AN INFLUENZA VIRUS MOLECULAR INFECTION MODEL AND DISCRETE EVENT, STOCHASTIC SIMULATION

by

RICHARD BURKE SQUIRES

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE AND ENGINEERING

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2008

ACKNOWLEDGEMENTS

I would like to thank all who contributed help and prayers to this endeavor. I must first that my wife Christine for the patience and wholehearted support! I also want to thank my committee members Dr Sajal Das and Professor Kalyan Basu for their advice, direction and support. Additionally, I would like to thank my boss, committee member and mentor Richard Scheuermann with whom my influenza virus work got its start as part of the BioHealthBase Bioinformatics Resource Center (BRC). I offer my thanks and appreciation for all those whom offered technical or scientific help or advice, namely Preetam Ghosh, Samik Ghosh and Eric Bortz. I must not forget to acknowledge my family and friends whom have offered numerous prayers for my project and me. Finally, I must thank God, whom I owe the greatest thanks for gifts given and for the ultimate inspiration for the project understanding that "all things happen for a reason" and that many "coincidences" occurred just in the right time and order to bring me to this project.

April 25, 2008

ABSTRACT

AN INFLUENZA VIRUS MOLECULAR INFECTION MODEL AND DISCRETE EVENT, STOCHASTIC SIMULATION

Richard Burke Squires, M.S.

The University of Texas at Arlington, 2008

Supervising Professor: Sajal Das

Influenza virus is responsible for the greatest pandemic in human history causing 20 – 40 million deaths worldwide during the 1918 flu season. In 1997 fears of a future pandemic arose with the discovery of a new strain of H5N1 avian influenza. In an effort to better understand the dynamics of influenza infection a new model of influenza virus infection at the molecular level has been developed. Comprising nineteen stages and five molecular types (mRNA, cRNA, vRNA, vRNP and proteins) the FluSim model consists of the major contributing pathogen factors to influenza infection. A discrete event, stochastic simulation based upon this model utilizing the SimJava framework has also been implemented. Simulation results are compared against experimental evidence as well as a separate Dizzy simulation implementation.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

ABSTRACT iv

LIST OF ILLUSTRATIONS ix

LIST OF TABLES xi

Chapter

1. INTRODUCTION 1

1.1 Influenza Infection, a Global Health Threat 1

1.2 Influenza, The Disease and Diagnosis 2

1.3 Influenza Epidemics and Pandemics 6

1.3.1 Influenza Epidemics 6

1.3.2 Influenza Pandemics 6

1.4 Influenza Prevention and Treatment 9

1.4.1 Vaccines 9

1.4.2 Antiviral Drugs 11

1.5 Purpose 12

2. INFLUENZA GENETICS AND PROTEOMICS 13

2.1 Influenza Virion 13

2.2 The Influenza Proteins 14

2.2.1 Surface Proteins 14

2.2.2 Proteins of the Viral Ribonucleoprotein 16

2.2.3 Matrix and Non-structural Proteins 17

3. INFLUENZA LIFE CYCLE 18

3.1 Introduction 18

3.2 Binding 19

3.3 Entry of the Virion into the Host Cell via Endocytosis 19

3.4 Fusion and Uncoating of the Virion 20

3.5 vRNP Import into the Nucleus 20

3.6 Viral RNA Transcription and Translation 21

3.7 vRNP Export 21

3.8 Virion Assembly and Release 21

4. INFLUENZA PATHWAYS 23

4.1 Biological Pathways 23

4.2 Reactome 23

4.3 Influenza Pathways 25

4.4 A Reactome Pathway Example 25

5. MODELING AND SIMULATION 28

5.1 Modeling and Simulation 28

5.2 Deterministic and Stochastic Models 30

5.3 Influenza Models 30

5.4 The Modeling Process 31

5.5 Simulation Framework 32
5.6 SimJava 32
6. FLUSIM 35
6.1 FluSim Model 35
6.2 Stages 35
6.3 Molecules 37
6.4 Ports 38
6.5 State Logic 39
6.6 Assumptions 41
6.7 Limitations 41
7. RESULTS 43
7.1 FluSim Molecular Measurements 43
7.1.1 mRNA Molecular Measurements 46
7.1.2 vRNA Molecular Measurements 48
7.1.3 cRNA Molecular Measurements 50
7.1.4 vRNP Molecular Measurements 51
7.1.5 Simulation Comparison with Experimental Results 51
7.1.6 Protein Molecular Measurements 52
7.2 Virion Measurements 54
7.3 Simulation Runtime 55
7.4 FluSim Quantification 56
7.4.1 Dizzy Components and Reactions 56

7.4.2 Comparison of FluSim discrete event simulation results to Dizzy results 58

7.5 Conclusions 63

Appendix

A. THESIS COMMITTEE QUESTIONS AND REPLY 63

REFERENCES 69

BIOGRAPHICAL INFORMATION 74

LIST OF ILLUSTRATIONS

Figure Page

1.1 Influenza virion 14

1.2 Wild aquatic birds are the main reservoir for influenza A viruses..... 16

1.3 Emergency hospital during 1918 influenza pandemic, Camp Funston, a subdivision of Fort Riley in Kansas 18

1.4 Genetic variability of the three primary pandemics of the 20th Century 19

1.5 The yearly influenza vaccine production timeline 21

2.1 Photomicrograph of influenza virions 24

2.2 The Influenza virus viral ribonucleoprotein (vRNP)..... 27

3.1 A schematic of the influenza life cycle..... 31

4.1 The anatomy of a Reactome pathway..... 35

4.2 Influenza life cycle pathways in the Reactome database..... 37

4.3 A localization view of influenza life cycle components and pathways..... 38

5.1 A schematic representation of the SimJava components..... 44

6.1 The FluSim influenza life cycle model..... 47

7.1 Influenza experimental evidence..... 56

7.2 mRNA transcriptions results for a single virion infecting a host cell..... 60

7.3 Viral RNA production in FluSim 61

7.4 Complementary RNA results of FluSim for each of the eight

influenza segments 63

7.5 Viral ribonucleoprotein results in FluSim..... 64

7.6 Protein results in FluSim..... 65

7.7 Virion particle release rates in FluSim 67

7.8 Dizzy model output for mRNA, cRNA and vRNA..... 70

7.9 Viral mRNA Dizzy results..... 70

LIST OF TABLES

Table		Page
1.1	World Health Organization Influenza Naming Standard	15
2.1	Influenza Proteins	25
6.1	The tags encoded within FluSim allowing for the passing of events to multiple stages for each of the eight RNAs or the ten proteins.	49
6.2	Processing times for the initial stages of the influenza life cycle	51
6.3	Processing times for mRNA synthesis and vRNA replication of the eight vRNP segments of influenza virus	51
7.1	Relative amounts of RNA	52
7.2	The Dizzy model components.	56
7.3	The reactions of our Dizzy model	56
7.4	Dizzy and FluSim cRNA Comparison	59
7.5	Dizzy and FluSim vRNA Comparison	60
7.6	Dizzy and FluSim mRNA Comparison	61
7.7	Protein count comparison	62

CHAPTER 1

INTRODUCTION

1.1 Influenza Infection, a Global Health Threat

The influenza virus has infected mankind for millennia being first described as an epidemic by Hippocrates in 412 BC (Kaplan, Webster, 1977). The influenza virus remains to this day a major cause of disease to such an extent that the National Institute of Health (NIH) classifies it as a re-emerging infectious disease. A virus is an infective agent, too small to be seen with the naked eye, which is able to multiply within cells of a host organism by utilizing the machineries and resources of the host cell. Furthermore, a virus typically consists of a nucleic acid, either ribonucleic acid (RNA) or deoxyribonucleic acid (DNA), surrounded by a protein coat or shell. In the case of influenza virus the nucleic acid is RNA, which is present as a single strand in each of eight segments. While outside of the cell, a single viral particle is referred to as a virion. All viruses are categorized by a taxonomy or family tree and influenza virus, in particular, is a member of the Orthomoyxoviridae family. Members of the Orthomoyxoviridae family are characterized by their segmented genomes of negative-sense and single stranded RNA (Baltimore 1971).

1.2 Influenza, The Disease and Diagnosis

The Influenza virus is the causative agent of the disease Influenza, which is a disease of the upper respiratory tract. The disease Influenza typically presents [the symptoms of] sudden fever and chills in addition to headache and sore throat, malaise, anorexia and a dry cough. The fever usually peaks within a day of the disease symptoms and lasts for 1 – 5 days. The symptoms of a person infected with influenza include a clear runny nose, flushed face, hot and moist skin and a general appearance of sickness [Cox, Subbarao, 1999].

Patients who exhibit the symptoms of influenza may have nasopharyngeal (nasal) and throat swabs taken for clinical testing, viral identification and epidemiological disease tracking. Various tests are commonly performed on the samples including virus isolation, detection of viral proteins, detection of viral nucleic acid, and serological diagnosis [Cox, Subbarao, 1999]. Historically the serological (blood) diagnosis helped to determine the serotype or subtype of the virus. The serotype or subtype of the virus as determined by serological analysis provides a rough measure of each of two of the external proteins on the influenza virus surface, namely hemagluttanin (HA) and neuraminidase (NA). Only two of the eleven proteins may be measured by serological analysis because they interact with the humoral immune system and thus elicit an immune response that may be tested in the laboratory. The remaining proteins exist in the interior of the virus or in the case of M2 in the viral envelope or outer shell as shown in figure 1.1.

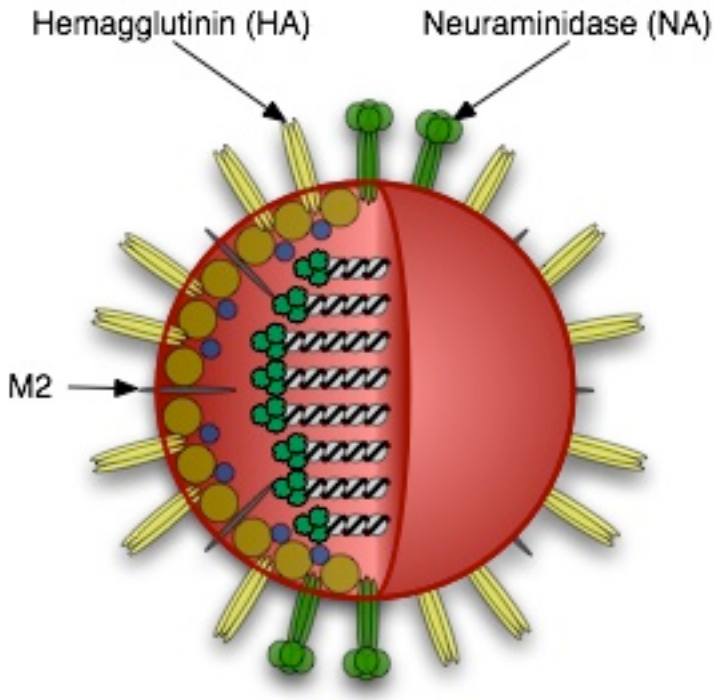


Figure 1.1. The Influenza virion. Exterior proteins hemagluttanin (HA), neuraminidase (NA) and matrix protein 2 (M2) are shown.

In addition to serological analysis, detection of viral RNA sequences and virus isolation may also be performed on clinical samples. Currently, virus isolation is the gold standard in laboratory diagnosis of influenza infection but the process takes a minimum of 1 – 2 days to complete thus preventing the techniques' use in guiding clinical treatment.

The need for a much quicker method of virus identification to assist in directing clinical treatment has been met with the development of new techniques in the detection of viral proteins. Though this technique is less sensitive then detection of virus nucleic acids it does provide the much-needed clinical guidance.

Finally, the detection of virus nucleic acid (RNA) is possible utilizing the techniques of reverse-transcription followed by the polymerase chain reaction (PCR), a form of molecular Xeroxing™. Reverse-transcription is the process by which the influenza virus RNA serves as a template for the production of a mirror image of the virus' RNA. This is followed by the performance of the polymerase chain reaction or PCR where the mirror image copy is in essence "xeroxed" producing many nearly perfect copies. The end result of these techniques is to provide enough material to make virus identification possible.

Once an influenza virus sequence is known, the strain is given a name. Influenza viruses are named according the standard nomenclature established by the World Health Organization (WHO). According to this standard an influenza virus name begins with the influenza virus type (A, B, or C), followed by the species from which the virus is isolated (omitted if human), the geographical location of the isolation, the isolate identifier, the year of isolation and in the case of influenza A viruses, the hemagluttanin (H) and neuraminidase (N) subtype. For example, the virus of H1N1 subtype isolated from humans in Puerto Rico in 1934 is: A/Puerto Rico/8/34 (H1N1). Currently there are 16 different hemagluttanin (H1 to H16) subtypes and 9 different neuraminidase (N1 to N9) subtypes known for influenza A viruses [Krug & Lamb 2001].

Table 1.1. World Health Organization Influenza Naming Standards

Influenza Type	Host (Non-human)	Geographical Location	Isolation Number	Isolation Year	Subtype
A		Puerto Rico	8	1934	(H1N1)
A	Goose	Guangdong	1	1996	(H5N1)

Among the three types of influenza, type A is the most prevalent in humans followed by type B and type C. In addition to humans, various other animals serve as influenza virus hosts. Common hosts include various types of birds (avian), hosts, pigs, horses, ferrets, and so on.

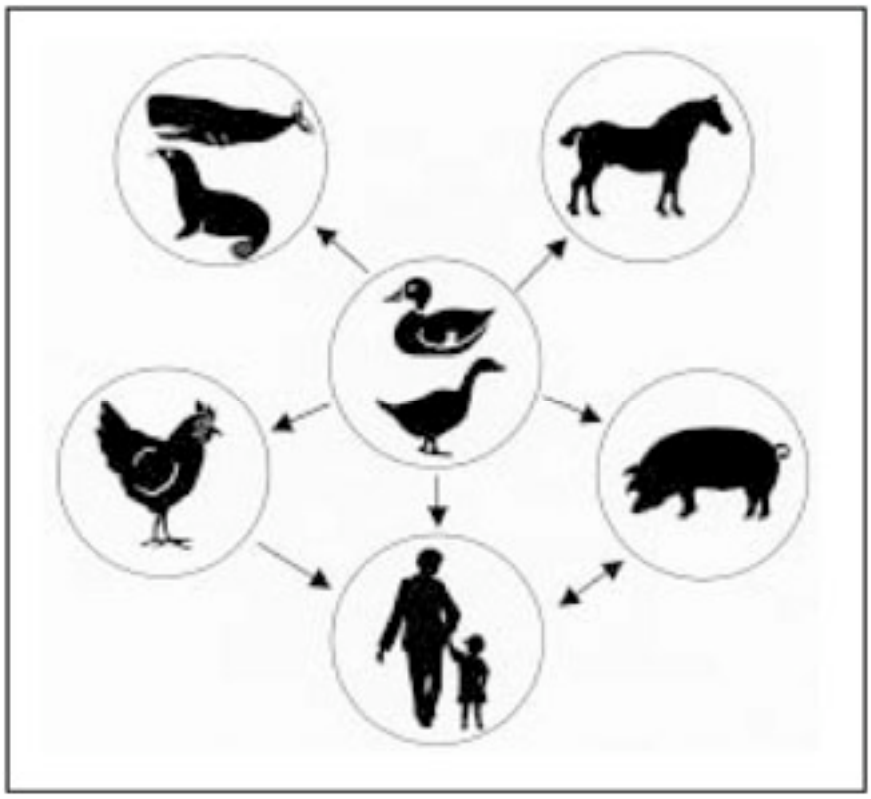


Figure 1.2. Wild aquatic birds are the main reservoir for influenza A viruses. The arrows indicate direction infection demonstrating that pigs and poultry can also infect humans. (Trampuz, Prabhu et al. 2004)

1.3 Influenza Epidemics and Pandemics

1.3.1 Influenza Epidemics

The influenza virus has a high rate of mutation especially in the surface proteins HA or NA, which may in turn cause significant enough confirmation changes to be classified as a new subtype and possibly cause an epidemic or a pandemic. An epidemic is a disease that affects a disproportionately large number of people within a population at the same time while a pandemic is an epidemic of global proportion. Influenza epidemics kill approximately 36,000 people each year ((ACIP) 2006).

1.3.2 Influenza Pandemics

In the 20th century an influenza pandemic, which would turn out to be the most devastating pandemic in recorded human history, occurred during 1918 and killed an estimates 20 – 40 million people worldwide. It became known as the "Spanish Flu" due to its suspected origin in Spain and had a subtype of "H1N1" (Crosby 1989).

The 20th century saw other pandemics including a 1957 pandemic from the H2N2 "Asian" influenza subtype and a 1968 pandemic resulting from the H3N2 influenza subtype. Neither the 1957 nor the 1968 pandemics approached the deadliness of the 1918 pandemic, killing only 70,000 and 30,000 Americans respectively (Thompson, Shay et al. 2003).

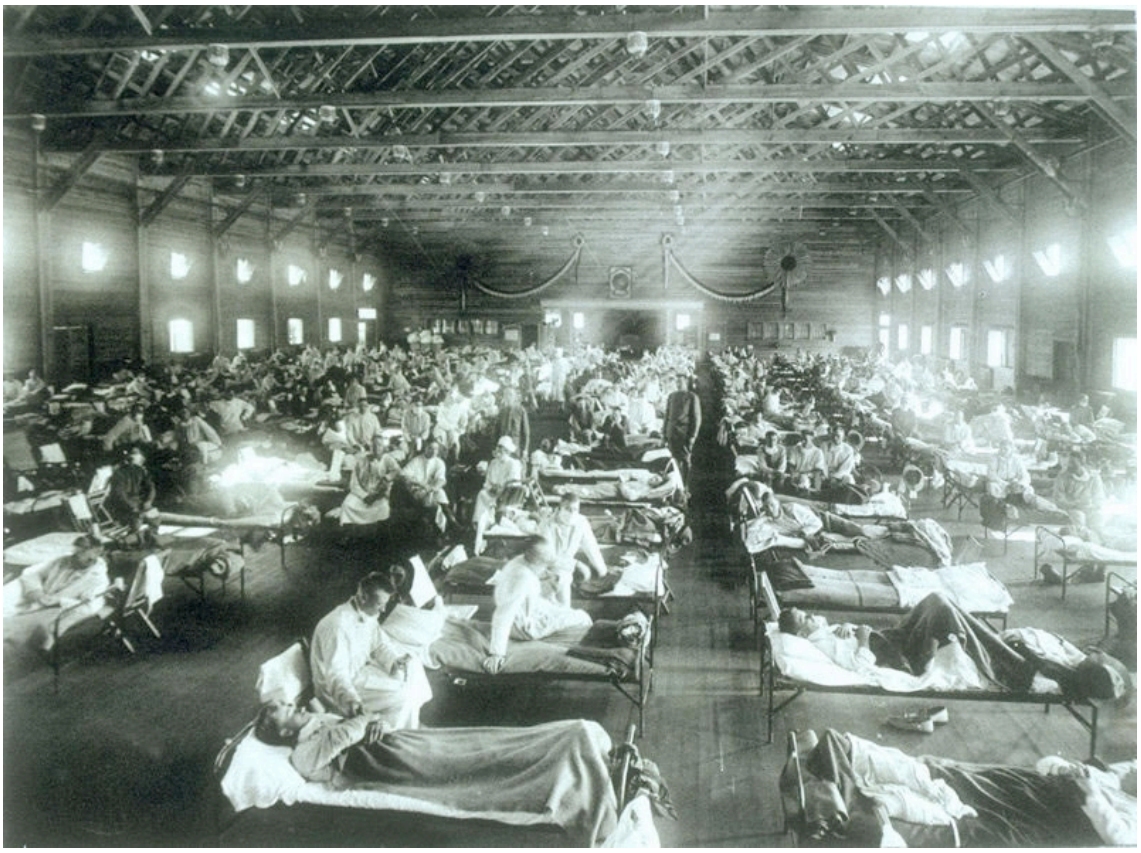


Figure 1.3. Emergency hospital during 1918 influenza pandemic, Camp Funston, a subdivision of Fort Riley in Kansas. (© National Museum of Health and Medicine, Armed Forces Institute of Pathology, Washington DC).

Pandemics result from a new and significant variation of the influenza virus that is introduced into a naive population. With each new pandemic it was discovered that mutations in the hemagluttanin ("H") proteins or the neuraminidase ("N") protein or both had taken place, as shown in figure 1.4, circumventing any immunological "memory" infected individuals possessed.

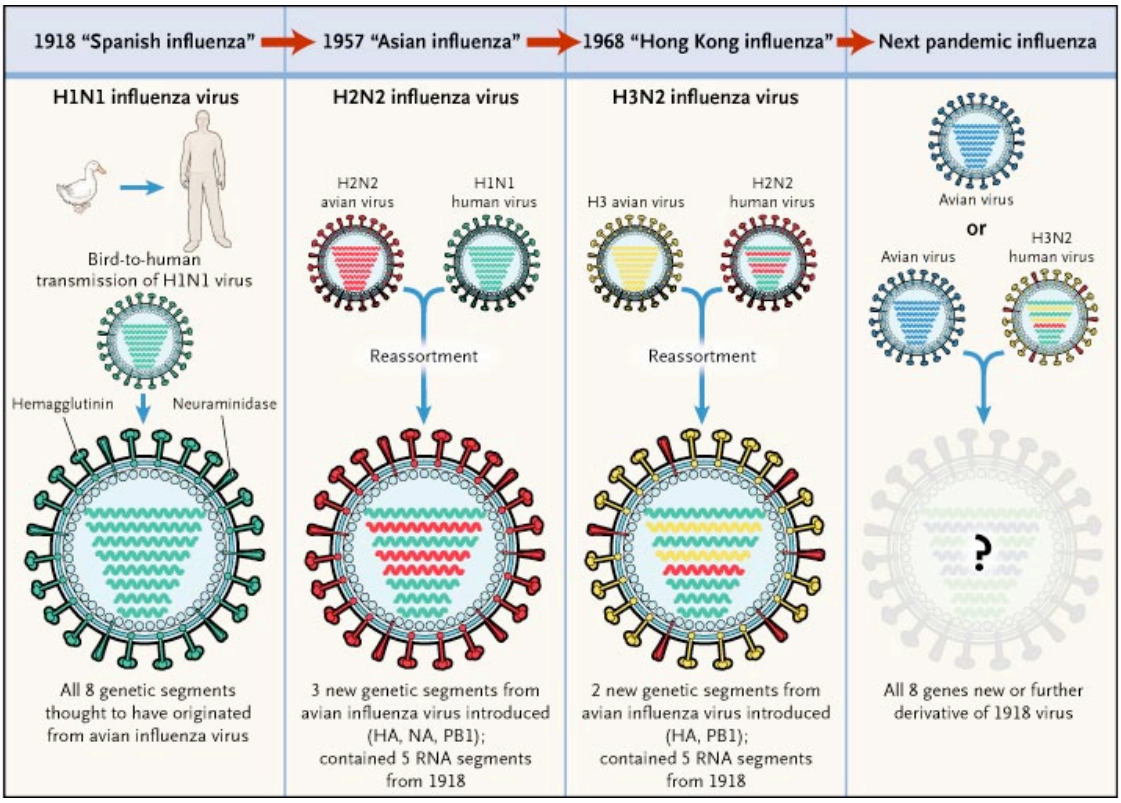


Figure 1.4. Genetic variability of the three primary pandemics of the 20th Century. (Belshe 2005)

The fear of a new potential pandemic surfaced in Hong Kong in 1997 with the discovery of the H5N1 influenza subtype also known as the "Avian Flu" or "Bird Flu". (Subbarao, Klimov et al. 1998) Avian species are known to be a reservoir for the influenza virus but direct transmission from bird to human was not previously known. With the appearance of the H5N1 subtype, the fear of a pandemic that could rival or surpass the 1918 "Spanish flu" became a possibility. This fear drives much of the current influenza research.

1.4 Influenza Prevention and Treatment

Yearly influenza epidemics as well as the fear of a possible pandemic would be minimal if there was a proven, preventable measure for such occurrences. Currently, there is not a single influenza vaccine that is capable to reliably protecting against multiple years of influenza epidemic strains nor it is possible to protect against an as-yet-unknown potential pandemic strain of influenza. There are a number of reasons why including the mutability of the influenza virus, the facts that the predominant influenza virus strains change from year to year as well as the vaccine production timeline and growing resistance to antiviral drugs. (Gerdil 2003)

1.4.1 Vaccines

Influenza vaccine production is a lengthy process, which involves some guesswork in the beginning. The process begins when the World Health Organization officials analyze the dominant circulating strains of influenza. From this process the predominant strains of influenza are classified and declared the official strains for the upcoming influenza vaccine. During the time of production of the vaccine the influenza virus can mutate thus making the official strain selections less effective. Once the dominant strains are classified, the Food and Drug Administration (FDA) distributes seed viruses to vaccine manufacturers. Each strain is prepared separately and is combined later. The strains will be incubated in chicken eggs and then extracted and purified. The virus particles are then chemically broken apart to disrupt the virus and the fragments are collected and combined. The influenza vaccine components then

undergo testing by the FDA. Once the vaccine components are tested they are then packaged for delivery and shipped. The entire process takes approximately 9 months from January to October.

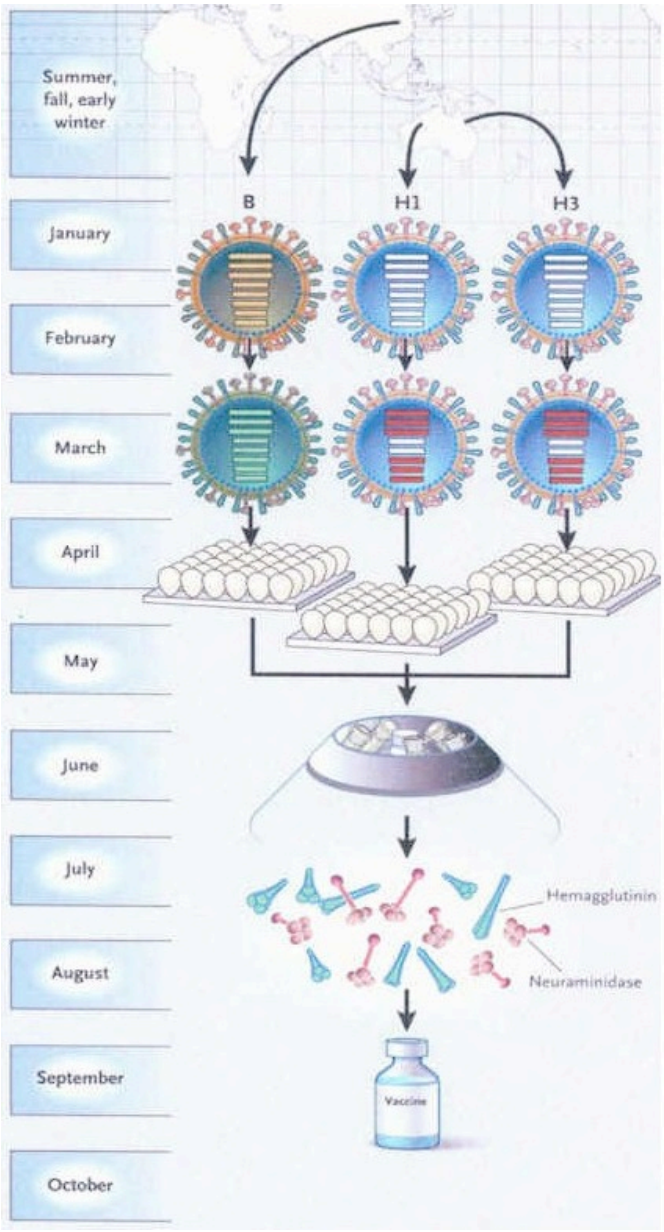


Figure 1.5. The yearly influenza vaccine production timeline begins in January with the isolation of the predominant serotype from influenza A and B virus and continues through October to produce a ready vaccine. This must be repeated each year.

While there have been update to the vaccine production process, namely recombinant vaccines (Roberts, Kretzschmar et al. 1998) that remove the need to incubate viruses in eggs, the fact remains that the flu vaccine must still be reconstituted and produced yearly and there is no guarantee that the vaccine will prevent pandemic influenza infection.

1.4.3 Antiviral Drugs

Influenza vaccines are not the only treatment option for influenza. While vaccines can prevent a person from being infected by the influenza virus they are specific to particular strains of influenza. For patients that may be infected already, antiviral drugs offer an opportunity to minimize the effects and duration of the disease. In the case of influenza there are two broad categories of antiviral drugs - neuraminidase inhibitors, which prevent the release of the virion from the host cell, and M2 ion channel inhibitors, which prevent the fusion and uncoating of the virion within the host cell endosome. These antiviral drugs work well to contain the virus and could possibly help to contain a pandemic strain of influenza but there are cases of resistance to these antivirals. Resistance to antiviral drugs is a quality of some influenza virus strains, which reduces the effectiveness of the drugs in these particular strains because of changes in the binding characteristics between the drugs and the influenza virus. (Bright, Shay et al. 2006)

1.5 Purpose

In an effort to gain a better understanding of the dynamics of the influenza virus I present here a stochastic simulation of influenza infection in a host cell. It is my goal that this work helps to increase understanding of the influenza virus infection that, despite the great efforts put forth over recent decades to understand it still eludes us today. The present thesis models only the dynamics of a single influenza virus in a host cell and does not include the modeling of the host immune response dynamics nor does it include the dynamics of the spreading of the newly created virion to other cells. I expect that this effort continues to evolve just as the virus itself does with the ultimate goal of assisting in the elucidation of the keys to unlocking the elusive nature of the influenza virus and its actions. We expect that once we develop the complete model, the model can be used to understand the dynamics of different drugs on this system and their quantitative effects on the different molecules of the cell.

CHAPTER 2

INFLUENZA GENETICS AND PROTEOMICS

2.1 Influenza Virion

The influenza virion consists of the outer protein lipid bilayer, which encompasses the viral ribonucleoproteins of the virus. The viral ribonucleoproteins are complexes that package the viral RNA for transport and stability. The influenza virion itself is highly plieomorphic or variable in shape though it is commonly spherical in form with an average diameter of 80-120 nm.

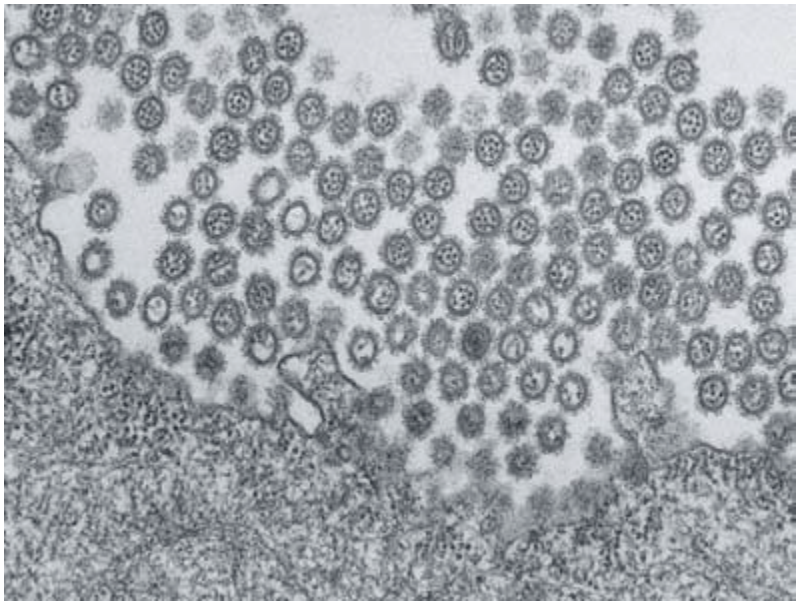


Figure 2.1 Photomicrograph of influenza virions. (Noda, Sagara et al. 2006)

2.2 The Influenza Proteins

There are eleven proteins encoded by the eight influenza viral ribonucleoproteins (vRNP). These include PB2, PB1, PB1-F2, PA, HA, NP, NA, M1, M2, NS1 and NS2/NEP. These proteins fall into the following broad categories: surface proteins, ribonucleoproteins proteins and matrix and non-structural proteins.

Table 2.1 Influenza Proteins

Segment	Name	Length	Description
1	PB2	2341	Polymerase (basic): cap binding
2	PB1	2341	Polymerase (basic): elongation
2	PB1-F2	382	PB1-F2
3	PA	2233	Polymerase (acidic): protease activity
4	HA	1778	Hemagluttanin
5	NP	1565	Nucleoprotein: RNA binding; transport of vRNA
6	NA	1413	Neuraminidase: release of virus
7	M1	1005	Matrix protein
7	M2	315	Membrane protein - ion channel
8	NS1	868	Non-structural protein; Host counter-measures protein
8	NEP	395	Non-structural protein

2.2.1 Surface Proteins

On the surface of the influenza virion one will notice a distribution of three different proteins, hemagluttanin (HA), neuraminidase (NA) and matrix protein 2 (M2). The hemagluttanin (HA) protein exists as a trimeric protein, composed of three copies of the hemagluttanin monomer. Tetramers of the hemagluttanin protein protrudes from the virus surface forming pockets that binds to a host cells sialic acid thus beginning the influenza life cycle. The configuration of the hemagluttanin tetramer determines which

of the two types of sialic acid the virion will bind to. Sialic acids are characterized by their configuration and their respective binding characteristics to the viral hemagluttanin. Sialic acids may have a $\alpha 2.3$ or $\alpha 2.6$ configuration. The $\alpha 2.3$ configuration preferentially binds with human influenza viruses, while the $\alpha 2.6$ linkages preferentially bind to avian influenza viruses.

In addition to the role hemagluttanin plays in binding the virion to the host cell, the hemagluttanin protein also plays an important role in the fusion and uncoating process, which facilitates the movement of the viral ribonucleoprotein (vRNP) from the virion to the host cell nucleus.

The neuraminidase (NA) protein exists as a tetrameric protein that also protrudes from the virus surface. The neuraminidase protein serves an important role in completing the release of a new virion from the host cell as the neuraminidase protein cleaves the sialic acids away from the hemagluttanin proteins also on the surface of the virion thus facilitating the release of the virion. The neuraminidase's release function is the focus of the antiviral drugs that works to prevent the release resulting a clumping of new virions.

Matrix protein number two (M2) exists as a tetrameric protein and like the neuraminidase and hemagluttanin proteins, penetrates the virus plasma membrane or outer coat. Unlike the HA and NA proteins, the M2 does not extend as far beyond the plasma membrane. The M2 protein plays in important role in the influenza life cycle by facilitating the incorporation of protons (H+) into the virion interior reducing the pH

and initiating a structural change in the hemagluttanin protein required for fusion of the virion to the endosome wall.

2.2.2 Proteins of the Viral Ribonucleoprotein

The influenza virus RNA is single stranded RNA, which is quickly degraded within a cell if not complexed with other proteins. The vRNP consists of one of the eight viral RNA segments complexed with multiple nucleoproteins (NP), and one of each of PA, PB1 and PB2.

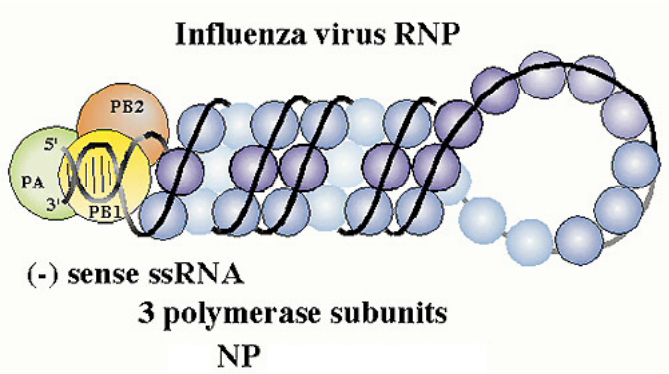


Figure 2.2 The Influenza virus viral ribonucleoprotein (vRNP) is comprised of one of each polymerase protein (PA, PB1, PB2) as well as numerous nucleoproteins (NP) shown in shades of blue. Influenza segment RNA is shown as a black line. (Portela and Digard 2002)

In the case of the influenza virus, the viral RNA forms a complex with a number of nucleoproteins (NP) and is topped off with the polymerases proteins as shown in figure 2.1 The nucleoproteins provide a valley that the vRNA strand binds to and stabilizes the RNA. On average there is one NP protein per 24 bases of viral RNA (Portela and Digard 2002).

In addition to the nucleoproteins, the vRNP also packages the three polymerase proteins. These polymerase proteins, PB2, PB1 and PA, which are encoded by segments 1, 2 and 3 respectively, are crucial for replication of the virus once the vRNP reaches the host cell nucleus.

2.2.3 Matrix and Non-structural Proteins

The remaining influenza virus proteins include another matrix protein and the non-structural proteins. In addition to the M2 matrix proteins that exist in the plasma membrane, segment 7 also encodes the M1 matrix protein. The M1 protein serves to anchor the surface proteins by establishing a layer adjacent to the plasma membrane. It is this layer of M1 proteins that the HA, NA and M2 bind to, thus providing an anchor for each protein to extend from.

The non-structural proteins serve important roles in the virion. The non-structural proteins 1 (NS1) is known to operate as a counter-measure to the host defense mechanisms against viral infection thus ensuring unencumbered replication of the influenza virus within the host cell (Wang, Li et al. 2000).

CHAPTER 3

INFLUENZA LIFE CYCLE

3.1 Introduction

To accurately model and simulate the influenza life cycle we must understand the process that the influenza virus goes through as it infects a host cell and how the components of influenza virus interact with each other. The influenza life cycle is comprised of 14 stages from binding to release. These stages include binding (or attachment) of the virion to the host cell, entry of the influenza virion into the host cell within an endosome, fusion of the virion to the host endosome and uncoating of the vRNP, importing of the vRNP into the host cell nucleus, the various viral RNA synthesis and translation stages, assembly and export of the vRNP from the host nucleus, budding of the new virion and finally release of new viral particles or virions.

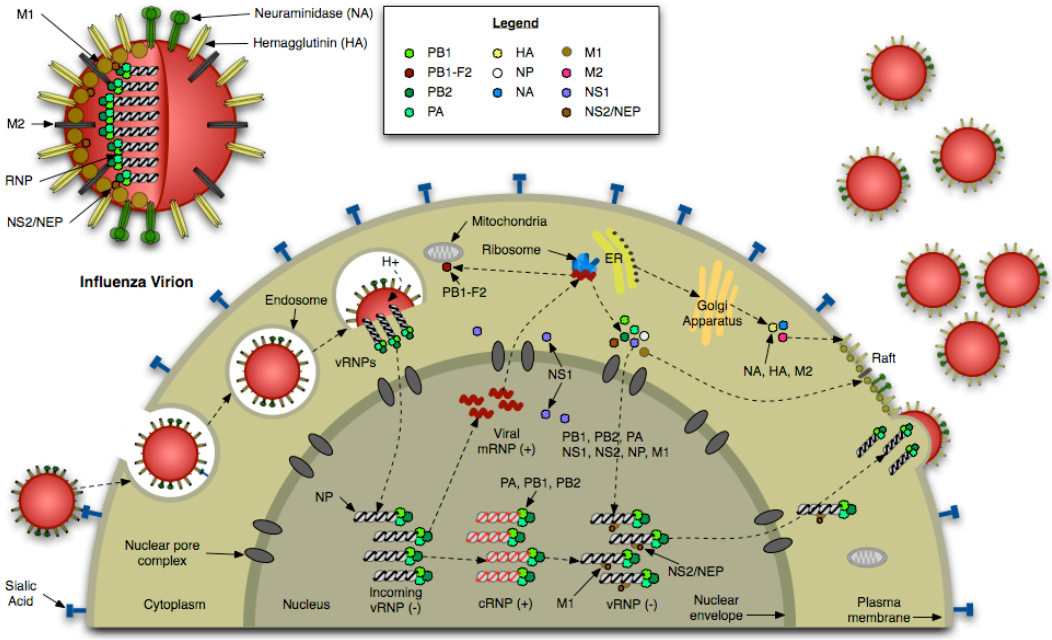


Figure 3.1. A schematic of the influenza life cycle displaying the 19 stages of influenza life cycle as described in the text.

3.2 Binding

The influenza viral life cycle begins with the binding of a virion to a host cell. Influenza virions come into contact with a host cell though the mechanism of inhalation in which one or more influenza viral particles are inhaled through the nasal passage and attach themselves to a hosts epithelial lung cell (Yoshimura, Kuroda et al. 1982). Influenza viruses bind via surface hemagluttanin (HA) proteins to sialic acids on the host cell surface. Sialic acids exist in two configurations, either the alpha 2,3 or the alpha 2,6 configuration of the galactose molecule.

In most circumstances, host cells posses one of the two sialic acid configurations with humans possessing the alpha 2,6 linkage and avian hosts possessing the alpha 2,3 linkage configuration. Only pigs possess both linkages. The specificity of influenza hemagluttanin for sialic acid in alpha 2,6 or alpha 2,3 configuration is a

feature restricting the transfer of influenza viruses between avian species and humans (Connor, Kawaoka et al. 1994; Gambaryan, Robertson et al. 1999; Mochalova, Gambaryan et al. 2003).

3.3 Entry of the Virion into the Host Cell via Endocytosis

Virus particles bound to the host cell surface can be internalized by one of four mechanisms. Most internalization appears to be mediated by clathrin-coated pits, but internalization via caveolae, macropinocytosis, and by non-clathrin, non-caveolae pathways has also been described for influenza viruses. Internalized viruses become internalized within an endosome or membrane bound compartment within the host cell (Matlin, Reggio et al. 1981; Yoshimura, Kuroda et al. 1982; Sieczkarski and Whittaker 2003; Lakadamyali, Rust et al. 2006).

3.4 Fusion and Uncoating of the Virion

Once internalized, acidification or infusion of protons (H+) into the virion via the M2 ion channel stimulates fusion of the viral and endosomal membranes. The acidification in turns causes a structural change in the viral hemagluttanin proteins and frees the fusion peptide of the hemagluttanins HA2 subunit to interact with the endosome membrane. The concerted structural change of several hemagluttanin molecules opens up a pore through which the viral RNP passes into the cytosol of the cell (Martin and Helenius 1991; Kemler, Whittaker et al. 1994; Whittaker, Kann et al. 2000).

3.5 vRNP Import into the Nucleus

Once the ribonucleoproteins have been released into the cytoplasm of the cell they are then transported into the host nucleus. The transport process begins with the recognition of the nuclear localization signal (NLS) located on each nucleoprotein of the vRNP by proteins known as karyopherins. A trimeric complex of the vRNP and two karopherin proteins forms and then binds with the nuclear pore and the host cell nuclear membrane. The vRNP are then transported and released into the cell nucleus (Martin and Helenius 1991; Kemler, Whittaker et al. 1994; Whittaker, Kann et al. 2000).

3.6 Viral RNA Transcription and Translation

Once inside the nucleus, influenza viral ribonucleic acids (vRNA) are then be transcribed and replicated. The viral negative-strand RNA (vRNA) serve as a template for the synthesis both of capped, polyadenylated viral messenger RNA (mRNA) and of full-length positive-strand RNA, "complementary RNA" (cRNA). The cRNA in turn serves as a template for the synthesis of new vRNA molecules. These three reactions are catalyzed by the viral RNA polymerases that enter the host cell nucleus with the viral RNP complex (Shapiro and Krug 1988; Mullin, Dalton et al. 2004; Vreede, Jung et al. 2004; Amorim and Digard 2006; Vreede and Brownlee 2006).

3.7 vRNP Export

With the completion of RNA transcription, assembly and nuclear export of the vRNP may then take place. Influenza viral RNA (vRNA) synthesized in the nucleus of

the infected host cell is packaged into ribonucleoprotein (RNP) complexes containing two viral proteins, NP and M1. Nucleoprotein (NP) provides a stable carrier to protect the sugar phosphate backbone of the vRNA. Matrix protein 1 (M1) is critical for export of the vRNP complex from the nucleus, mediating the interaction of the vRNP complex with the viral NEP/NS2 protein (Boulo, Akarsu et al. 2007).

3.8 Virion Assembly and Release

Parallel to vRNP export, surface proteins HA, NA and M2 are synthesized in the cytoplasm and secreted into endoplasmic reticulum where the proteins are folded and glycosolated. Hemagluttanin proteins are then assembled into trimers while NA and M2 are assembled into tetramers and all are exported to the Golgi apparatus. The final assembly steps include association of the HA, NA and M2 into lipid rafts along the plasma membrane. The packaging of the eight vRNP occurs by selective incorporation or random incorporation. Upon completion of budding the influenza virion is released by way of neuraminidase enzymatic release of sialic acid. (Leser and Lamb 2005; Boulo, Akarsu et al. 2007)

CHAPTER 4

INFLUENZA PATHWAYS

In our quest to model the influenza life cycle it is crucial to understand the pathways that the pathogen, the infecting influenza virion, and its components traverse within the infected host cell. Biological pathways play an important role in understanding how each of the various proteins that comprise an organism operate within that organism and potentially how they interact with other organisms as is the case with the influenza virus and a host cell.

4.1 Biological Pathways

Biological pathways describe how molecules such as proteins interact with other molecules to accomplish the functions of a cell. Biological functions cover all of the processes required to sustain an organism, reproduce, defend itself as well as respond to the surrounding environment. Each molecule can act as a component in one or more pathways within the network of interactions of an organism.

4.2 Reactome

In recent years numerous biological pathway databases have been created in an effort to quantify and capture biological pathways. Efforts such as BioCyc (Karp, Ouzounis et al. 2005), KEGG (Kanehisa and Goto 2000) and Reactome (Joshi-Tope,

Gillespie et al. 2005) have succeeded in creating repositories of biological pathway knowledge. While these three databases do not represent the entirety of biological pathway knowledge they do represent three of the most significant efforts.

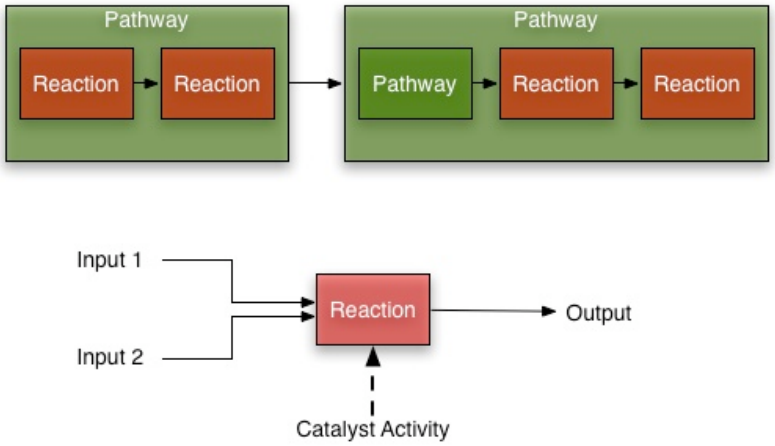


Figure 4.1. The anatomy of a Reactome pathway. In Reactome, a pathway is comprised of one or more reactions. Each reaction (bottom) has input and output as well as the capability to facilitate a catalyst. Pathways (top) may contain only reactions as shown in the upper left or a pathway may contain some combination of pathways and reactions.

The Reactome project (www.Reactome.org) is "a curated knowledgebase of biological pathways" consisting of human pathways along with orthologous or equivalent pathways in other organisms. The influenza life cycle pathways are one of two non-human pathways within the Reactome database (HIV being the other). A significant advantage that the Reactome database has over other pathway databases is the ability to facilitate the representation of both metabolic pathways and signal transduction pathways. An additional feature that will become important for future modeling efforts is the ability to export Reactome pathways in various open pathway exchange formats including BioPAX and SBML (Hucka, Finney et al. 2003).

4.3 Influenza Pathways

On behalf of the BioHealthBase Bioinformatic Resource Center (BRC) for Biodefense and Emerging/Re-emerging Infectious Disease we contributed the initial framework of the influenza life cycle and influenza host-pathogen interactions to the Reactome database. Utilizing this framework the pathways of the entire influenza life cycle have been annotated in the Reactome database with the help of many influenza experts. These pathways help to form the foundation of this model and simulation.

The influenza additions to the Reactome project include the initial framework of influenza life cycle infection as well as the pathway details for all processes therein. Additionally, we have added preliminary pathways focusing on host-pathogen interaction of the influenza virus and host cell.

4.4 A Reactome Pathway Example

Echoing the stages of influenza infection described in Chapter 3 the Reactome pathways represent the molecular details of each biological pathways as best understood at the present time. An outline of the influenza life cycle can be seen in figure 4.2. To complement the influenza pathways we have also added the host immune response pathways of the toll-like receptor-3 (TLR3) and retinoic acid inducible gene I (RIG-I). The TLR3 and RIG-I pathways are response pathways of the host cells, which induce the interferon-alpha/beta cascade. The details used to describe pathways in the Reactome project provide critical information for the modeling and simulation process. An example of this is graphically represented in figure 4.3 in which the entire influenza

life cycle pathway from the Reactome project has been exported in a biological pathway exchange format, BioPAX and then imported into the pathway display software Cytoscape. Utilizing the Cerebral plug in, the influenza life cycle components are display according to their location in the cell using on the Reactome pathway details.

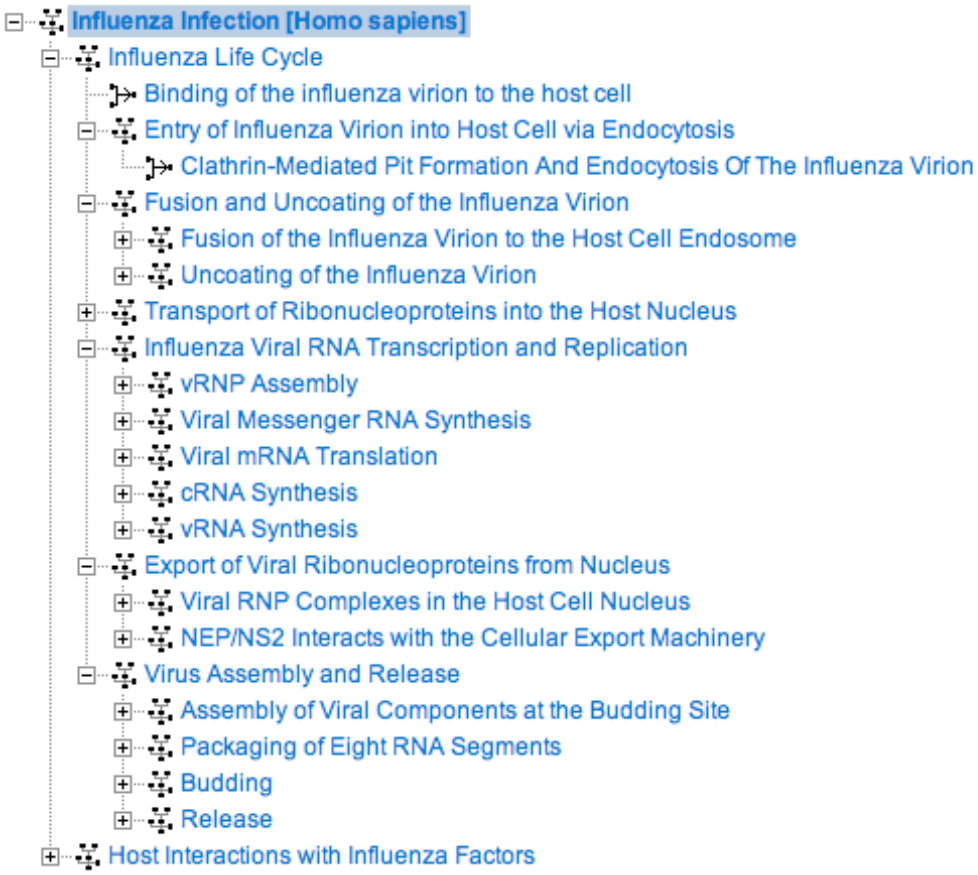


Figure 4.2. Influenza life cycle pathways in the Reactome database.

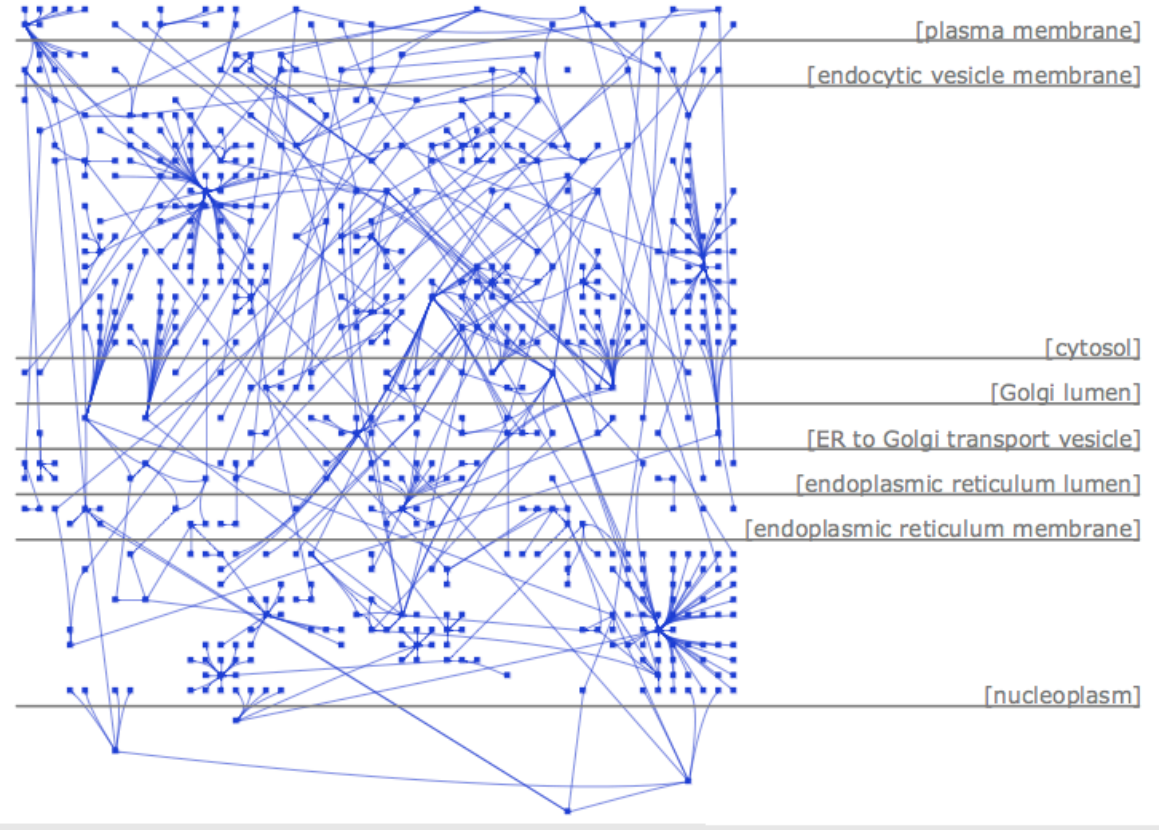


Figure 4.3. A localization view of influenza life cycle components and pathways. Using the Reactome pathway data for influenza life cycle, the components and pathways are imported into CytoScape and graphed using the Cerebral plug-in, which allows molecules to be grouped by cellular compartment. Show here is a reproduction of cellular location beginning at the top and working inwards towards the cell nucleus. (Barsky, Gardy et al. 2007).

CHAPTER 5

MODELING AND SIMULATION

5.1 Modeling and Simulation

Modeling is the process of abstracting a system, existing or proposed, for the purpose of understanding how it functions. A system is abstracted in an effort to conserve resources such as time or processing power while concentrating on the most important aspects of the system, which must come at the expense of the details of that system. The process of abstraction further allows the modeler to obtain the desired information about the system being modeled within an abbreviated timeframe or allowing researchers to simulate something such as influenza infection that would be unethical to test in a real person.

Models fall into two broad categories, physical (modeling the physical world) and logical (modeling the relational behavior of the physical resources). We will only consider logical models in out treatment of this subject. Logical models include approximations and assumptions about both the structural and quantitative aspects of the physical system we wish to model. (Kelton, Sadowski et al. 1997)

A mathematical model, which is a subset of logical models, is an abstract model that uses mathematical equations to describe the behavior of a system. Eykhof (1974) defined a mathematical model as “a representation of the essential aspects of an existing

system (or a system to be constructed) which presents knowledge of that system in usable form".

Mathematical models can be classified as belonging to one of several categories including: static vs. dynamic or deterministic vs. stochastic. Static models do not account for the changes in time while dynamic model do. Deterministic model have a deterministic or unchanging values for their model while stochastic models allow for changes representing randomness. Each category of models has advantages and disadvantages as well as a best fit to a particular scenario. In the case of modeling a biological system such as the influenza life cycle we believe the stochastic model provides the best fit. Biological molecules within a cell are by nature stochastic, moving and interacting in a seemingly random manner. The mathematical treatment of a stochastic process falls within the scope of random algebra. Using the algebra of random numbers one can solve very simple stochastic process. An important distinction between mathematical modeling and the discrete event, stochastic modeling used in FluSim is the fact that mathematical modeling works best for systems in a steady state. Utilizing discrete event, stochastic models one is better able to model systems that are not in the steady state such as influenza infection where the process being modeled has discrete events that occur only when a previous event has been completed. In this case events can only take place if they are triggered by their preceding event.

By putting a model in motion, simulation allows one to observe a models behavior over time. The behavior of a model is evaluated numerically especially with

the help of simulation software. Later in this chapter we will discuss the particular simulation framework used to develop the FluSim model and simulation.

5.2 Deterministic and stochastic models

Predictable and definite laws govern deterministic processes. In a deterministic model every set of variable states is uniquely determined by parameters in the model and by sets of previous states of these variables. Therefore, deterministic models will always perform the same way for a given set of initial conditions. Conversely, in a stochastic model, randomness is introduced, and variable states are not described by unique values, but rather by probability distributions. Over the course of multiple runs average values can be computed for variable states based upon a stochastic model. In some cases, stochastic models are used to simulate deterministic systems, which include smaller-scale phenomena that cannot either be accurately observed or modeled. As such, these small-scale phenomena are effectively unpredictable. A good stochastic model manages to represent the average effect of unresolved dynamics on larger-scale phenomena in terms of a ranged value.

5.3 Influenza Models

The application of modeling to the field of influenza biology has produced numerous models with the wide range of viewpoints from those detailing a specific aspects of influenza infection such as the M2 ion channel (Forrest, Kukol et al. 2000), to models exploring the kinetics of influenza infection from a purely mathematical

viewpoint (Baccam, Beauchemin et al. 2006), to broader influenza infection models affecting populations including a cellular automaton model of influenza infection (Beauchemin, Samuel et al. 2005). What is missing is an examination of influenza life cycle dynamics at the molecular level. The dynamic understanding of the influenza life cycle at the molecular level may help to identify the critical pathways or molecules in the life cycle while offering the ability to measure the impact of changes to potential critical pathways or molecules in the influenza life cycle dynamics. Our motivation for this project is driven by this opportunity to find critical pathways or molecules in the influenza life cycle in hopes of elucidating further treatments or vaccine opportunities for influenza infection.

5.4 The Modeling Process

Believing that a stochastic model will best reflect the underlying biology, we have constructed a model of influenza infection within a host cell utilizing a discrete event stochastic model for the task. The process of designing our model includes determining the various entities of the model, the initial characteristics that define each entity as well as each entities behavior in the system. We continue by defining how the entities interact with each other and finally we define a termination condition for the simulation. Once we have the model defined, we implement it and run the simulation of the model. Eventually we hope to perturb the model while observing the changes in the behavior of the system. We can then test our system against previously experimentally determined perturbations before we commencing with novel perturbations while

confirming them in the laboratory. Our treatment here will be to model the important stages of influenza infection using experimental evidence of influenza infection as our guide.

5.5 Simulation Framework

In developing a new model and simulation we have the choice of starting from scratch with regards to programming or utilizing one of the many tools or frameworks in existence. In our case the possibility of hosting the simulation in some form on a website directed us towards Java based frameworks. In reviewing the Java simulation landscape in 2006 one solution stood out amongst the competition in terms of maturity of the framework and documentation as well as providing the flexibility of later web hosting (McNab and Howell 1998; Ghosh, Ghosh et al. 2006). That framework is the SimJava2 framework developed by McNab and Howell.

5.6 SimJava

SimJava is a discrete event, process oriented simulation package originally based on HASE++, a C++ simulation library. SimJava provides the capability to create interacting entities or states with ports linking one entity to another. Through these ports entities pass events to one another. The SimJava framework contains Java classes covering items such as entities, ports, predicates and animations all of which may be extended using the Java programming language to customize the system for our particular use.

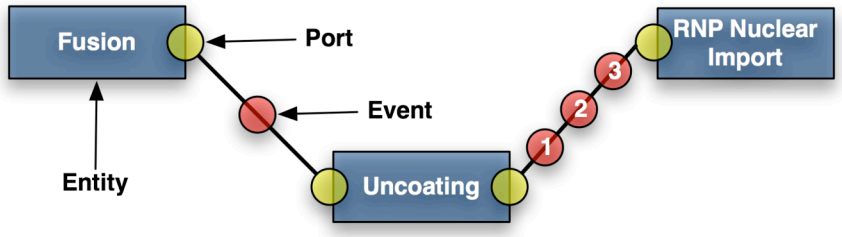


Figure 5.1. A schematic representation of the SimJava components (entity, port and event) and an event passing from the fusion entity through an outgoing port to the Uncoating entity's incoming port. The event represents a single virion. In cases where multiple molecules or events occur between the same stages as is the case for the 8 vRNP events traveling between the Uncoating stage and the RNP nuclear import stage, each event is tagged with a unique identifier.

In the SimJava framework an entity is an independent and distinct object or process that the modeler wishes to represent. Entity can have built-in logic as we shall see with some of the FluSim entities. In the case of SimJava entity considerations are also guided by the fact that statistics may only be gathered on entities and their respective ports. Therefore, if one wishes to gather statistics utilizing the built in statistical capabilities of SimJava, an entity declaration is required.

Each entity in SimJava is linked to one or more entities through ports as shown in Figure 5.1. For an entity to receive events it must have one or more incoming ports and for an entity to send events it must have one or more outgoing ports. In the main SimJava class the modeler initializes each entity then links outgoing ports to incoming ports for the appropriate entities. Once the entities and ports are initialized the simulation can then send an event from one connected entity to the designated entity. In our examples, the synthesis of mRNA becomes an entity while a port to protein

translation provides us the ability to model protein synthesis. In SimJava an event is scheduled for delivery with the Sim_schedule method. The Sim_schedule event sends an event consisting of the sending entity, a delay after which it should be processed and a tag to identify the particular event type. As we shall see our model uses the tag identifier to distinguish which of the molecules (e.g. PB2, PB1, PA) of a particular type (mRNA, cRNA, vRNA, etc) is being sent from an entity for example. Additionally, the influenza state that is the source of an event is included as data in the event occurrence, which allows us to pass an entities event through a molecular entity such as a vRNP or vRNA molecule so we can gather statistics on both the influenza state as well as the molecular entity.

With version 2.0, SimJava now includes significant enhancements including built in statistical and reporting capabilities as well as random number generators tailored for use in stochastic simulations.

CHAPTER 6

FLUSIM

6.1 FluSim Model

The FluSim model utilizes the SimJava framework to create a discrete event temporal based stochastic model and simulation of the influenza life cycle. The scope of the present simulation of FluSim is the influenza virus pathways without considering the host system's immune system defensive responses. Utilizing the generic SimJava sim_entity class for the various FluSim influenza life cycle stages and molecule types we link them together using SimJava sim_ports.

6.2 Stages

Before constructing the FluSim model great consideration was given to how best utilize the SimJava framework we selected as well as the level of detail we wanted to model based upon our previous work with the Reactome influenza life cycle pathways. We ultimately wanted to find the best compromise between the details necessary to accurately capture the dynamics of influenza infection while abstracting enough to create the best model possible.

We began by examining the various stages of influenza life cycle and decided to model the following steps: Binding, Internalization, Actin-Dependant Transport, Dynein-Directed Transport, End-Directed Transport, Fusion and Uncoating, RNP

Nuclear Import, mRNA Transcription, Viral Protein Translation, Protein secretion, M1 cytoplasmic translocation, cRNA synthesis, Viral Protein Nuclear Import, vRNA Replication, RNP Assembly, RNP nuclear export from the nucleus, Virion assembly and Release of the virion from the host cell. We believe that these 19 steps provide the necessary and sufficient detail to model influenza life cycle dynamics. A schematic of the FluSim model may be seen in Figure 6.1.

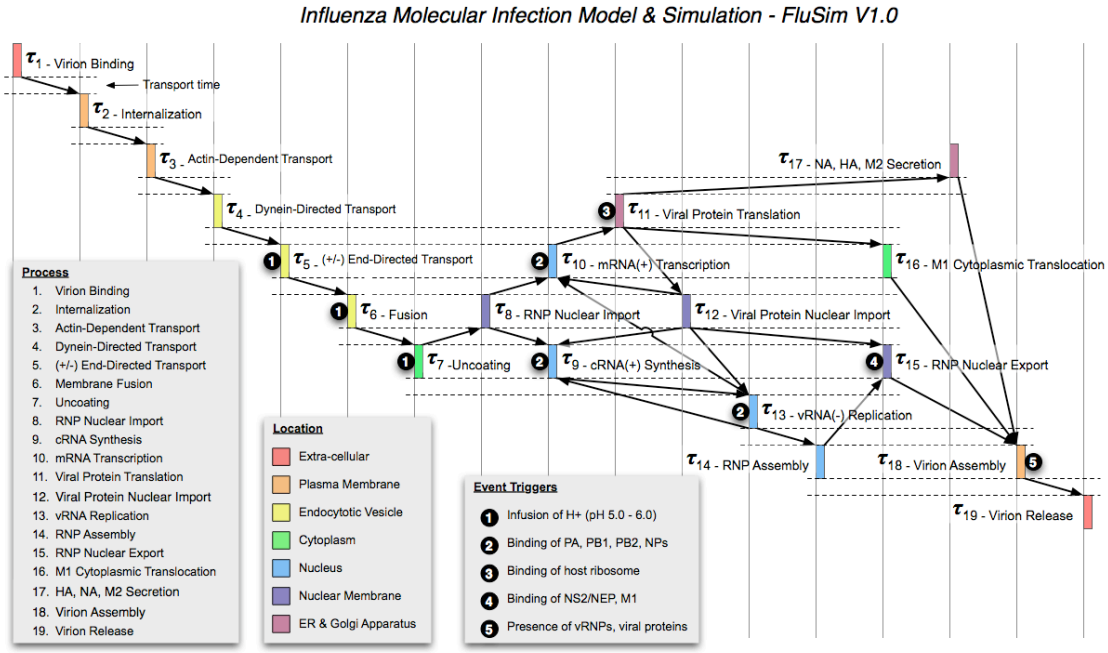


Figure 6.1. The FluSim influenza life cycle model showing the twelve stages of the influenza virus life cycle as well as direction of flow of messages within the model.

Each FluSim stage is a separate class within the FluSim package that extends the generic SimJava sim_entity class. All of the logic of each stage is incorporated into the stages class. For example, the primary transcription of the vRNP into mRNA upon initial entry into the host cell nucleus and the subsequent synthesis of cRNA from the

same vRNP is coded into the RNP nuclear import class. This is accomplished by scheduling primary transcription of the vRNP within the vRNP_import class to occur almost immediately after arriving in the vRNP_import stage echoing the arrival of the vRNP into the nucleus. The vRNP_import class then schedules cRNA synthesis after the primary mRNA transcription. The total synthesis time used for the FluSim model 10 hours post infection or 36,000 seconds at which time virions have been observed as have bursting cells. Further resources have carried the time of infection to 24 hours. (Zdanov and Bukrinskaja 1969; Lars Mˆhler 2005).

6.3 Molecules

A primary goal of our work is to match molecular output with experimental evidence. In order to do this we needed to know the counts of each molecule type which we further broke down into each particular molecule.

To this end we implemented code to count each molecule type and each particular molecule within each type. These new counters included representations for the viral ribonucleoprotein (vRNP), the viral RNA (vRNA), the complementary RNA (cRNA) as well as the messenger RNA (mRNA) and finally proteins. Within each stage particular molecular details are encoded. The process time for transcription is calculated based upon the length of the sequence divided by the rate of transcription (50 base pairs per second) and synthesis rate is based upon the length of the sequence divided by the synthesis rate (800 base pair per second) with the major contributor to each being the time for the polymerase initiation (Berg, Tymoczko et al. 2002).

6.4 Ports

Once we have defined the stages we then connect them to allow events to pass from one stage to another appropriate stage. To accomplish this SimJava uses the concept of ports. Each class has input as well as output ports where appropriate.

Table 6.1. The tags encoded within FluSim allowing for the passing of events to multiple stages for each of the eight RNAs or the ten proteins.

Molecule	Place holder	Destination Stage	Pathway	RNA or Protein Number
NS RNP Segment to mRNA Transcription	1	10	1	08
NA protein to secretion stage	1	17	1	06

In order to differentiate between vRNP entities sent to the mRNA synthesis stage for initial transcription and a vRNP entity sent to the cRNA_synthesis stages we have created a systematic way of describing the stages and molecular entities within FluSim. We begin by assigning each state an identification number or as it is known in SimJava a “tag” that is a number starting at 100,000. The ten thousands place indicates the destination stage number as shown in Figure 6.1. The hundreds place indicates the outgoing port that directly links to an incoming port on the recipient. Utilizing these six digit numbers as a base then we add the segment or protein number to each of the state tags. The vRNP from vRNP_import directed towards mRNA_synthesis state becomes 108101 – 108108 with the PB2 vRNPs tag being 108101. A problem arises when there is more then one destination for a given stage as is the case with vRNP nuclear import, vRNA replication and mRNA transcription stages. To accommodate this we arbitrarily

assign each of the multiple output ports between states a number in hundreds of our tag so the PB2 vRNP destine for the mRNA_synthesis stage is tagged 108101 and the PB2 vRNP destine for cRNA_synthesis is tagged 108201. We assume that molecules destine for a particular state are committed to that state and are not randomly taken up by other states where possible.

6.5 State Logic

The dynamic nature of influenza virus is not completely understood but some experimental evidence does exists to help describe the necessary logic to direct molecules along particular ports as described above..

We have also tried to be accurate as possible with the assembly of complexes and of the budding virion as authentic as possible. To this end the vRNA synthesis state only proceeds if the three polymerase proteins are present as well as enough NP protein to stabilize the newly replicated vRNA. Similarly, the budding state proceeds only when enough proteins and the eight vRNP are present.

In Table 6.1 we see the approximate times for the beginning stages of influenza life cycle as well as the state times for vRNP import and export. We let the simulation computer the budding and release times. Timings for mRNA synthesis and vRNA replication of the 8 vRNP segments of influenza are shown in Table 6.2. The rate of synthesis is 50 base pairs/second and the rate of synthesis also 50 base pair / second (Berg, Tymoczko et al. 2002). The simulation will generate the events of vRNP assembly and Budding event times.

Table 6.2 Processing times for the initial stages of the influenza life cycle.

Event	Value (sec)	Reference
Virus binding time	0	(We assume binding has occurred)
Internalization	60	(Lakadamyali, Rust et al. 2003; Brandenburg and Zhuang 2007)
Actin-Dependant transport	240	As above
Dynein dependant transport time	10	As above
End directed transport	300	As above
Fusion & Uncoating	1800	(Sieczkarski and Whittaker 2003)
vRNP Import	600	(Martin and Helenius 1991)
vRNP Export	600	(Martin and Helenius 1991)

Table 6.3 Processing times for mRNA synthesis and vRNA replication of the 8 vRNP segments of influenza virus.

Segment	Length	Initiation (sec)	mRNA Synthesis (sec)	vRNA replication (sec)
PB2	2341	46.0	46.82	46.82
PB1	2341	46.0	46.82	46.82
PA	2233	46.0	44.66	44.66
HA	1775	46.0	35.5	35.5
NP	1565	46.0	31.3	31.3
NA	1413	46.0	28.26	28.26
M	1027	46.0	20.54	20.54
NS	890	46.0	17.8	17.8

The times of transcription and synthesis are different for each vRNP due to the length of the viral RNA segments just as the translation times will be different for each of the 10 proteins based upon the rate of translation (18 amino acid residues/second) and the length of each protein.

Table 6.4 Processing times for influenza mRNA translation.

Protein	Initiation (sec)	Value (sec)	Reference
PB2	50.0	42.17	

PB1	50.0	42.06	18 amino acids residues/second. (Edwards, Kane et al. 1991; Darzacq, Singer et al. 2005)
PA	50.0	39.78	
HA	50.0	31.44	
NP	50.0	27.67	
NA	50.0	25.22	
M1	50.0	14.00	
M2	50.0	5.39	
NS1	50.0	12.78	
NS2	50.0	6.72	

6.6 Assumptions

The process of modeling as described in Chapter 5 describes the need for abstraction in the modeling process. Abstraction allows for the most important aspects of the system being modeled to be the primary influenza on the output of the model and simulation. In order to abstract such a system numerous assumptions must be made. Here we describe the assumptions that went into the design of the FluSim model.

Our assumptions include: an excess of host cell components including nucleotides, amino acid molecules, ribosomes, others cellular molecules not directly modeled, the stability of the host cell (no apoptosis or cell death), a lack of cross infection by multiple influenza virus to the same host cell. We further assume a negligible host immune response to influenza infection.

6.7 Limitations

In this endeavor we have strived to create a model of the influenza life cycle dynamics that closely mirrors a true infection. We have achieved in designing a model that represents the various state of influenza life cycle as well as some of the major

molecular entities. Despite our best efforts limitations in the modeling system, the Java programming language and hardware prevent a true implementation of the current model. Limitations include restricting the degradation of RNA molecules to 2 iterations much less then expected of a eukaryotic system (Alberts, Bray et al. 2002). I believe that many of these limitations can be overcome by a parallel implementation of the same framework and model or by reviewing the field of frameworks and choosing a new framework that can accommodate the necessary resources.

In the absence of such limitations we believe that we would see an earlier arrival of virion production (through parallel treatment of the simulations events) as well as better distributed output.

CHAPTER 7

RESULTS

To simulate molecular dynamics of influenza infection in a host cell we have constructed a discrete event, stochastic model. Utilizing the SimJava v2 framework we are able to model the nineteen states of influenza life cycle along with the five classes of viral molecules involved in influenza life cycle (vRNA, cRNA, mRNA, vRNP, proteins) as well as the production of virions. In the chapter we will discuss the results of the model and simulation and how they correspond to known measurements as well as how they represent previously undetermined measurements.

7.1 FluSim Molecular Measurements

Utilizing the model described in Chapter 6, we developed a SimJava implementation of the model. The simulation simulates the dynamics of the different stages and captures the molecular counts and then outputs a graph for each of the measured viral molecule types including mRNA, vRNA, cRNA, vRNPs, proteins and virions produced. Experimental evidence provides evidence on the relative concentrations for viral mRNA, cRNA and vRNA as seen in Figure 7.1. Influenza researchers have measured the concentrations of some influenza molecules including the amounts of cRNA and mRNA produced over time (Vreede, Jung et al. 2004) as well as some individual segments of vRNA and mRNA (Ng, Li et al. 2008) but an extensive

survey of individual segment dynamics over the course of time has not been performed yet. From this limited amount of measurement data we can derive some basic characteristics of the dynamics of the various molecules in real infection and use the information to compare the modeling results in the influenza infection. In Figure 7.1 we see to the left an image of the relative amounts of mRNA, cRNA and vRNA in relation to a know quantity of 5S rRNA during a time course experiment of influenza virus infection with time points taken at 3, 4.5, 6 and 9 hours post infection (Ervin Fodor, personal communication). This data is then plotted to the right. We observe that the level of cRNA increases slightly from 3 to 4.5 hpi and remains stable for the remainder of the infection. Messenger RNA increases exponentially from 2 hpi to 4.5 hpi after which degradation decreases the amount of mRNA over then remaining course of infection. Finally, vRNA increases exponentially from 2 hpi through 6 hpi then levels off for the remainder of infection. The nature of these behaviors we wish to determine in our simulation model.

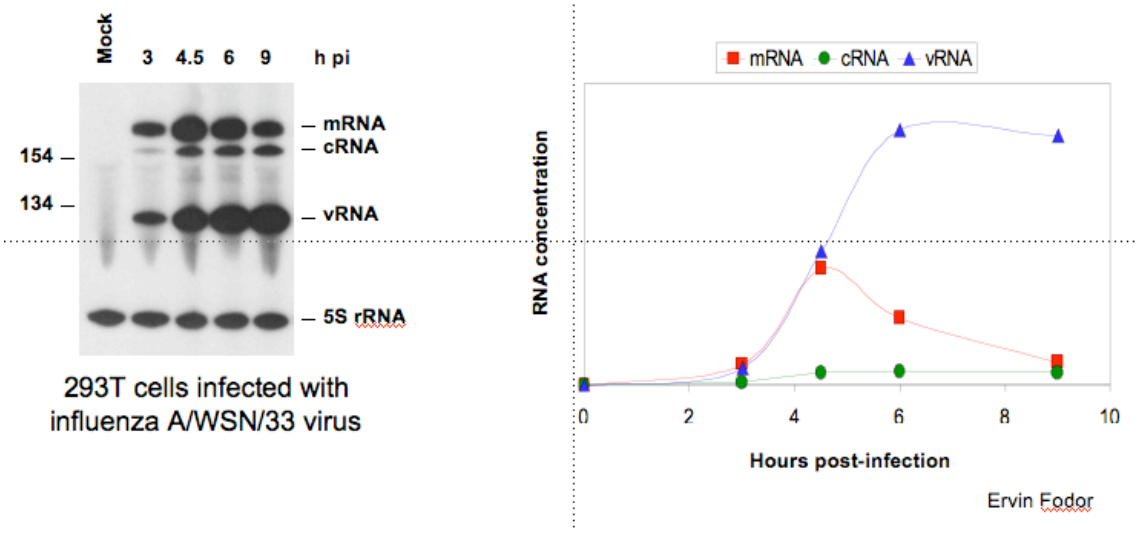


Figure 7.1. Influenza experimental data. To the left we see an image of the relative amounts of cRNA, mRNA and vRNA in influenza infection at time points of 3, 4.5, 6, and 9 hours post infection (hpi). This data is then plotted to the right showing the relative amounts of cRNA, mRNA and vRNA.

7.1.1 mRNA Molecular Measurements

Recall that primary mRNA transcription takes place utilizing the imported vRNP as a template. The mRNA molecules that are produced are exported out of the nucleus and translated into protein through the work of the host cell ribosome machinery and the endoplasmic reticulum (ER). In influenza infection, we observe two mRNA processes, namely mRNA production and mRNA degradation. The viral mRNA molecules, through a process known as "cap-snatching" (Shih and Krug 1996), obtain a nucleotide string or "cap" that facilitates the start of mRNA transcription in addition to providing some stability to the nascent mRNA molecule being transcribed. As the early phase of influenza infection gives way to the later phase the transcription rate levels off one can observe the degradation rate take over as the major contributor to mRNA concentration (Shapiro, Gurney et al. 1987).

In our simulation we see in Figure 7.2 that mRNA transcription shows the characteristic exponential increase and slow decrease that we saw in the experimental results. In our case, while the numbers are close to what is expected, as we shall discuss in our Dizzy comparison, the timing is off since the peak of mRNA production is shown experimentally at about 4.5 hpi. One major reason for the shortened time is due to the small number of precursor mRNA caps in the system. Instead of the 220,000 caps in a normal host cell or 44,000 per virion assuming a multiplicity of infection (MOI) of 5 we have use a 10-fold smaller amount to enable use to view the behavior of the entire system in a reasonable time. Replacing the number of caps and rerunning the system will be done once the system is optimized for speed.

Once mRNA production has leveled off the effects of mRNA degradation can truly be observed both in experimental evidence and in our simulation. In experimental evidence mRNA degradation begins approximately 4 - 5 hpi.

In our results we observe the characteristic exponential increase of mRNA production in the system beginning at approximately 2900 seconds post infection (spi). At approximately 3100 spi we observe the end of the exponential production as the system runs our of precursor mRNA caps from the host to prime viral mRNA synthesis. In order to complete simulations in a reasonable time we have reduced the number of the caps from the expected 44,000 to a value of 4,400. This explains the end of mRNA transcription and the gradual degradation of the mRNAs that occur at a much earlier time then observed experimentally. We expect that returning the number of mRNA caps to the calculated value would help to make FluSim output closer to experimental results.

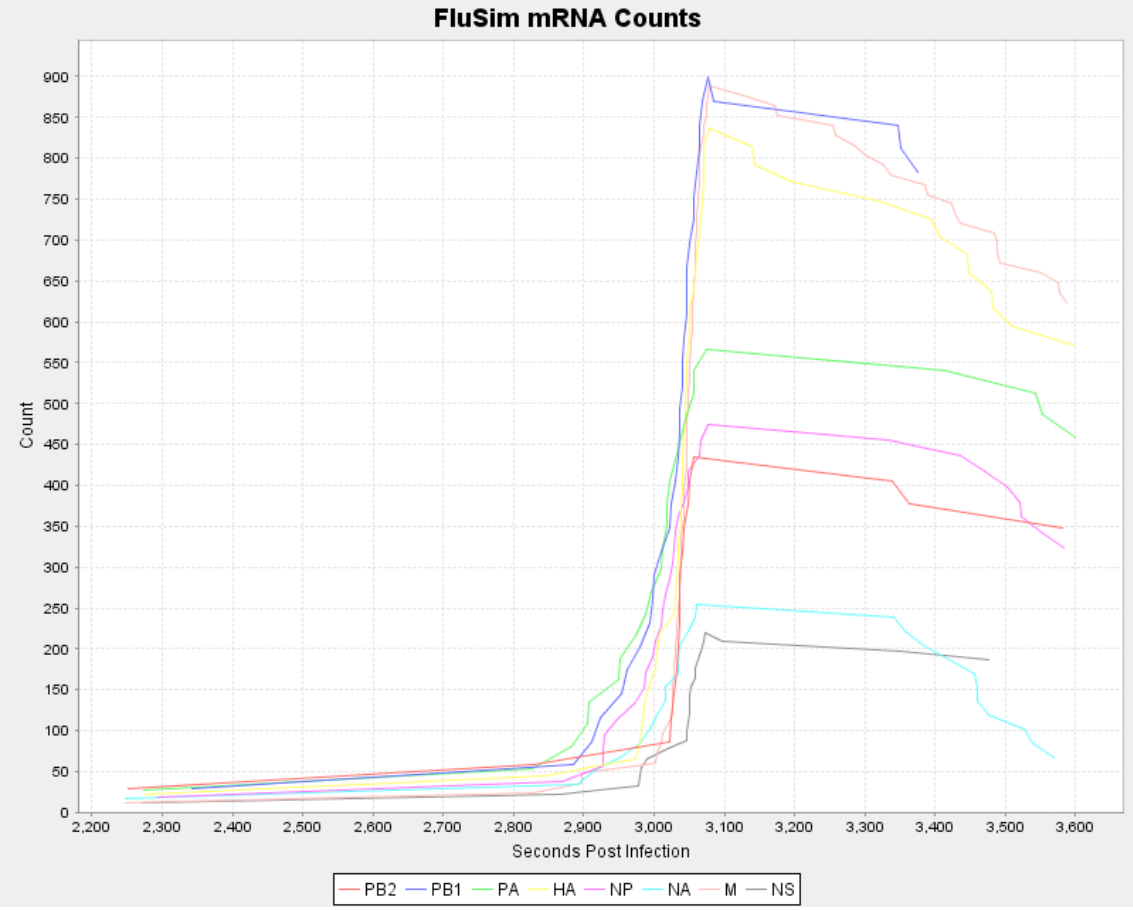


Figure 7.2 mRNA transcriptions results for a single virion infecting a host cell. An increase in mRNA production can be seen beginning around the 3100 seconds hour post infection (spi) then plateaus at approximately 3300 spi. This results is an abbreviate representation of infection as the number of precursor mRNA caps is set at 4400 instead of the expected 44,000 to purposes of speeding the model up. Though the simulation is sped up we observe the similar dynamics of the molecular entities to experimental results. As we shall see these number agree with the Dizzy model.

7.1.2 vRNA Molecular Measurements

Viral RNA is replicated from the positive stranded cRNA utilizing the viral polymerase complex (PA, PB1, PB2) and is stabilized by binding to a nucleoprotein proteins forming viral ribonucleoproteins (vRNPs) (Vreede, Jung et al. 2004). Viral

RNAs then serve as templates for further mRNA or cRNA production as well as for export as vRNP for inclusion in new virions. Figure 7.3 shows vRNA replication in the FluSim model. Viral RNA replication begins about 3000 spi and continues to climb for the remaining of the simulation. While PB2 vRNA replication shows an exponential like growth the remaining vRNAs show slower then expected replication. Taking multiple runs of FluSim while changing the initial random number seed then averaging the values across all runs should give a better representation of the vRNA replication behavior.

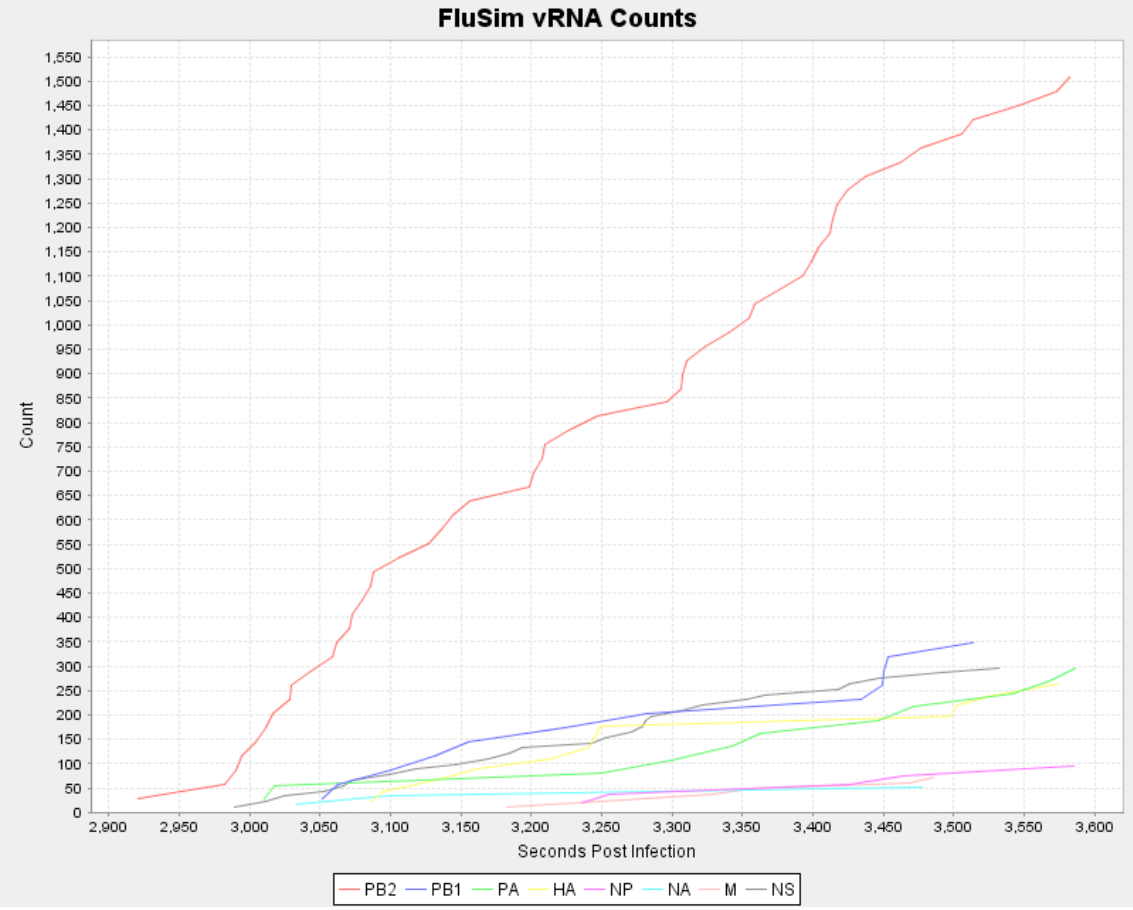


Figure 7.3. Viral RNA production in FluSim.

7.1.3 cRNA Molecular Measurements

Subsequent to mRNA transcription the vRNA / vRNP serve as templates for replication of positive copies of the vRNA in the form of cRNA, which in turn serve as templates for production of further vRNA. Initially scientists proposed that cRNA production began once there was sufficient concentration of NP protein in the infected cell but currently the stabilization model proposed by Vreede, et al (Vreede, Jung et al. 2004) is the preferred model for understanding cRNA replication in an infected cell. In the stabilization model, cRNA is replicated immediately following initial transcription of the imported vRNP. However, these initial cRNA molecules are degraded since they lack the stability provided by being bound to the nucleoprotein complex. We utilize this stability model in our system.

Complementary RNA synthesis rates from our simulation are shown in Figure 7.4. In our model cRNA rates were set to mirror the experimental results as described above with cRNA synthesis rates being less then vRNA synthesis rates. We observe experimentally that cRNA synthesis is less then vRNA replication by a 5 or 10 fold difference. This is also observed in our model with the cRNA synthesis mean being approximately 10 fold lower then vRNA replication.

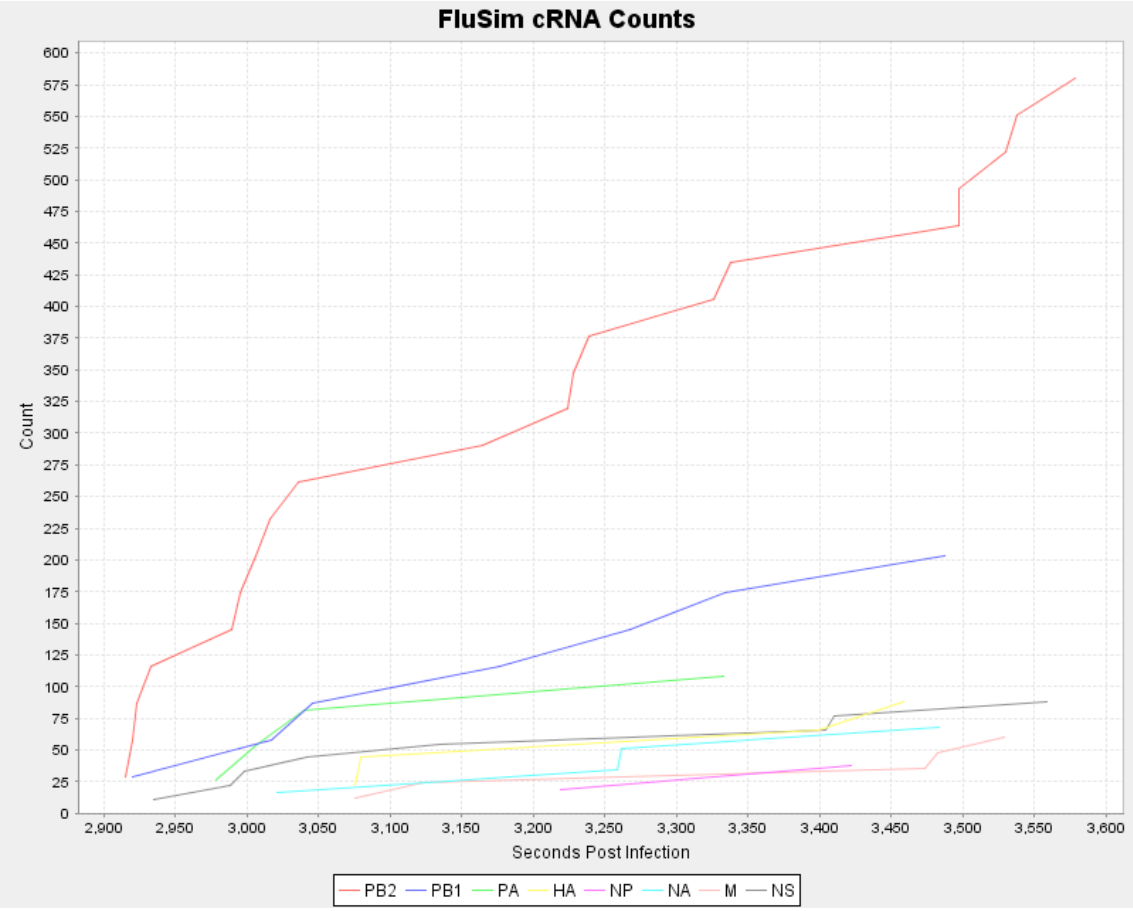


Figure 7.4 Complementary RNA results of FluSim for each of the eight influenza segments. The cRNA replication rate is approximately 1/5th to 1/10th the vRNA rate, which mirrors experimental results. There is a great variety to the various segment rates that demonstrates the stochastic nature of the model.

7.1.4 vRNP Molecular Measurements

The first of experimentally undetermined molecular measurements is that of the viral ribonucleoprotein. What is knows is that at six hours post infection the presences of vRNPs can be detected while virions cannot (Enami, Qiao et al. 1993). The vRNPs are exported out of the nucleus for packaging into the virions for budding and release. We see in our simulation in Figure 7.5 that the vRNP production begins at

approximately 3000 spi and continues linearly for the remainder of the simulation. This is short of the experimental evidence that shows that vRNP are first visible at 6 hpi. It remains to be determined what behavior causes the shift from cRNA synthesis and mRNA transcription to vRNP production. Simple the presence of M1 protein carrying the export signal may not be the case as that logic is built into this system.

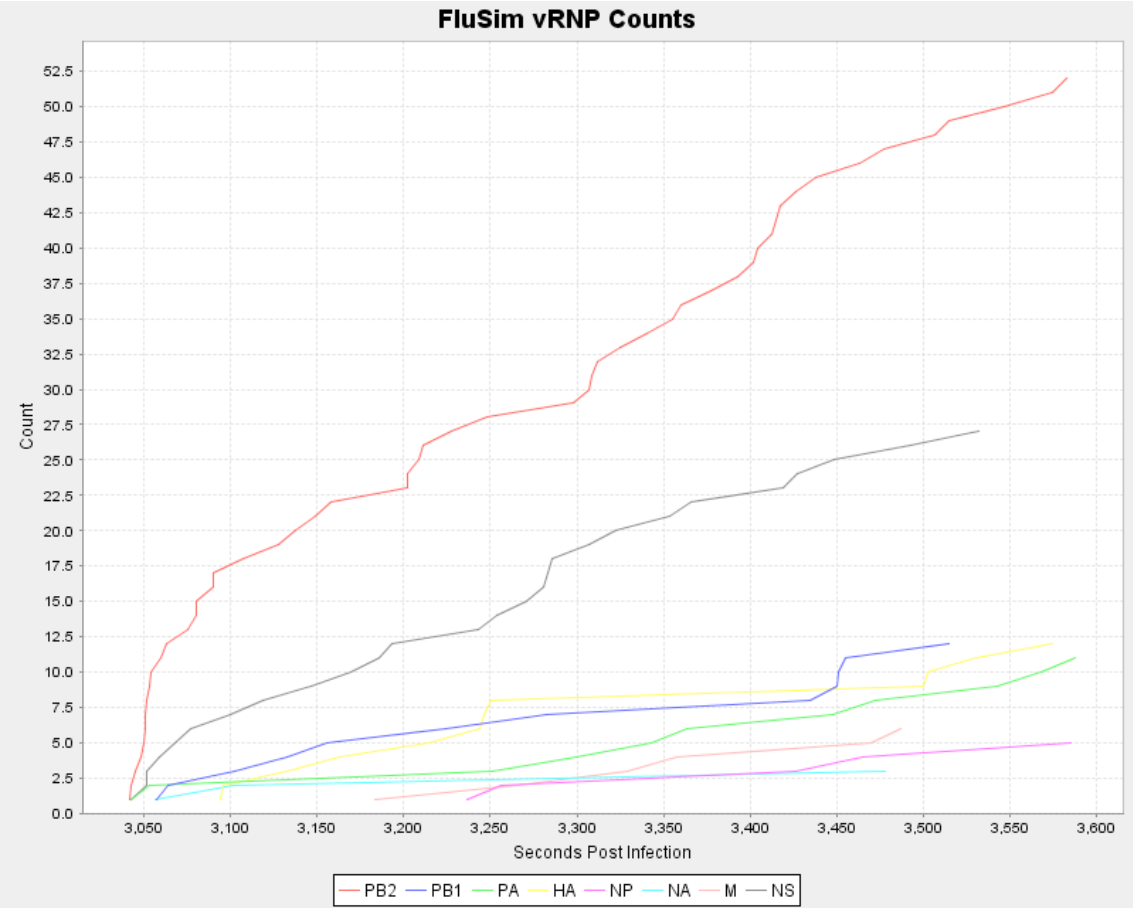


Figure 7.5 Viral ribonucleoprotein results in FluSim begins approximately four hours post infection and continues linearly throughout infection. Viral RNPs are produced before being exported for packaging into a new virion.

7.1.5 Simulation Comparison with Experimental Results

The experimental results we have obtained present relative amount of mRNA, cRNA and vRNA. Therefore it is not possible to do a quantitative comparison to our results. We can however do a qualitative comparison as seen in Table 7.1. Our comparison shows that our simulation produces the same relative amounts of mRNA and cRNA but falls short in the production of vRNA. A better comparison can be made once multiple iterations of our simulation are performed and averaged.

Table 7.1 Relative amounts of RNA

Molecular ratio	*Experiment Results*	*Simulation Results*	*Remarks*
mRNA/mRNA	*1*	*1*	
cRNA/mRNA	*1/5*	*100 / 500*	*Match*
vRNA/mRNA	*2*	*500 / 500*	*vRNA low*

7.1.6 Protein Molecular Measurements

Protein production is measured in FluSim for each of the 10 proteins. For simplicity sake the newly discovered protein PB1-F2 has been omitted from the model. Figure 7.6 shows the protein translation rate of the simulation. In our case the output of protein production varies greatly. First and foremost the usage of proteins is incorporated in the system so the figure below shows the level of unutilized protein in the system at any given time. Additionally, we know from previous experiments that the production rate of the M2 andNS2/NEP protein is approximately 10% of the primary transcript, which accounts for the lowest production rates in our model.

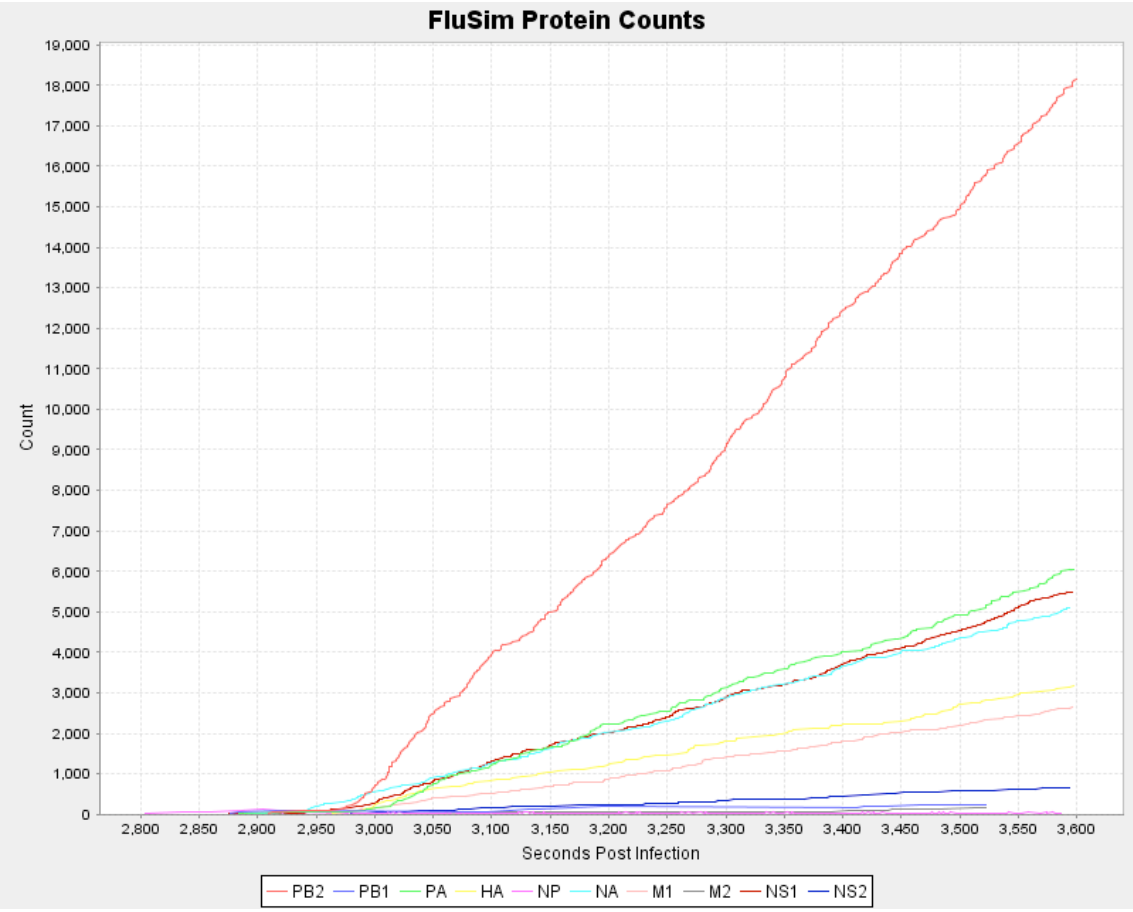


Figure 7.6 Protein results in FluSim is shown. One can observe the stochastic nature of the simulation in the great variety of start times for each of the protein translations. Given the approximate number of proteins required per virion the protein production seems under-representative of required values.

7.2 Virion Measurements

The ultimate measurement of the infectivity of a virus is its ability to reproduce providing the capability to infect other cells and ultimately other hosts. In our model we check for the presence of the eight distinct vRNPs as well as the presences of the right number of proteins in a virion before a virion is produced. At this time, our system is

not producing virions due to the lack of proteins. We expect that the proposed increase of host mRNA caps and the resulting mRNA increase and subsequent protein increase will fix this once again allowing our system to produce virions. For illustrative purposes virions production from a previous iteration of the system is shown in figure 7.7.

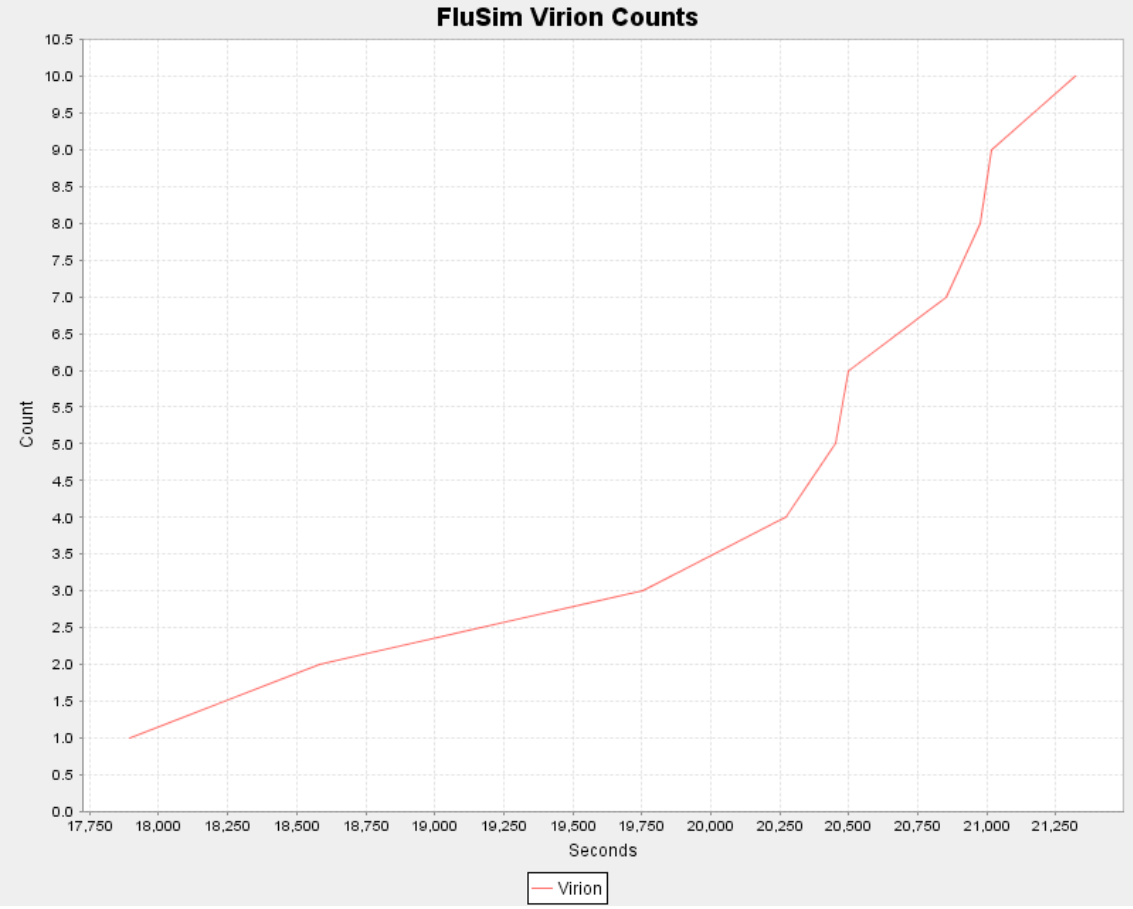


Figure 7.7 Virion particle release rates in FluSim. Each virion is released only after each of the required vRNP is present and an approximate number of proteins found in each virion is present. At this time we are not subtracting the proteins from the counts considering the low protein translation rates. With the increase in protein production rates, removing the proteins as virions are produced will allow for true virion production rates.

7.3 Simulation Runtime

These results were obtained by running the simulation for 1 hour of simulation infection time. We strive to extend the simulation time to 10 hours post infection to represent the period of influenza infection known to have visible virions. It is important to note that our results reflect the molecule infection process resulting from a single virion infecting a host cell. In multiplicity of infection (MOI) or the number of virions infecting a single host cell range in experiments ranges from 0.5 to 50. (Chu and Whittaker 2004; Jones, Turpin et al. 2006) We look forward to expanding our capacity to model increased MOI as well as increased times in the future.

7.4 FluSim Quantification

In an effort to compare our discrete event stochastic simulation qualitatively we have implemented a simplified model using the Dizzy software (URL). The Dizzy software is chemical kinetics simulation software package designed by Stephen Ramsey of the Institute for Systems Biology (ISB). The Dizzy software provides one with the capability to design a model of a system, such as influenza infection in a host cell, and then simulate it using any of the included simulations engines. In our case to we used the implementation of the Gillespie stochastic algorithm for simulating chemical reaction kinetics. The Gillespie stochastic algorithm is an appropriate for modeling the kinetics of a set of coupled chemical reactions, taking into account stochastic effects from low copy numbers of the chemical species. (Ramsey, Orrell et al. 2005)

7.4.1 Dizzy Components and Reactions

To setup a Dizzy model, one defines the components of the system (in our case the various RNA molecule types, proteins as well as basic components of the host cell such a pre-mRNAs) and then one utilizes these components in simplified reactions of each of the major contributing reactions to the model. We see in Table 7.2 the components defined in the Dizzy model and in Table 7.3 we see the reactions in our Dizzy model.

Table 7.2 The Dizzy model components.

Dizzy Components	**Description**
vRNA(1…8) = 0	Viral RNA
vRNPN(1…8) = 0;	Initial Viral ribonucleoproteins
P(1…10) = 0;	Proteins
cRNA(1…8) = 0;	Complementary ribonucleoproteins
mRNA(1…8) = 0;	Viral mRNAs
Cap = 4400;	Host precursor mRNA caps

Table 7.3 The reactions of our Dizzy model. The reaction number is shown in column one, the reaction itself in column two and the rate equation in column three.

Primary transcription		
r1	vRNP1 + cap -> vRNP1 + mRNA1	1/(46 + 47) * 0.95;
r2	vRNP2 + cap -> vRNP2 + mRNA2	1/(46 + 47) * 0.95;
r3	vRNP3 + cap -> vRNP3 + mRNA3	1/(46 + 45) * 0.95;
r4	vRNP4 + cap -> vRNP4 + mRNA4	1/(46 + 36) * 0.95;
r5	vRNP5 + cap -> vRNP5 + mRNA5	1/(46 + 31) * 0.95;
r6	vRNP6 + cap -> vRNP6 + mRNA6	1/(46 + 28) * 0.95;
r7	vRNP7 + cap -> vRNP7 + mRNA7	1/(46 + 21) * 0.95;
r8	vRNP8 + cap -> vRNP8 + mRNA8	1/(46 + 18) * 0.95;

Table 7.3 Continued.

mRNA degradation		
r11	mRNA1 ->	0.0005;
r12	mRNA2 ->	0.0005;
r13	mRNA3 ->	0.0005;
r14	mRNA4 ->	0.0005;
r15	mRNA5 ->	0.0005;
r16	mRNA6 ->	0.0005;
r17	mRNA7 ->	0.0005;
r18	mRNA8 ->	0.0005;
Translation		
r21	mRNA1 -> mRNA1 + P1	(1/(50 + 42));
r22	mRNA2 -> mRNA2 + P2	(1/(50 + 42));
r23	mRNA3 -> mRNA3 + P3	(1/(50 + 40));
r24	mRNA4 -> mRNA4 + P4	(1/(50 + 31));
r25	mRNA5 -> mRNA5 + P5	(1/(50 + 28));
r26	mRNA6 -> mRNA6 + P6	(1/(50 + 25));
r27	mRNA7 -> mRNA7 + P7	(1/(50 + 14));
r28	mRNA7 -> mRNA7 + P8	(1/(50 + 5)) * 0.1;
r29	mRNA8 -> mRNA8 + P9	(1/(50 + 13));
r20	mRNA8 -> mRNA8 + P10	(1/(50 + 7)) * 0.1;
cRNA synthesis		
r41	P1 + P2 + P3 + P5 + vRNA1 -> cRNA1 + vRNA1	(1/(46 + 47)) * 0.05;
r42	P1 + P2 + P3 + P5 + vRNA2 -> cRNA2 + vRNA2	(1/(46 + 47)) * 0.05;
r43	P1 + P2 + P3 + P5 + vRNA3 -> cRNA3 + vRNA3	(1/(46 + 45)) * 0.05;
r44	P1 + P2 + P3 + P5 + vRNA4 -> cRNA4 + vRNA4	(1/(46 + 36)) * 0.05;
r45	P1 + P2 + P3 + P5 + vRNA5 -> cRNA5 + vRNA5	(1/(46 + 31)) * 0.05;
r46	P1 + P2 + P3 + P5 + vRNA6 -> cRNA6 + vRNA6	(1/(46 + 28)) * 0.05;
r47	P1 + P2 + P3 + P5 + vRNA7 -> cRNA7 + vRNA7	(1/(46 + 21)) * 0.05;
r48	P1 + P2 + P3 + P5 + vRNA8 -> cRNA8 + vRNA8	(1/(46 + 18)) * 0.05;
cRNA degradation		
r141	cRNA1 ->	0.0012;
r142	cRNA2 ->	0.0012;
r143	cRNA3 ->	0.0012;
r144	cRNA4 ->	0.0012;
r145	cRNA5 ->	0.0012;
r146	cRNA6 ->	0.0012;
r147	cRNA7 ->	0.0012;
r148	cRNA8 ->	0.0012;

Table 7.3 Continued.

vRNA synthesis		
r251	P1 + P2 + P3 + P5 + cRNA1 -> cRNA1 + vRNA1	(1/(46 + 47));
r252	P1 + P2 + P3 + P5 + cRNA2 -> cRNA2 + vRNA2	(1/(46 + 47));
r253	P1 + P2 + P3 + P5 + cRNA3 -> cRNA3 + vRNA3	(1/(46 + 45));
r254	P1 + P2 + P3 + P5 + cRNA4 -> cRNA4 + vRNA4	(1/(46 + 36));
r255	P1 + P2 + P3 + P5 + cRNA5 -> cRNA5 + vRNA5	(1/(46 + 31));
r256	P1 + P2 + P3 + P5 + cRNA6 -> cRNA6 + vRNA6	(1/(46 + 28));
r257	P1 + P2 + P3 + P5 + cRNA7 -> cRNA7 + vRNA7	(1/(46 + 21));
r258	P1 + P2 + P3 + P5 + cRNA8 -> cRNA8 + vRNA8	(1/(46 + 18));
Viral mRNA transcription		
r361	vRNA1 + cap -> vRNA1 + mRNA1	(1/(46 + 47)) * 0.95;
r362	vRNA2 + cap -> vRNA2 + mRNA2	(1/(46 + 47)) * 0.95;
r363	vRNA3 + cap -> vRNA3 + mRNA3	(1/(46 + 45)) * 0.95;
r364	vRNA4 + cap -> vRNA4 + mRNA4	(1/(46 + 36)) * 0.95;
r365	vRNA5 + cap -> vRNA5 + mRNA5	(1/(46 + 31)) * 0.95;
r366	vRNA6 + cap -> vRNA6 + mRNA6	(1/(46 + 28)) * 0.95;
r367	vRNA7 + cap -> vRNA7 + mRNA7	(1/(46 + 21)) * 0.95;
r368	vRNA8 + cap -> vRNA8 + mRNA8	(1/(46 + 18)) * 0.95;

7.4.2 Comparison of FluSim discrete event simulation results to Dizzy results

We can see in Figure 7.8 the Dizzy simulation results for cRNA in our model. The comparison of cRNA synthesis is summarized in Table 7.4. We observe that the median value of 100 matches exactly between the Dizzy model and FluSim model while the ranges of 25 to 600 cRNAs match almost exactly as well.

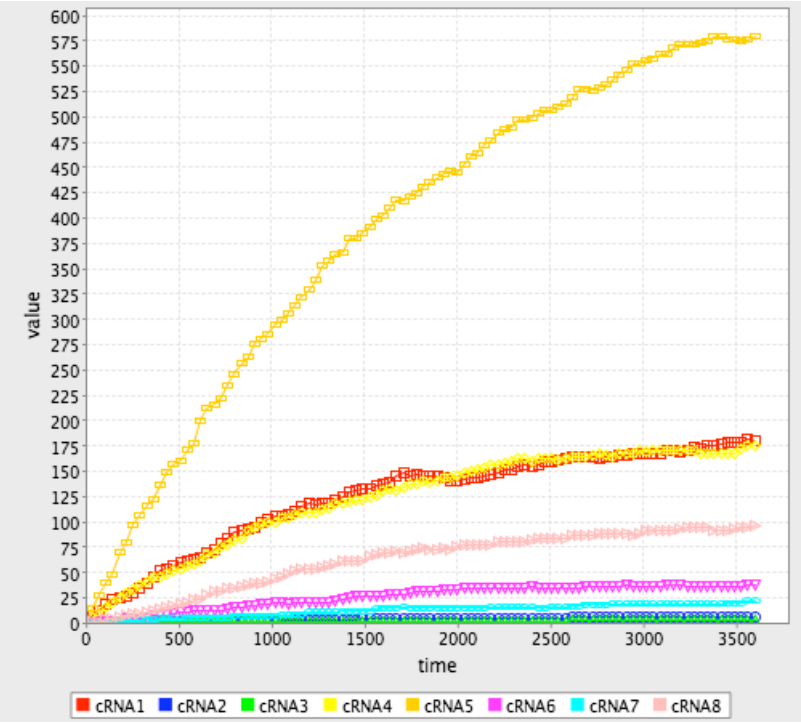


Figure 7.8 Dizzy model output for each of the eight cRNA segments based upon the Gillespie stochastic algorithm. Time is seconds.

Table 7.4 Dizzy and FluSim cRNA Comparison.

	Dizzy	FluSim
Minimum	~25	~35
Maximum	~600	~600
Median Value	~100	~100

Viral RNA (vRNA) production is shown in Figure 7.9. We observe a range from 25 to ~2200 vRNAs produced with a median value of 300. These value matches closely with the FluSim values of 50 to 1500 for the range and a median value of 300 vRNAs produced. Through multiple iteration of the simulation it was observed that the median value remains the same.

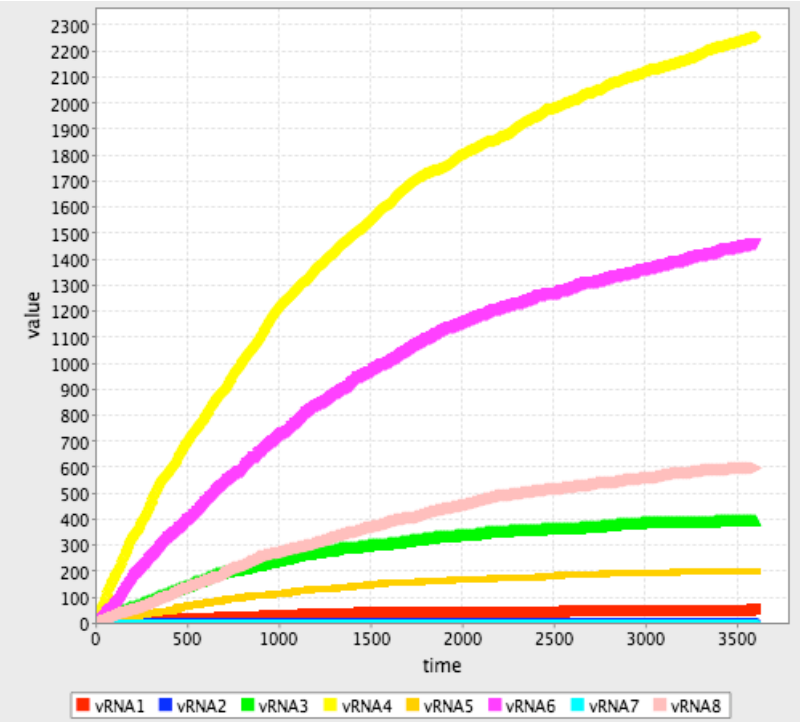


Figure 7.9 Dizzy model output for each of the eight vRNA segments based upon the Gillespie stochastic algorithm. Time is seconds.

Table 7.5 Dizzy and FluSim vRNA Comparison.

	Dizzy	FluSim
Minimum	~25	~50
Maximum	~2200	~1500
Median Value	~300	~300

Messenger RNA value comparison is shown in Table 7.6. In this table we see that the mean values is approximately the same (500 versus 450) while the ranges is a little tighter for Dizzy results (400 – 1050) the results then the FluSim simulation (200 – 900) results.

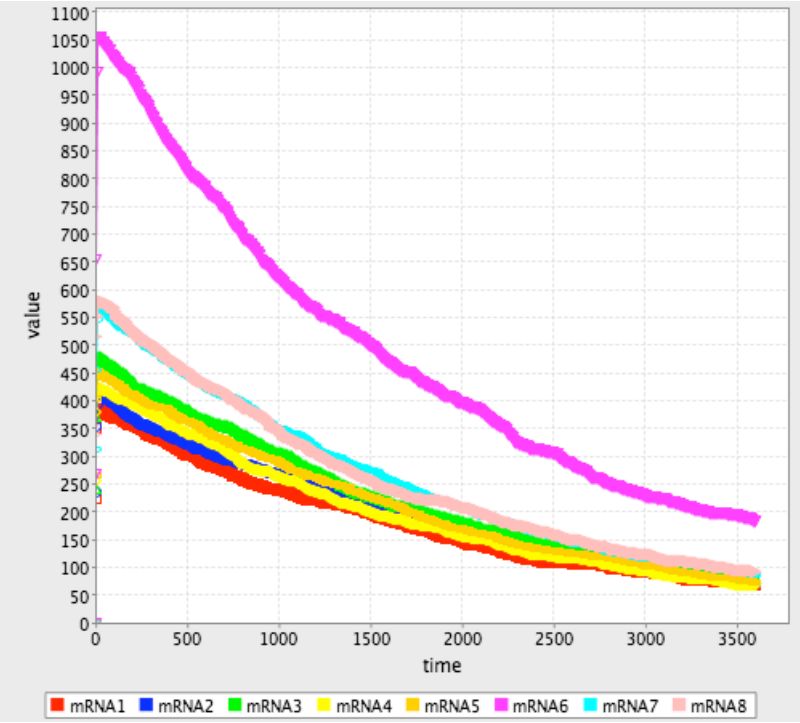


Figure 7.10 Dizzy model output for each of the eight mRNA segments based upon the Gillespie stochastic algorithm. Time is seconds.

Table 7.6 Dizzy and FluSim mRNA Comparison.

	Dizzy	FluSim
Minimum	~400	~200
Maximum	~1050	~900
Median Value	~500	~450

In summary the current FluSim discrete event closely matches the numbers generated in the Dizzy model, which confirms that the FluSim source code is accurately modeling the expected reactions. Where both the models fall short is the fact that the timing of both shows.

7.4.3 Protein Count Comparison with Dizzy Simulation

Protein production counts in the FluSim and Dizzy simulations vary greatly both between individual proteins and between individual simulations iterations. The counts in Table 7.7 are an average of two iterations of each of the simulations. In order to truly gauge the behavior of our simulations we first must perform enough iterations to obtain statistically significant results for our output. Despite the fact that we only have a couple iterations of each simulation, we can observe some expected behaviors. We observe that the splice products of the matrix and non-structural proteins are produced at approximately 10% of the primary transcript as programmed. Additionally in most cases we observe smaller counts for the polymerase proteins and the nucleoprotein. Furthermore, the difference between the FluSim and Dizzy NP results can be accounted for by the fact that the specific amount of NPs required for each protein are coded into the FluSim model while a single NP is required in the Dizzy model.

Table 7.7 Protein count comparison.

Proteins	Dizzy Avg	FluSim Avg
PB2	1250	10750
PB1	300	4250
PA	1650	5625
HA	12000	5500
NP	5125	100
NA	16250	3750
M1	15500	6250
M2	1775	750
NS1	15000	3750
NS2	1750	500

7.5 Conclusions

In conclusion we believe that we have presented a model of influenza infection and a simulation of that model using discrete event simulation technique. The current model only includes the virus infection paths. We have shown that it is possible to model and simulate influenza infection at the molecular level and that the nineteen stages in our model represent the major contributing factors to influenza infection in a host. The immune system response due to virus infection is not part of the current model and that is for future work. The qualitative behavior of the molecular dynamics of few molecules is compared with the published experimental results. To validate the quantative results of the simulation, a separate model using the similar pathways were developed using Gillespie simulation. The quantities comparison is not very satisfactory.

It is out hope that the development of this system that will provide the framework for future models as we strive to develop a cycle of modeling, simulation and laboratory testing to best utilize computational models and simulation in the larger scope of understanding influenza infection and viral infection in general.

APPENDIX A

THESIS COMMITTEE QUESTIONS AND REPLY

Thesis Defense Concerns and Questions
April 12, 2008

1. Considering expectations.

There are numerous aspects to considering expectations. First of all, expectations should be considered before reviewing the results. Expectations should be considered for each important aspect of the model and simulations. The aspects to consider include:

- Molecular behavior (vRNA, vRNA, cRNA, mRNA, protein molecules)
  - First appearance
  - Generation rate
  - Degradation rate
- Virion
  - First appearance
  - Generation rates
- Experimental milestones (
  - vRNP first detectable at 6 hpi
  - Virions first detectable at 10 hpi
- The first appearance of each stage in the simulation
- The length of each stage in the simulation

Once expectations have been defined, one should then consider if the current simulation results meet the expectations. If the simulation results do not meet expectations consider why not? One should consider whether the model is wrong or are the expectations wrong?

To evaluate if the simulation is wrong, multiple runs of the simulation should be performed and an average taken. The average should be compared with expectations. If the average is different compared to expectations, then the individual stage or molecules that do not meet expectations should be investigated. This investigation should include reviewing what knowledge went into the model design as well as a new review of the particular aspect of the system to look for missed information.

Expectations can also be confirmed or refined based upon a review of the literature.

2. Viral mRNA Splicing (M, NS segments)

   The splicing of segment 7 and 8 mRNAs results in the M2 and NS2/NEP proteins respectively. These are what I was mistakenly calling these "secondary transcript." Splicing the original transcript and excising an intron produces them. What is important is that I think the implementations in FluSim is correct as they are in the minority (10%) of the primary transcript.

3. Host precursor mRNA molecules

   To avoid confusion between the virion mRNA molecules host precursor mRNAs should be referred to as "caps" in light of the process of the viral polymerase complex removing the 5' methyl cap from the host precursor mRNA being called cap-snatching. The number of caps is import in that they are required for viral mRNA transcription to proceed.

4. mRNA output

   Output of the mRNA molecules shows an exponential increase and then a sudden drop off. The sudden drop off is the result of mRNA degradation rate that is too high. The rate will be corrected in the simulation.

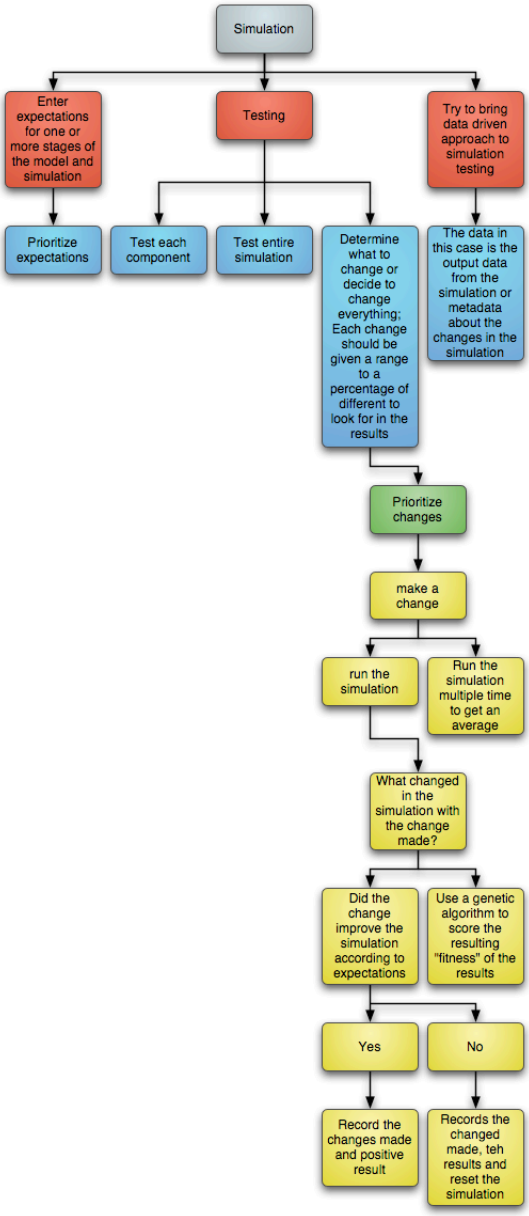
5. Molecule complexes

   Molecule complexes are not considered specifically in the SimJava. I am assuming they are worked into the rates. A complex is built into the logic of the export of the vRNA as vRNPs. Areas where complexes can be considered include: the oligomerization or combining of single HA, NA, M2 proteins into tetramer or trimers, polymerase complex formation, in addition to the import of proteins or vRNP using host proteins.

6. Process of testing model

   Testing of a model and a simulation could be done in a systematic way. The flowchart below describes a possible process for testing a model and simulation.

   Utilize the principles of unit testing from software engineering; each stage of our model could be tested to make sure it behaves as expected. Each stage test would be given scenarios that produce positive and negative output from the state or Java class. Following successful testing of each of the stages the entire system combined could then be tested.



a. In ironic twist could we also use a genetic algorithm to test the "fitness" of each simulation based upon our expectations

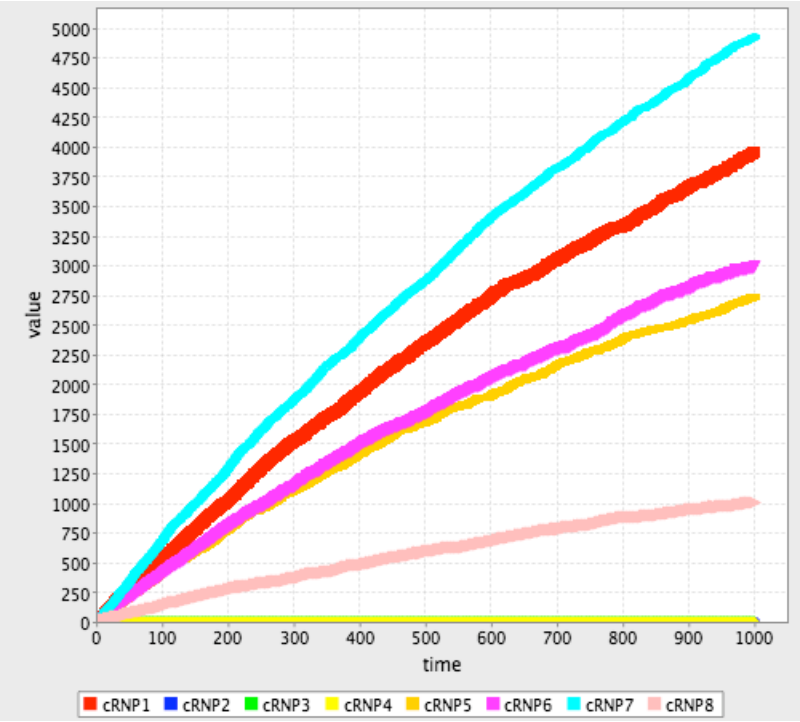
Figure 1. Flow chart of model testing proposal.

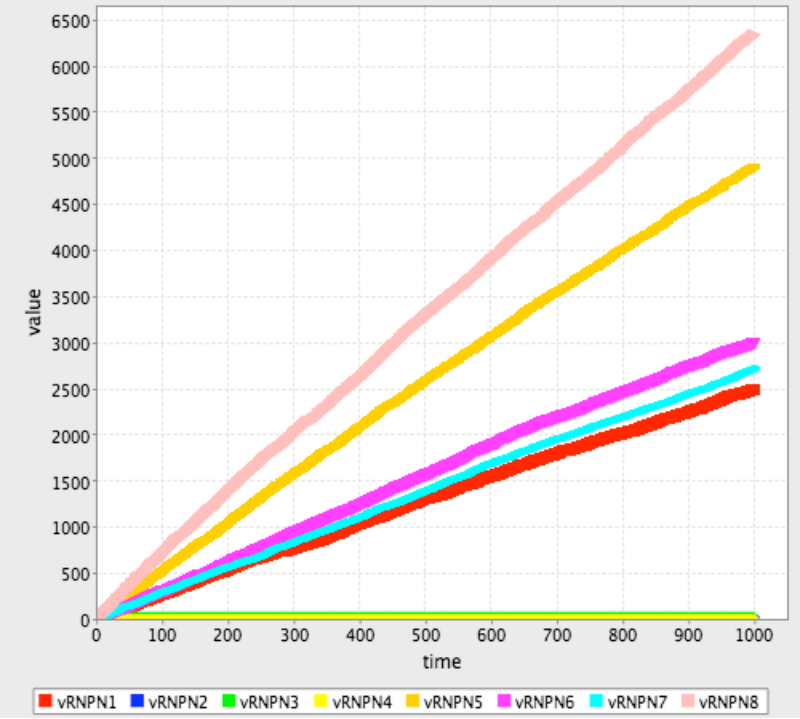
7. Synthetic data

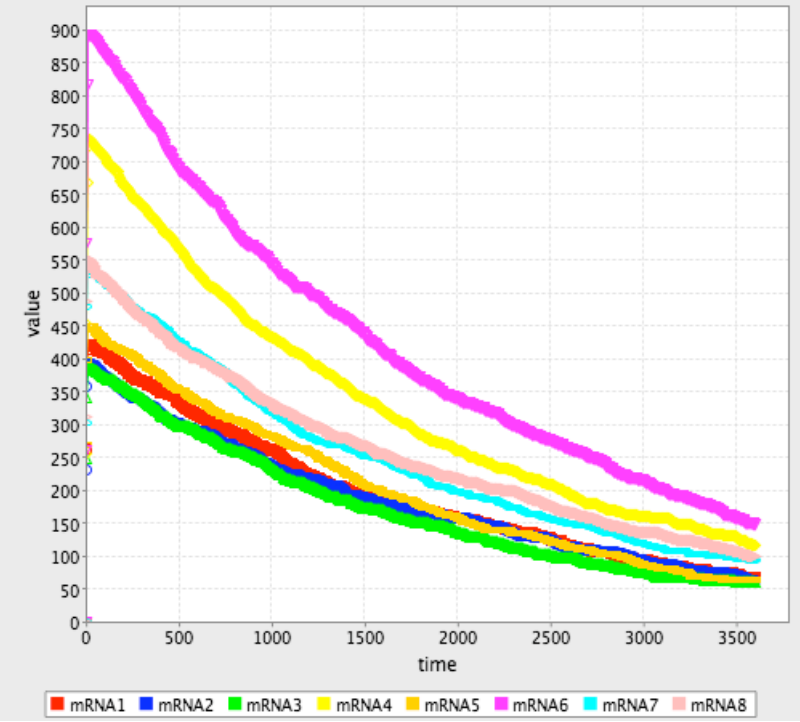
Synthetic data would be useful in testing a simulation if the synthetic data were used for rate values especially as a component of the testing above. The synthetic data could be pseudo-randomly generated within a predetermined lower and upper bound and then fed into the simulation. The simulation could then be tested as described above.

8. Dizzy results

I found an error in the Dizzy file. I have included correct graphs below. I will be working with the Dizzy file and the SimJava file to best mirror the reactions and rates in each. At this time the graphs below show that the polymerase and NP segments are is short supply (with the exception of segment 1) which are reflected in the SimJava model as they are the rate limiting components required for RNA synthesis and replication.




value
time
vRNPN1 vRNPN2 vRNPN3 vRNPN4 vRNPN5 vRNPN6 vRNPN7 vRNPN8


value
time
mRNA1 mRNA2 mRNA3 mRNA4 mRNA5 mRNA6 mRNA7 mRNA8

REFERENCES

(ACIP), A. C. o. I. P. (2006). "Prevention and Control of Influenza Recommendations of the Advisory Committee on Immunization Practices (ACIP)."

Alberts, B., D. Bray, et al. (2002). "The Molecular Biology of the Cell (ed.)." New York: Garland Science.

Amorim, M. J. and P. Digard (2006). "Influenza A virus and the cell nucleus." Vaccine **24**(44-46): 6651-6655.

Baccam, P., C. Beauchemin, et al. (2006). "Kinetics of influenza a virus infection in humans." J Virol **80**(15): 7590-9.

Barsky, A., J. L. Gardy, et al. (2007). "Cerebral: a Cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation." Bioinformatics: btm057.

Beauchemin, C., J. Samuel, et al. (2005). "A simple cellular automaton model for influenza A viral infections." Journal of Theoretical Biology **232**(2): 223-234.

Belshe, R. B. (2005). "The Origins of Pandemic Influenza -- Lessons from the 1918 Virus." N Engl J Med **353**(21): 2209-2211.

Berg, J., J. Tymoczko, et al. (2002). Biochemistry Fifth Edition, WH Freeman and Company.

Boulo, S., H. Akarsu, et al. (2007). "Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes." Virus Research **124**(1-2): 12-21.

Brandenburg, B. and X. Zhuang (2007). "Virus trafficking - learning from single-virus tracking." Nat Rev Micro **5**(3): 197-208.

Bright, R. A., D. K. Shay, et al. (2006). "Adamantane Resistance Among Influenza A Viruses Isolated Early During the 2005-2006 Influenza Season in the United States." JAMA **295**(8): 891-894.

Chu, V. C. and G. R. Whittaker (2004). "Influenza virus entry and infection require host cell N-linked glycoprotein." Proc Natl Acad Sci U S A **101**(52): 18153-8.

Connor, R. J., Y. Kawaoka, et al. (1994). "Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates." Virology **205**(1): 17-23.

Crosby, A. (1989). America's Forgotten Pandemic. Cambridge, U.K., Cambridge Univ. Press.

Darzacq, X., R. H. Singer, et al. (2005). "Dynamics of transcription and mRNA export." Current Opinion in Cell Biology **17**(3): 332-339.

Edwards, A. M., C. M. Kane, et al. (1991). "Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro." J. Biol. Chem. **266**(1): 71-75.

Enami, K., Y. Qiao, et al. (1993). "An Influenza Virus Temperature-Sensitive Mutant Defective in the Nuclear-Cytoplasmic Transport of the Negative-Sense Viral RNAs." Virology **194**(2): 822-827.

Forrest, L. R., A. Kukol, et al. (2000). "Exploring Models of the Influenza A M2 Channel: MD Simulations in a Phospholipid Bilayer." Biophys. J. **78**(1): 55-69.

Gambaryan, A. S., J. S. Robertson, et al. (1999). "Effects of egg-adaptation on the receptor-binding properties of human influenza A and B viruses." Virology **258**(2): 232-9.
Gerdil, C. (2003). "The annual production cycle for influenza vaccine." Vaccine **21**(16): 1776-9.
Ghosh, S., P. Ghosh, et al. (2006). "iSimBioSys: a discrete event simulation platform for 'in silico' study of biological systems." in proc. Simulation Symposium, 2006. 39th Annual.
Hucka, M., A. Finney, et al. (2003). "The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models." Bioinformatics **19**(4): 524-531.
Jones, J. C., E. A. Turpin, et al. (2006). "Inhibition of influenza virus infection by a novel antiviral peptide that targets viral attachment to cells." J Virol **80**(24): 11960-7.
Joshi-Tope, G., M. Gillespie, et al. (2005). "Reactome: a knowledgebase of biological pathways." Nucl. Acids Res. **33**(suppl_1): D428-432.
Kanehisa, M. and S. Goto (2000). "KEGG: Kyoto Encyclopedia of Genes and Genomes." Nucl. Acids Res. **28**(1): 27-30.
Karp, P. D., C. A. Ouzounis, et al. (2005). "Expansion of the BioCyc collection of pathway/genome databases to 160 genomes." Nucl. Acids Res. **33**(19): 6083-6089.
Kelton, D. W., R. P. Sadowski, et al. (1997). Simulation with Arena.
Kemler, I., G. Whittaker, et al. (1994). "Nuclear Import of Microinjected Influenza Virus Ribonucleoproteins." Virology **202**(2): 1028-1033.
Lakadamyali, M., M. J. Rust, et al. (2003). "Visualizing infection of individual influenza viruses." PNAS **100**(16): 9280-9285.
Lakadamyali, M., M. J. Rust, et al. (2006). "Ligands for Clathrin-Mediated Endocytosis Are Differentially Sorted into Distinct Populations of Early Endosomes." Cell **124**(5): 997-1009.
Lars Mˆhler, D. F. H. S. U. R. (2005). "Mathematical model of influenza A virus production in large-scale microcarrier culture." Biotechnology and Bioengineering **90**(1): 46-58.
Leser, G. P. and R. A. Lamb (2005). "Influenza virus assembly and budding in raft-derived microdomains: a quantitative analysis of the surface distribution of HA, NA and M2 proteins." Virology **342**(2): 215-27.
Martin, K. and A. Helenius (1991). "Transport of incoming influenza virus nucleocapsids into the nucleus." J. Virol. **65**(1): 232-244.
Matlin, K. S., H. Reggio, et al. (1981). "Infectious entry pathway of influenza virus in a canine kidney cell line." J Cell Biol **91**(3 Pt 1): 601-13.
McNab, R. and F. W. Howell (1998). "simjava: a discrete event simulation package for Java with applications in computer systems modelling." in proc. First International Conference on Web-based Modelling and Simulation, San Diego CA, Society for Computer Simulation, Jan 1998.

Mochalova, L., A. Gambaryan, et al. (2003). "Receptor-binding properties of modern human influenza viruses primarily isolated in Vero and MDCK cells and chicken embryonated eggs." Virology **313**(2): 473-80.
Mullin, A. E., R. M. Dalton, et al. (2004). "Increased amounts of the influenza virus nucleoprotein do not promote higher levels of viral genome replication." J Gen Virol **85**(12): 3689-3698.
Ng, S. S., O. T. Li, et al. (2008). "Heterologous influenza vRNA segments with identical non-coding sequences stimulate viral RNA replication in trans." Virol J **5**: 2.
Noda, T., H. Sagara, et al. (2006). "Architecture of ribonucleoprotein complexes in influenza A virus particles." Nature **439**(7075): 490-492.
Portela, A. and P. Digard (2002). "The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication." J Gen Virol **83**(4): 723-734.
Ramsey, S., D. Orrell, et al. (2005). "Dizzy: stochastic simulation of large-scale genetic regulatory networks." J Bioinform Comput Biol **3**(2): 415-36.
Roberts, A., E. Kretzschmar, et al. (1998). "Vaccination with a Recombinant Vesicular Stomatitis Virus Expressing an Influenza Virus Hemagluttanin Provides Complete Protection from Influenza Virus Challenge." J. Virol. **72**(6): 4704-4711.
Shapiro, G. I., T. Gurney, Jr., et al. (1987). "Influenza virus gene expression: control mechanisms at early and late times of infection and nuclear-cytoplasmic transport of virus-specific RNAs." J. Virol. **61**(3): 764-773.
Shapiro, G. I. and R. M. Krug (1988). "Influenza virus RNA replication in vitro: synthesis of viral template RNAs and virion RNAs in the absence of an added primer." J. Virol. **62**: 2285.
Shih, S. R. and R. M. Krug (1996). "Surprising function of the three influenza viral polymerase proteins: selective protection of viral mRNAs against the cap-snatching reaction catalyzed by the same polymerase proteins." Virology **226**(2): 430-5.
Sieczkarski, S. B. and G. R. Whittaker (2003). "Differential Requirements of Rab5 and Rab7 for Endocytosis of Influenza and Other Enveloped Viruses." Traffic **4**(5): 333-343.
Subbarao, K., A. Klimov, et al. (1998). "Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness." Science **279**: 393.
Thompson, W. W., D. K. Shay, et al. (2003). "Mortality Associated With Influenza and Respiratory Syncytial Virus in the United States." JAMA **289**(2): 179-186.
Trampuz, A., R. M. Prabhu, et al. (2004). "Avian Influenza: A New Pandemic Threat?" Mayo Clin Proc **79**(4): 523-530.
Vreede, F. T. and G. G. Brownlee (2006). "Influenza virion-derived viral ribonucleoproteins synthesize both mRNA and cRNA in vitro." J Virol.

Vreede, F. T., T. E. Jung, et al. (2004). "Model Suggesting that Replication of Influenza Virus Is Regulated by Stabilization of Replicative Intermediates." J. Virol. **78**(17): 9568-9572.

Wang, X., M. Li, et al. (2000). "Influenza A Virus NS1 Protein Prevents Activation of NF-kappa B and Induction of Alpha/Beta Interferon." J. Virol. **74**(24): 11566-11573.

Whittaker, G. R., M. Kann, et al. (2000). "VIRAL ENTRY INTO THE NUCLEUS." Annual Review of Cell and Developmental Biology **16**(1): 627-651.

Yoshimura, A., K. Kuroda, et al. (1982). "Infectious Cell Entry Mechanism of Influenza Virus." J. Virol. **43**(1): 284-293.

Zdanov, V. M. and A. G. Bukrinskaja (1969). "Myxoviruses reproduction." Medicina, Moscow, Russia.

BIOGRAPHICAL INFORMATION

Richard Burke Squires, received his Bachelor of Science degree from the University of Dallas, Irving, Texas majoring in Biochemistry with a concentration in Computer Science beginning his parallel paths in the biological and computer sciences. Burke then entered the laboratory, learning molecular biology techniques that would benefit him as he continued to straddle the world of biology and computer science. His work then took him to the Center for Biomedical Inventions (CBI) at the University of Texas Southwestern Medical Center at Dallas (UT Southwestern). At the CBI Burke began his formal work in bioinformatics beginning in 1998. Jumping in the opportunity to continue his bioinformatics work Burke joined the biotechnology start-up Eliance, which would be bought out by MacroGenics a couple of eyars later. He later returned to UT Southwestern to his present position with the BioHealthBase Bioinformatics Resource Center (BRC). It was with the BioHealthBase BRC that Burke was introduced to influenza virus. He looks forward to continuing to develop the FluSim system while pursuing a Doctorial degree. Burke and his wife Christine live in Irving and are expecting their first child.