory Syncytial Virus. *Proc Natl Acad Sci U A* **1984**, *81*, 7683–7.

91. Zimmer, G.; Budz, L.; Herrler, G. Proteolytic Activation of Respiratory Syncytial Virus Fusion Protein CLEAVAGE AT TWO FURIN CONSENSUS SEQUENCES. *J. Biol. Chem.* **2001**, *276*, 31642–31650, doi:10.1074/jbc.M102633200.

92. González-Reyes, L.; Ruiz-Argüello, M.B.; García-Barreno, B.; Calder, L.; López, J.A.; Albar, J.P.; Skehel, J.J.; Wiley, D.C.; Melero, J.A. Cleavage of the Human Respiratory Syncytial Virus Fusion Protein at Two Distinct Sites Is Required for Activation of Membrane Fusion. *Proc. Natl. Acad. Sci.* **2001**, *98*, 9859–9864, doi:10.1073/pnas.151098198.

93. Gruber, C.; Levine, S. Respiratory Syncytial Virus Polypeptides. III. The Envelope-Associated Proteins. *J. Gen. Virol.* **1983**, *64 (Pt 4)*, 825–832, doi:10.1099/0022-1317-64-4-825.

94. Day, N.D.; Branigan, P.J.; Liu, C.; Gutshall, L.L.; Luo, J.; Melero, J.A.; Sarisky, R.T.; Del Vecchio, A.M. Contribution of Cysteine Residues in the Extracellular Domain of the F Protein of Human Respiratory Syncytial Virus to Its Function. *Virol. J.* **2006**, *3*, 34, doi:10.1186/1743-422X-3-34.

95. Zhao, X.; Singh, M.; Malashkevich, V.N.; Kim, P.S. Structural Characterization of the Human Respiratory Syncytial Virus Fusion Protein Core. *Proc. Natl. Acad. Sci.* **2000**, *97*, 14172–14177, doi:10.1073/pnas.260499197.

96. Levine, S. Polypeptides of Respiratory Syncytial Virus. *J. Virol.* **1977**, *21*, 427–431, doi:10.1128/JVI.21.1.427-431.1977.

97. Hendricks, D.A.; Baradaran, K.; McIntosh, K.; Patterson, J.L. Appearance of a Soluble Form of the G Protein of Respiratory Syncytial Virus in Fluids of Infected Cells. *J. Gen. Virol.* **1987**, *68 (Pt 6)*, 1705–1714, doi:10.1099/0022-1317-68-6-1705.

98. Hendricks, D.A.; McIntosh, K.; Patterson, J.L. Further Characterization of the Soluble Form of the G Glycoprotein of Respiratory Syncytial Virus. *J. Virol.* **1988**, *62*, 2228–2233.

99. Roberts, S.R.; Lichtenstein, D.; Ball, L.A.; Wertz, G.W. The Membrane-Associated and Secreted Forms of the Respiratory Syncytial Virus Attachment Glycoprotein G Are Synthesized from Alternative Initiation Codons. *J. Virol.* **1994**, *68*, 4538–4546.

100. Gorman, J.J.; Ferguson, B.L.; Speelman, D.; Mills, J. Determination of the Disulfide Bond Arrangement of Human Respiratory Syncytial Virus Attachment

(G) Protein by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Protein Sci. Publ. Protein Soc.* **1997**, *6*, 1308–1315.

101. Olmsted, R.A.; Elango, N.; Prince, G.A.; Murphy, B.R.; Johnson, P.R.; Moss, B.; Chanock, R.M.; Collins, P.L. Expression of the F Glycoprotein of Respiratory Syncytial Virus by a Recombinant Vaccinia Virus: Comparison of the Individual Contributions of the F and G Glycoproteins to Host Immunity. *Proc. Natl. Acad. Sci.* **1986**, *83*, 7462–7466, doi:10.1073/pnas.83.19.7462.

102. Tripp, R.A.; Jones, L.P.; Haynes, L.M.; Zheng, H.; Murphy, P.M.; Anderson, L.J. CX3C Chemokine Mimicry by Respiratory Syncytial Virus G Glycoprotein. *Nat. Immunol.* **2001**, *2*, 732–738, doi:10.1038/90675.

103. Chirkova, T.; Lin, S.; Oomens, A.G.P.; Gaston, K.A.; Boyoglu-Barnum, S.; Meng, J.; Stobart, C.C.; Cotton, C.U.; Hartert, T.V.; Moore, M.L.; et al. CX3CR1 Is an Important Surface Molecule for Respiratory Syncytial Virus Infection in Human Airway Epithelial Cells. *J. Gen. Virol.* **2015**, *96*, 2543–2556, doi:10.1099/vir.0.000218.

104. Johnson, S.M.; McNally, B.A.; Ioannidis, I.; Flano, E.; Teng, M.N.; Oomens, A.G.; Walsh, E.E.; Peeples, M.E. Respiratory Syncytial Virus Uses CX3CR1 as a Receptor on Primary Human Airway Epithelial Cultures. *PLOS Pathog.* **2015**, *11*, e1005318, doi:10.1371/journal.ppat.1005318.

105. Karron, R.A.; Buonagurio, D.A.; Georgiu, A.F.; Whitehead, S.S.; Adamus, J.E.; Clements-Mann, M.L.; Harris, D.O.; Randolph, V.B.; Udem, S.A.; Murphy, B.R.; et al. Respiratory Syncytial Virus (RSV) SH and G Proteins Are Not Essential for

Viral Replication in Vitro: Clinical Evaluation and Molecular Characterization of a Cold-Passaged, Attenuated RSV Subgroup B Mutant. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 13961–13966, doi:10.1073/pnas.94.25.13961.

106. Collins, P.L.; Mottet, G. Membrane Orientation and Oligomerization of the Small Hydrophobic Protein of Human Respiratory Syncytial Virus. *J Gen Virol* **1993**, *74 (Pt 7)*, 1445–50, doi:10.1099/0022-1317-74-7-1445.

107. Gan, S.-W.; Tan, E.; Lin, X.; Yu, D.; Wang, J.; Tan, G.M.-Y.; Vararattanavech, A.; Yeo, C.Y.; Soon, C.H.; Soong, T.W.; et al. The Small Hydrophobic Protein of the Human Respiratory Syncytial Virus Forms Pentameric Ion Channels. *J. Biol. Chem.* **2012**, *287*, 24671–24689, doi:10.1074/jbc.M111.332791.

108. Gonzalez, M.E.; Carrasco, L. Viroporins. *FEBS Lett.* **2003**, *552*, 28–34, doi:10.1016/S0014-5793(03)00780-4.

109. Bukreyev, A.; Whitehead, S.S.; Murphy, B.R.; Collins, P.L. Recombinant Respiratory Syncytial Virus from Which the Entire SH Gene Has Been Deleted Grows Efficiently in Cell Culture and Exhibits Site-Specific Attenuation in the Respiratory Tract of the Mouse. *J Virol* **1997**, *71*, 8973–82.

110. Whitehead, S.S.; Bukreyev, A.; Teng, M.N.; Firestone, C.-Y.; Claire, M.S.; Elkins, W.R.; Collins, P.L.; Murphy, B.R. Recombinant Respiratory Syncytial Virus Bearing a Deletion of Either the NS2 or SH Gene Is Attenuated in Chimpanzees. *J. Virol.* **1999**, *73*, 3438–3442.

111. Carter, S.D.; Dent, K.C.; Atkins, E.; Foster, T.L.; Verow, M.; Gorny, P.; Harris, M.; Hiscox, J.A.; Ranson, N.A.; Griffin, S.; et al. Direct Visualization of the Small

Hydrophobic Protein of Human Respiratory Syncytial Virus Reveals the Structural Basis for Membrane Permeability. *FEBS Lett.* **2010**, *584*, 2786–2790, doi:10.1016/j.febslet.2010.05.006.

112. Olmsted, R.A.; Collins, P.L. The 1A Protein of Respiratory Syncytial Virus Is an Integral Membrane Protein Present as Multiple, Structurally Distinct Species. *J. Virol.* **1989**, *63*, 2019–2029.

113. Ghildyal, R.; Mills, J.; Murray, M.; Vardaxis, N.; Meanger, J. Respiratory Syncytial Virus Matrix Protein Associates with Nucleocapsids in Infected Cells. *J. Gen. Virol.* **2002**, *83*, 753–757, doi:10.1099/0022-1317-83-4-753.

114. Ghildyal, R.; Baulch-Brown, C.; Mills, J.; Meanger, J. The Matrix Protein of Human Respiratory Syncytial Virus Localises to the Nucleus of Infected Cells and Inhibits Transcription. *Arch. Virol.* **2003**, *148*, 1419–1429, doi:10.1007/s00705-003-0112-y.

115. Ghildyal, R.; Ho, A.; Jans, D.A. Central Role of the Respiratory Syncytial Virus Matrix Protein in Infection. *FEMS Microbiol. Rev.* **2006**, *30*, 692–705, doi:10.1111/j.1574-6976.2006.00025.x.

116. Shahriari, S.; Wei, K.; Ghildyal, R. Respiratory Syncytial Virus Matrix (M) Protein Interacts with Actin In Vitro and in Cell Culture. *Viruses* **2018**, *10*, doi:10.3390/v10100535.

117. Gould, P.S.; Easton, A.J. Coupled Translation of the Second Open Reading Frame of M2 MRNA Is Sequence Dependent and Differs Significantly within the

Subfamily Pneumovirinae. *J. Virol.* **2007**, *81*, 8488–8496, doi:10.1128/JVI.00457-07.

118. Ahmadian, G.; Randhawa, J.S.; Easton, A.J. Expression of the ORF-2 Protein of the Human Respiratory Syncytial Virus M2 Gene Is Initiated by a Ribosomal Termination-Dependent Reinitiation Mechanism. *EMBO J.* **2000**, *19*, 2681–2689, doi:10.1093/emboj/19.11.2681.
119. Tran, T.-L.; Castagné, N.; Dubosclard, V.; Noinville, S.; Koch, E.; Moudjou, M.; Henry, C.; Bernard, J.; Yeo, R.P.; Eléouët, J.-F. The Respiratory Syncytial Virus M2-1 Protein Forms Tetramers and Interacts with RNA and P in a Competitive Manner. *J. Virol.* **2009**, *83*, 6363–6374, doi:10.1128/JVI.00335-09.
120. Mason, S.W.; Aberg, E.; Lawetz, C.; DeLong, R.; Whitehead, P.; Liuzzi, M. Interaction between Human Respiratory Syncytial Virus (RSV) M2-1 and P Proteins Is Required for Reconstitution of M2-1-Dependent RSV Minigenome Activity. *J. Virol.* **2003**, *77*, 10670–10676, doi:10.1128/JVI.77.19.10670-10676.2003.
121. Hardy, R.W.; Wertz, G.W. The Cys3-His1 Motif of the Respiratory Syncytial Virus M2-1 Protein Is Essential for Protein Function. *J. Virol.* **2000**, *74*, 5880–5885, doi:10.1128/JVI.74.13.5880-5885.2000.
122. Asenjo, A.; Calvo, E.; Villanueva, N. Phosphorylation of Human Respiratory Syncytial Virus P Protein at Threonine 108 Controls Its Interaction with the M2-1 Protein in the Viral RNA Polymerase Complex. *J. Gen. Virol.* **2006**, *87*, 3637–3642, doi:10.1099/vir.0.82165-0.

123. Rincheval, V.; Lelek, M.; Gault, E.; Bouillier, C.; Sitterlin, D.; Blouquit-Laye, S.; Galloux, M.; Zimmer, C.; Eleouet, J.-F.; Rameix-Welti, M.-A. Functional Organization of Cytoplasmic Inclusion Bodies in Cells Infected by Respiratory Syncytial Virus. *Nat. Commun.* **2017**, *8*, 563, doi:10.1038/s41467-017-00655-9.

124. Collins, P.L.; Hill, M.G.; Camargo, E.; Grosfeld, H.; Chanock, R.M.; Murphy, B.R. Production of Infectious Human Respiratory Syncytial Virus from Cloned CDNA Confirms an Essential Role for the Transcription Elongation Factor from the 5' Proximal Open Reading Frame of the M2 MRNA in Gene Expression and Provides a Capability for Vaccine Development. *Proc. Natl. Acad. Sci.* **1995**, *92*, 11563–11567, doi:10.1073/pnas.92.25.11563.

125. Collins, P.L.; Hill, M.G.; Cristina, J.; Grosfeld, H. Transcription Elongation Factor of Respiratory Syncytial Virus, a Nonsegmented Negative-Strand RNA Virus. *Proc Natl Acad Sci U A* **1996**, *93*, 81–5.

126. Fearns, R.; Collins, P.L. Role of the M2-1 Transcription Antitermination Protein of Respiratory Syncytial Virus in Sequential Transcription. *J Virol* **1999**, *73*, 5852–64.

127. Bermingham, A.; Collins, P.L. The M2-2 Protein of Human Respiratory Syncytial Virus Is a Regulatory Factor Involved in the Balance between RNA Replication and Transcription. *Proc Natl Acad Sci U A* **1999**, *96*, 11259–64.

128. Cheng, X.; Park, H.; Zhou, H.; Jin, H. Overexpression of the M2-2 Protein of Respiratory Syncytial Virus Inhibits Viral Replication. *J. Virol.* **2005**, *79*, 13943–13952, doi:10.1128/JVI.79.22.13943-13952.2005.

129. Tawar, R.G.; Stéphane Duquerroy; Vonrhein, C.; Varela, P.F.; Damier-Piolle, L.; Castagné, N.; MacLellan, K.; Bedouelle, H.; Bricogne, G.; Bhella, D.; et al. Crystal Structure of a Nucleocapsid-Like Nucleoprotein-RNA Complex of Respiratory Syncytial Virus. *Science* **2009**, *326*, 1279–1283, doi:10.1126/science.1177634.

130. Castagne, N. Biochemical Characterization of the Respiratory Syncytial Virus P-P and P-N Protein Complexes and Localization of the P Protein Oligomerization Domain. *J. Gen. Virol.* **2004**, *85*, 1643–1653, doi:10.1099/vir.0.79830-0.

131. Groskreutz, D.J.; Babor, E.C.; Monick, M.M.; Varga, S.M.; Hunninghake, G.W. Respiratory Syncytial Virus Limits α Subunit of Eukaryotic Translation Initiation Factor 2 (EIF2α) Phosphorylation to Maintain Translation and Viral Replication. *J. Biol. Chem.* **2010**, *285*, 24023–24031, doi:10.1074/jbc.M109.077321.

132. García-Barreno, B.; Delgado, T.; Melero, J.A. Identification of Protein Regions Involved in the Interaction of Human Respiratory Syncytial Virus Phosphoprotein and Nucleoprotein: Significance for Nucleocapsid Assembly and Formation of Cytoplasmic Inclusions. *J. Virol.* **1996**, *70*, 801–808.

133. Khattar, S.K.; Yunus, A.S.; Collins, P.L.; Samal, S.K. Deletion and Substitution Analysis Defines Regions and Residues within the Phosphoprotein of Bovine Respiratory Syncytial Virus That Affect Transcription, RNA Replication, and Interaction with the Nucleoprotein. *Virology* **2001**, *285*, 253–269, doi:10.1006/viro.2001.0960.

134. Asenjo, A.; Rodríguez, L.; Villanueva, N. Determination of Phosphorylated Residues from Human Respiratory Syncytial Virus P Protein That Are Dynamically Dephosphorylated by Cellular Phosphatases: A Possible Role for Serine 54. *J. Gen. Virol.* **2005**, *86*, 1109–1120, doi:10.1099/vir.0.80692-0.

135. Hardy, R.W.; Wertz, G.W. The Product of the Respiratory Syncytial Virus M2 Gene ORF1 Enhances Readthrough of Intergenic Junctions during Viral Transcription. *J. Virol.* **1998**, *72*, 520–526.

136. Curran, J.; Marq, J.B.; Kolakofsky, D. An N-Terminal Domain of the Sendai Paramyxovirus P Protein Acts as a Chaperone for the NP Protein during the Nascent Chain Assembly Step of Genome Replication. *J. Virol.* **1995**, *69*, 849–855.

137. Liuzzi, M.; Mason, S.W.; Cartier, M.; Lawetz, C.; McCollum, R.S.; Dansereau, N.; Bolger, G.; Lapeyre, N.; Gaudette, Y.; Lagace, L.; et al. Inhibitors of Respiratory Syncytial Virus Replication Target Cotranscriptional MRNA Guanylylation by Viral RNA-Dependent RNA Polymerase. *J. Virol.* **2005**, *79*, 13105–13115, doi:10.1128/JVI.79.20.13105-13115.2005.

138. Cao, D.; Gao, Y.; Roesler, C.; Rice, S.; D'Cunha, P.; Zhuang, L.; Slack, J.; Domke, M.; Antonova, A.; Romanelli, S.; et al. Cryo-EM Structure of the Respiratory Syncytial Virus RNA Polymerase. *Nat. Commun.* **2020**, *11*, 1–9, doi:10.1038/s41467-019-14246-3.

139. Yakhini, Z.; Jurisica, I. Cancer Computational Biology. *BMC Bioinformatics* **2011**, *12*, 120, doi:10.1186/1471-2105-12-120.

140. Hogeweg, P. The Roots of Bioinformatics in Theoretical Biology. *PLoS Comput. Biol.* **2011**, *7*, doi:10.1371/journal.pcbi.1002021.

141. Zhao, K.; Henderson, E.; Bullard, K.; Oberste, M.S.; Burns, C.C.; Jorba, J. PoSE: Visualization of Patterns of Sequence Evolution Using PAML and MATLAB. *BMC Bioinformatics* **2018**, *19*, 364, doi:10.1186/s12859-018-2335-7.

142. Dehouck, Y.; Kwasigroch, J.M.; Gilis, D.; Rooman, M. PoPMuSiC 2.1: A Web Server for the Estimation of Protein Stability Changes upon Mutation and Sequence Optimality. *BMC Bioinformatics* **2011**, *12*, 151, doi:10.1186/1471-2105-12-151.

143. Dong, Q.; Menon, R.; Omenn, G.S.; Zhang, Y. Structural Bioinformatics Inspection of NeXtProt PE5 Proteins in the Human Proteome. *J Proteome Res* **2015**, *14*, 3750–61, doi:10.1021/acs.jproteome.5b00516.

144. Collins, P.L.; Murphy, B.R. New Generation Live Vaccines against Human Respiratory Syncytial Virus Designed by Reverse Genetics. *Proc Am Thorac Soc* **2005**, *2*, 166–73, doi:10.1513/pats.200501-011AW.

145. Karron, R.A.; Buchholz, U.J.; Collins, P.L. Live-Attenuated Respiratory Syncytial Virus Vaccines. *Curr Top Microbiol Immunol* **2013**, *372*, 259–84, doi:10.1007/978-3-642-38919-1_13.

146. Roberts, J.N.; Graham, B.S.; Karron, R.A.; Munoz, F.M.; Falsey, A.R.; Anderson, L.J.; Marshall, V.; Kim, S.; Beeler, J.A. Challenges and Opportunities in RSV Vaccine Development: Meeting Report from FDA/NIH Workshop. *Vaccine* **2016**, *34*, 4843–4849, doi:10.1016/j.vaccine.2016.07.057.

147. Dudas, R.A.; Karron, R.A. Respiratory Syncytial Virus Vaccines. *Clin. Microbiol. Rev.* **1998**, *11*, 430–439.

148. Giersing, B.K.; Vekemans, J.; Nava, S.; Kaslow, D.C.; Moorthy, V. Report from the World Health Organization's Third Product Development for Vaccines Advisory Committee (PDVAC) Meeting, Geneva, 8–10th June 2016. *Vaccine* **2017**, S0264410X16312051, doi:10.1016/j.vaccine.2016.10.090.

149. Modjarrad, K.; Giersing, B.; Kaslow, D.C.; Smith, P.G.; Moorthy, V.S. WHO Consultation on Respiratory Syncytial Virus Vaccine Development Report from a World Health Organization Meeting Held on 23–24 March 2015. *Vaccine* **2016**, *34*, 190–197, doi:10.1016/j.vaccine.2015.05.093.

150. Linfield, D.T.; Rezaee, F. RSV: Available Prophylactic Options and Vaccines in Clinical Trials. *Burd. Respir. Syncytial Virus Infect. Young* **2019**, doi:10.5772/intechopen.84851.

151. Martínez-Castilla, L.P.; Rodríguez-Sotres, R. A Score of the Ability of a Three-Dimensional Protein Model to Retrieve Its Own Sequence as a Quantitative Measure of Its Quality and Appropriateness. *PLOS ONE* **2010**, *5*, e12483, doi:10.1371/journal.pone.0012483.

152. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303, doi:10.1093/nar/gky427.

153. Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Cassarino, T.G.; Bertoni, M.; Bordoli, L.; et al. SWISS-MODEL: Modelling Protein Tertiary and Quaternary Structure Using Evolutionary Information. *Nucleic Acids Res.* **2014**, *42*, W252–W258, doi:10.1093/nar/gku340.

154. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S.; et al. CDD/SPARCLE: The Conserved Domain Database in 2020. *Nucleic Acids Res.* **2020**, *48*, D265–D268, doi:10.1093/nar/gkz991.

155. Altschul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Res* **1997**, *25*, 3389–402.

156. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410, doi:10.1016/S0022-2836(05)80360-2.

157. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and Applications. *BMC Bioinformatics* **2009**, *10*, 421, doi:10.1186/1471-2105-10-421.

158. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X Version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948, doi:10.1093/bioinformatics/btm404.

159. Thompson, J.D.; Gibson, T.J.; Plewniak, F.; Jeanmougin, F.; Higgins, D.G. The CLUSTAL_X Windows Interface: Flexible Strategies for Multiple Sequence Alignment Aided by Quality Analysis Tools. *Nucleic Acids Res.* **1997**, *25*, 4876–4882, doi:10.1093/nar/25.24.4876.

160. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680, doi:10.1093/nar/22.22.4673.

161. Olson, S.A. MacVector: An Integrated Sequence Analysis Program for the Macintosh. *Comput. Anal. Seq. Data* **1994**, 195–201, doi:10.1385/0-89603-276-0:195.

162. Rastogi, P.A. MacVector. In *Bioinformatics Methods and Protocols*; Misener, S., Krawetz, S.A., Eds.; Methods in Molecular Biology™; Humana Press: Totowa, NJ, 1999; pp. 47–69 ISBN 978-1-59259-192-3.

163. Olson, S.A. Mac Vector: Protein Analysis. *Comput. Anal. Seq. Data* **1994**, 237–246, doi:10.1385/0-89603-276-0:237.

164. Thompson, J.D.; Gibson, T.J.; Higgins, D.G. Multiple Sequence Alignment Using ClustalW and ClustalX. *Curr. Protoc. Bioinforma.* **2003**, *00*, 2.3.1-2.3.22, doi:10.1002/0471250953.bi0203s00.

165. Sigrist, C.J.A.; de Castro, E.; Cerutti, L.; Cuche, B.A.; Hulo, N.; Bridge, A.; Bougueleret, L.; Xenarios, I. New and Continuing Developments at PROSITE. *Nucleic Acids Res.* **2013**, *41*, D344–D347, doi:10.1093/nar/gks1067.

166. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: Protein Structure and Function Prediction. *Nat. Methods* **2015**, *12*, 7–8, doi:10.1038/nmeth.3213.

167. Xu, D.; Zhang, J.; Roy, A.; Zhang, Y. Automated Protein Structure Modeling in CASP9 by I-TASSER Pipeline Combined with QUARK-Based Ab Initio Folding and FG-MD-Based Structure Refinement. *Proteins Struct. Funct. Bioinforma.* **2011**, *79*, 147–160, doi:10.1002/prot.23111.

168. Zhang, Y. I-TASSER Server for Protein 3D Structure Prediction. *BMC Bioinformatics* **2008**, *9*, 40, doi:10.1186/1471-2105-9-40.

169. Berman, H.; Henrick, K.; Nakamura, H.; Markley, J.L. The Worldwide Protein Data Bank (WwPDB): Ensuring a Single, Uniform Archive of PDB Data. *Nucleic Acids Res* **2007**, *35*, D301-3, doi:10.1093/nar/gkl971.

170. Berman, H.; Henrick, K.; Nakamura, H. Announcing the Worldwide Protein Data Bank. *Nat. Struct. Mol. Biol.* **2003**, *10*, 980–980, doi:10.1038/nsb1203-980.

171. Bernstein, F.C.; Koetzle, T.F.; Williams, G.J.B.; Meyer, E.F.; Brice, M.D.; Rodgers, J.R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank: A Computer-Based Archival File for Macromolecular Structures. *J. Mol. Biol.* **1977**, *112*, 535–542, doi:10.1016/S0022-2836(77)80200-3.

172. Zhang, W.; Yang, J.; He, B.; Walker, S.E.; Zhang, H.; Govindarajoo, B.; Virtanen, J.; Xue, Z.; Shen, H.-B.; Zhang, Y. Integration of QUARK and I-TASSER for Ab Initio Protein Structure Prediction in CASP11. *Proteins Struct. Funct. Bioinforma.* **2016**, *84*, 76–86, doi:10.1002/prot.24930.

173. Roy, A.; Kucukural, A.; Zhang, Y. I-TASSER: A Unified Platform for Automated Protein Structure and Function Prediction. *Nat. Protoc.* **2010**, *5*, 725–738, doi:10.1038/nprot.2010.5.

174. Zhang, Y.; Skolnick, J. SPICKER: A Clustering Approach to Identify near-Native Protein Folds. *J. Comput. Chem.* **2004**, *25*, 865–871, doi:10.1002/jcc.20011.

175. Gonnelli, G.; Rooman, M.; Dehouck, Y. Structure-Based Mutant Stability Predictions on Proteins of Unknown Structure. *J. Biotechnol.* **2012**, *161*, 287–293, doi:10.1016/j.jbiotec.2012.06.020.

176. Cozzetto, D.; Kryshtafovych, A.; Ceriani, M.; Tramontano, A. Assessment of Predictions in the Model Quality Assessment Category. *Proteins Struct. Funct. Bioinforma.* **2007**, *69*, 175–183, doi:10.1002/prot.21669.

177. Benkert, P.; Biasini, M.; Schwede, T. Toward the Estimation of the Absolute Quality of Individual Protein Structure Models. *Bioinformatics* **2011**, *27*, 343–350, doi:10.1093/bioinformatics/btq662.

178. Benkert, P.; Künzli, M.; Schwede, T. QMEAN Server for Protein Model Quality Estimation. *Nucleic Acids Res.* **2009**, *37*, W510-514, doi:10.1093/nar/gkp322.

179. Mariani, V.; Biasini, M.; Barbato, A.; Schwede, T. LDDT: A Local Superposition-Free Score for Comparing Protein Structures and Models Using Distance Difference Tests. *Bioinformatics* **2013**, *29*, 2722–2728, doi:10.1093/bioinformatics/btt473.

180. Schwede, T.; Sali, A.; Honig, B.; Levitt, M.; Berman, H.M.; Jones, D.; Brenner, S.E.; Burley, S.K.; Das, R.; Dokholyan, N.V.; et al. Outcome of a Workshop on

Applications of Protein Models in Biomedical Research. *Structure* **2009**, *17*, 151–159, doi:10.1016/j.str.2008.12.014.

181. Boeckmann, B.; Bairoch, A.; Apweiler, R.; Blatter, M.-C.; Estreicher, A.; Gasteiger, E.; Martin, M.J.; Michoud, K.; O'Donovan, C.; Phan, I.; et al. The SWISS-PROT Protein Knowledgebase and Its Supplement TrEMBL in 2003. *Nucleic Acids Res.* **2003**, *31*, 365–370, doi:10.1093/nar/gkg095.

182. Schwede, T. SWISS-MODEL: An Automated Protein Homology-Modeling Server. *Nucleic Acids Res.* **2003**, *31*, 3381–3385, doi:10.1093/nar/gkg520.

183. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612, doi:10.1002/jcc.20084.

184. Masso, M.; Vaisman, I.I. AUTO-MUTE: Web-Based Tools for Predicting Stability Changes in Proteins Due to Single Amino Acid Replacements. *Protein Eng. Des. Sel.* **2010**, *23*, 683–687, doi:10.1093/protein/gzq042.

185. Masso, M.; Vaisman, I.I. Accurate Prediction of Stability Changes in Protein Mutants by Combining Machine Learning with Structure Based Computational Mutagenesis. *Bioinformatics* **2008**, *24*, 2002–2009, doi:10.1093/bioinformatics/btn353.

186. Le Nouen, C.; Brock, L.G.; Luongo, C.; McCarty, T.; Yang, L.; Mehedi, M.; Wimmer, E.; Mueller, S.; Collins, P.L.; Buchholz, U.J.; et al. Attenuation of

Human Respiratory Syncytial Virus by Genome-Scale Codon-Pair Deoptimization. *Proc Natl Acad Sci U A* **2014**, *111*, 13169–74, doi:10.1073/pnas.1411290111.

187. Le Nouen, C.; McCarty, T.; Brown, M.; Smith, M.L.; Lleras, R.; Dolan, M.A.; Mehedi, M.; Yang, L.; Luongo, C.; Liang, B.; et al. Genetic Stability of Genome-Scale Deoptimized RNA Virus Vaccine Candidates under Selective Pressure. *Proc Natl Acad Sci U A* **2017**, *114*, E386–E395, doi:10.1073/pnas.1619242114.

188. Coleman, J.R.; Papamichail, D.; Skiena, S.; Futcher, B.; Wimmer, E.; Mueller, S. Virus Attenuation by Genome-Scale Changes in Codon Pair Bias. *Science* **2008**, *320*, 1784–1787, doi:10.1126/science.1155761.

189. Graham, B.S.; Gilman, M.S.A.; McLellan, J.S. Structure-Based Vaccine Antigen Design. *Annu. Rev. Med.* **2019**, *70*, 91–104, doi:10.1146/annurev-med-121217-094234.

190. Dormitzer, P.R.; Ulmer, J.B.; Rappuoli, R. Structure-Based Antigen Design: A Strategy for next Generation Vaccines. *Trends Biotechnol.* **2008**, *26*, 659–667, doi:10.1016/j.tibtech.2008.08.002.

191. Masso, M.; Vaisman Structure-Based Prediction of Protein Activity Changes: Assessing the Impact of Single Residue Replacements. *Conf Proc IEEE Eng Med Biol Soc* **2011**, *2011*, 3221–4, doi:10.1109/IEMBS.2011.6090876.

192. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32, doi:10.1023/A:1010933404324.

193. Tjong, H.; Qin, S.; Zhou, H.-X. PI2PE: Protein Interface/Interior Prediction Engine. *Nucleic Acids Res.* **2007**, *35*, W357-362, doi:10.1093/nar/gkm231.

194. Caldararu, O.; Mehra, R.; Blundell, T.L.; Kepp, K.P. Systematic Investigation of the Data Set Dependency of Protein Stability Predictors. *J. Chem. Inf. Model.* **2020**, *60*, 4772–4784, doi:10.1021/acs.jcim.0c00591.

195. Pucci, F.; Bernaerts, K.V.; Kwasigroch, J.M.; Rooman, M. Quantification of Biases in Predictions of Protein Stability Changes upon Mutations. *Bioinforma. Oxf. Engl.* **2018**, *34*, 3659–3665, doi:10.1093/bioinformatics/bty348.

196. Conzelmann, K.K. Reverse Genetics of Mononegavirales. In *Biology of Negative Strand RNA Viruses: The Power of Reverse Genetics*; Kawaoka, Y., Ed.; Current Topics in Microbiology and Immunology; Springer: Berlin, Heidelberg, 2004; pp. 1–41 ISBN 978-3-662-06099-5.

197. Buchholz, U.J.; Finke, S.; Conzelmann, K.K. Generation of Bovine Respiratory Syncytial Virus (BRSV) from CDNA: BRSV NS2 Is Not Essential for Virus Replication in Tissue Culture, and the Human RSV Leader Region Acts as a Functional BRSV Genome Promoter. *J Virol* **1999**, *73*, 251–9.

198. Luongo, C.; Winter, C.C.; Collins, P.L.; Buchholz, U.J. Respiratory Syncytial Virus Modified by Deletions of the NS2 Gene and Amino Acid S1313 of the L Polymerase Protein Is a Temperature-Sensitive, Live-Attenuated Vaccine Candidate That Is Phenotypically Stable at Physiological Temperature. *J. Virol.* **2013**, *87*, 1985–1996, doi:10.1128/JVI.02769-12.

199. Jin, H.; Zhou, H.; Cheng, X.; Tang, R.; Munoz, M.; Nguyen, N. Recombinant Respiratory Syncytial Viruses with Deletions in the NS1, NS2, SH, and M2-2

Genes Are Attenuated in Vitro and in Vivo. *Virology* **2000**, *273*, 210–218, doi:10.1006/viro.2000.0393.

200. Teng, M.N.; Collins, P.L. Altered Growth Characteristics of Recombinant Respiratory Syncytial Viruses Which Do Not Produce NS2 Protein. *J. Virol.* **1999**, *73*, 466–473.

201. Jin, H.; Cheng, X.; Traina-Dorge, V.L.; Park, H.J.; Zhou, H.; Soike, K.; Kemble, G. Evaluation of Recombinant Respiratory Syncytial Virus Gene Deletion Mutants in African Green Monkeys for Their Potential as Live Attenuated Vaccine Candidates. *Vaccine* **2003**, *21*, 3647–3652, doi:10.1016/S0264-410X(03)00426-2.

202. Teng, M.N.; Spann, K.; Bakre, A.; Hiscox, J.; Tripp, R.A.; Wu, W. Human Respiratory Syncytial Virus Non-Structural Protein NS1 Modifies MiR-24 Expression via Transforming Growth Factor-β. *J. Gen. Virol.* **2015**, *96*, 3179–3191, doi:10.1099/jgv.0.000261.

203. Bakre, A.; Mitchell, P.; Coleman, J.K.; Jones, L.P.; Saavedra, G.; Teng, M.; Tompkins, S.M.; Tripp, R.A. Respiratory Syncytial Virus Modifies MicroRNAs Regulating Host Genes That Affect Virus Replication. *J. Gen. Virol.* **2012**, *93*, 2346–2356, doi:10.1099/vir.0.044255-0.

204. Swedan, S.; Andrews, J.; Majumdar, T.; Musiyenko, A.; Barik, S. Multiple Functional Domains and Complexes of the Two Nonstructural Proteins of Human Respiratory Syncytial Virus Contribute to Interferon Suppression and Cellular Location. *J. Virol.* **2011**, *85*, 10090–10100, doi:10.1128/JVI.00413-11.

205. Bitko, V.; Barik, S. An Endoplasmic Reticulum-Specific Stress-Activated Caspase (Caspase-12) Is Implicated in the Apoptosis of A549 Epithelial Cells by Respiratory Syncytial Virus. *J. Cell. Biochem.* **2001**, *80*, 441–454.

206. Goswami, R.; Majumdar, T.; Dhar, J.; Chattopadhyay, S.; Bandyopadhyay, S.K.; Verbovetskaya, V.; Sen, G.C.; Barik, S. Viral Degradasome Hijacks Mitochondria to Suppress Innate Immunity. *Cell Res.* **2013**, *23*, 1025–1042, doi:10.1038/cr.2013.98.

207. Lukacs, N.W.; Moore, M.L.; Rudd, B.D.; Berlin, A.A.; Collins, R.D.; Olson, S.J.; Ho, S.B.; Peebles, R.S. Differential Immune Responses and Pulmonary Pathophysiology Are Induced by Two Different Strains of Respiratory Syncytial Virus. *Am. J. Pathol.* **2006**, *169*, 977–986, doi:10.2353/ajpath.2006.051055.

208. Real-Time PCR (QPCR) - US Available online: https://www.thermofisher.com/us/en/home/life-science/pcr/real-time-pcr.html (accessed on 31 January 2020).

209. Gilman, M.S.A.; Castellanos, C.A.; Chen, M.; Ngwuta, J.O.; Goodwin, E.; Moin, S.M.; Mas, V.; Melero, J.A.; Wright, P.F.; Graham, B.S.; et al. Rapid Profiling of RSV Antibody Repertoires from the Memory B Cells of Naturally Infected Adult Donors. *Sci. Immunol.* **2016**, *1*, doi:10.1126/sciimmunol.aaj1879.

210. Goodwin, E.; Gilman, M.S.A.; Wrapp, D.; Chen, M.; Ngwuta, J.O.; Moin, S.M.; Bai, P.; Sivasubramanian, A.; Connor, R.I.; Wright, P.F.; et al. Infants Infected with Respiratory Syncytial Virus Generate Potent Neutralizing Antibodies That

Lack Somatic Hypermutation. *Immunity* **2018**, *48*, 339-349.e5, doi:10.1016/j.immuni.2018.01.005.

211. Whitehead, S.S.; Hill, M.G.; Firestone, C.Y.; Claire, M.S.; Elkins, W.R.; Murphy, B.R.; Collins, P.L. Replacement of the F and G Proteins of Respiratory Syncytial Virus (RSV) Subgroup A with Those of Subgroup B Generates Chimeric Live Attenuated RSV Subgroup B Vaccine Candidates. *J. Virol.* **1999**, *73*, 9773–9780.
212. Noton, S.L.; Tremaglio, C.Z.; Fearns, R. Killing Two Birds with One Stone: How the Respiratory Syncytial Virus Polymerase Initiates Transcription and Replication. *PLOS Pathog.* **2019**, *15*, e1007548, doi:10.1371/journal.ppat.1007548.
213. Sun, Y.; López, C. Preparation of Respiratory Syncytial Virus with High or Low Content of Defective Viral Particles and Their Purification from Viral Stocks. *BIO-Protoc.* **2016**, *6*, doi:10.21769/BioProtoc.1820.
214. Sun, Y.; López, C.B. The Innate Immune Response to RSV: Advances in Our Understanding of Critical Viral and Host Factors. *Vaccine* **2017**, *35*, 481–488, doi:10.1016/j.vaccine.2016.09.030.
215. Sun, Y.; Kim, E.J.; Felt, S.A.; Taylor, L.J.; Agarwal, D.; Grant, G.R.; López, C.B. A Specific Sequence in the Genome of Respiratory Syncytial Virus Regulates the Generation of Copy-Back Defective Viral Genomes. *PLOS Pathog.* **2019**, *15*, e1007707, doi:10.1371/journal.ppat.1007707.

216. Krempl, C.; Murphy, B.R.; Collins, P.L. Recombinant Respiratory Syncytial Virus with the G and F Genes Shifted to the Promoter-Proximal Positions. *J. Virol.* **2002**, *76*, 11931–11942, doi:10.1128/JVI.76.23.11931-11942.2002.

217. Mehedi, M.; McCarty, T.; Martin, S.E.; Nouën, C.L.; Buehler, E.; Chen, Y.-C.; Smelkinson, M.; Ganesan, S.; Fischer, E.R.; Brock, L.G.; et al. Actin-Related Protein 2 (ARP2) and Virus-Induced Filopodia Facilitate Human Respiratory Syncytial Virus Spread. *PLOS Pathog.* **2016**, *12*, e1006062, doi:10.1371/journal.ppat.1006062.

218. Ramaswamy, M.; Shi, L.; Varga, S.M.; Barik, S.; Behlke, M.A.; Look, D.C. Respiratory Syncytial Virus Nonstructural Protein 2 Specifically Inhibits Type I Interferon Signal Transduction. *Virology* **2006**, *344*, 328–339, doi:10.1016/j.virol.2005.09.009.

219. Swedan, S.; Musiyenko, A.; Barik, S. Respiratory Syncytial Virus Nonstructural Proteins Decrease Levels of Multiple Members of the Cellular Interferon Pathways. *J. Virol.* **2009**, *83*, 9682–9693, doi:10.1128/JVI.00715-09.

220. Tian, B.; Zhang, Y.; Luxon, B.A.; Garofalo, R.P.; Casola, A.; Sinha, M.; Brasier, A.R. Identification of NF-KappaB-Dependent Gene Networks in Respiratory Syncytial Virus-Infected Cells. *J. Virol.* **2002**, *76*, 6800–6814, doi:10.1128/jvi.76.13.6800-6814.2002.

221. Zhang, Y.; Luxon, B.A.; Casola, A.; Garofalo, R.P.; Jamaluddin, M.; Brasier, A.R. Expression of Respiratory Syncytial Virus-Induced Chemokine Gene Networks in

Lower Airway Epithelial Cells Revealed by CDNA Microarrays. *J. Virol.* **2001**, *75*, 9044–9058, doi:10.1128/JVI.75.19.9044-9058.2001.

222. Miller, A.L.; Bowlin, T.L.; Lukacs, N.W. Respiratory Syncytial Virus-Induced Chemokine Production: Linking Viral Replication to Chemokine Production in Vitro and in Vivo. *J. Infect. Dis.* **2004**, *189*, 1419–1430, doi:10.1086/382958.

223. Krishnan, S.; Halonen, M.; Welliver, R.C. Innate Immune Responses in Respiratory Syncytial Virus Infections. *Viral Immunol.* **2004**, *17*, 220–233, doi:10.1089/0882824041310612.

224. Sheeran, P.; Jafri, H.; Carubelli, C.; Saavedra, J.; Johnson, C.; Krisher, K.; Sánchez, P.J.; Ramilo, O. Elevated Cytokine Concentrations in the Nasopharyngeal and Tracheal Secretions of Children with Respiratory Syncytial Virus Disease. *Pediatr. Infect. Dis. J.* **1999**, *18*, 115–122.

225. Gotoh, B.; Komatsu, T.; Takeuchi, K.; Yokoo, J. Paramyxovirus Accessory Proteins as Interferon Antagonists. *Microbiol Immunol* **2001**, *45*, 787–800.

226. Weber-Gerlach, M.; Weber, F. To Conquer the Host, Influenza Virus Is Packing It In: Interferon-Antagonistic Strategies beyond NS1. *J. Virol.* **2016**, *90*, 8389–8394, doi:10.1128/JVI.00041-16.

227. Elliott, J.; Lynch, O.T.; Suessmuth, Y.; Qian, P.; Boyd, C.R.; Burrows, J.F.; Buick, R.; Stevenson, N.J.; Touzelet, O.; Gadina, M.; et al. Respiratory Syncytial Virus NS1 Protein Degrades STAT2 by Using the Elongin-Cullin E3 Ligase. *J Virol* **2007**, *81*, 3428–36, doi:10.1128/JVI.02303-06.

228. Straub, C.P.; Lau, W.-H.; Preston, F.M.; Headlam, M.J.; Gorman, J.J.; Collins, P.L.; Spann, K.M. Mutation of the Elongin C Binding Domain of Human Respiratory Syncytial Virus Non-Structural Protein 1 (NS1) Results in Degradation of NS1 and Attenuation of the Virus. *Virol. J.* **2011**, *8*, 252, doi:10.1186/1743-422X-8-252.

229. Tang, A.; Chen, Z.; Cox, K.S.; Su, H.-P.; Callahan, C.; Fridman, A.; Zhang, L.; Patel, S.B.; Cejas, P.J.; Swoyer, R.; et al. A Potent Broadly Neutralizing Human RSV Antibody Targets Conserved Site IV of the Fusion Glycoprotein. *Nat. Commun.* **2019**, *10*, doi:10.1038/s41467-019-12137-1.

230. Levitz, R.; Wattier, R.; Phillips, P.; Solomon, A.; Lawler, J.; Lazar, I.; Weibel, C.; Kahn, J.S. Induction of IL-6 and CCL5 (RANTES) in Human Respiratory Epithelial (A549) Cells by Clinical Isolates of Respiratory Syncytial Virus Is Strain Specific. *Virol. J.* **2012**, *9*, 190, doi:10.1186/1743-422X-9-190.

231. Elliott, M.B.; Tebbey, P.W.; Pryharski, K.S.; Scheuer, C.A.; Laughlin, T.S.; Hancock, G.E. Inhibition of Respiratory Syncytial Virus Infection with the CC Chemokine RANTES (CCL5). *J. Med. Virol.* **2004**, *73*, 300–308, doi:10.1002/jmv.20091.

232. Okabayashi, T.; Kojima, T.; Masaki, T.; Yokota, S.; Imaizumi, T.; Tsutsumi, H.; Himi, T.; Fujii, N.; Sawada, N. Type-III Interferon, Not Type-I, Is the Predominant Interferon Induced by Respiratory Viruses in Nasal Epithelial Cells. *Virus Res.* **2011**, *160*, 360–366, doi:10.1016/j.virusres.2011.07.011.

233. Lo, M.S.; Brazas, R.M.; Holtzman, M.J. Respiratory Syncytial Virus Nonstructural Proteins NS1 and NS2 Mediate Inhibition of Stat2 Expression and Alpha/Beta Interferon Responsiveness. *J. Virol.* **2005**, *79*, 9315–9319, doi:10.1128/JVI.79.14.9315-9319.2005.

234. Kotelkin, A.; Prikhod'ko, E.A.; Cohen, J.I.; Collins, P.L.; Bukreyev, A. Respiratory Syncytial Virus Infection Sensitizes Cells to Apoptosis Mediated by Tumor Necrosis Factor-Related Apoptosis-Inducing Ligand. *J. Virol.* **2003**, *77*, 9156–9172, doi:10.1128/JVI.77.17.9156-9172.2003.

235. Kroemer, G.; Galluzzi, L.; Vandenabeele, P.; Abrams, J.; Alnemri, E.S.; Baehrecke, E.H.; Blagosklonny, M.V.; El-Deiry, W.S.; Golstein, P.; Green, D.R.; et al. Classification of Cell Death: Recommendations of the Nomenclature Committee on Cell Death 2009. *Cell Death Differ.* **2009**, *16*, 3–11, doi:10.1038/cdd.2008.150.

236. Keshavarz, M.; Solaymani-Mohammadi, F.; Miri, S.M.; Ghaemi, A. Oncolytic Paramyxoviruses-Induced Autophagy; a Prudent Weapon for Cancer Therapy. *J. Biomed. Sci.* **2019**, *26*, 48, doi:10.1186/s12929-019-0542-9.

237. Caspase-Glo® 3/7 Assay System | Caspase 3 Activity Assay | Apoptosis Detection Available online: https://www.promega.com/products/cell-health-assays/apoptosis-assays/caspase_glo-3_7-assay-systems/ (accessed on 31 January 2020).

BIOGRAPHY

Thomas C. McCarty graduated from West Springfield High School, Springfield, Virginia, in 1982. He received his Bachelor of Science from University of Wisconsin-Madison in 1986. He received his Master of Science from Johns Hopkins University in 2000. Mr. McCarty is currently employed as a research scientist with the Laboratory of Infectious Diseases, RNA Virus Section in the National Institute of Allergy and Infectious Disease (NIAID), part of the National Institutes of Health (NIH).