STOCHASTIC MODELING OF INTRACELLULAR INFLUENZA PRODUCTION

by

Dylan Barth

Submitted in partial fulfillment of the

requirements for Departmental Honors in

the Department of Physics and Astronomy

Texas Christian University

Fort Worth, Texas

May 7, 2018

STOCHASTIC MODELING OF INTRACELLULAR INFLUENZA PRODUCTION

Project Approved:

Supervising Professor: Hana Dobrovolny, Ph. D

Department of Physics and Astronomy

Brent Cooper, Ph. D

Department of Psychology

William Roche, Ph. D

Department of Philosophy

ABSTRACT

Influenza is a ubiquitous virus that has a high rate of mutation. Vaccinations and antiviral medications must change with a similarly rapid pace to be effective against this capricious infection. At the inception of the mutation, influenza may produce some combinations of wild type surface proteins, wild type RNA, and mutated versions of both. To better aid understanding of these mutations and pave a path for combating mutated virions, we modify a mathematical model of intracellular replication dynamics created by Heldt et al. to include genetic mutations at varying times. We then measure the effect of the mutation time on the combination of wild type and mutated surface proteins and RNA.

Stochastic Modeling of Intracellular Influenza Production

Dylan Barth

May 6, 2018

Abstract

Influenza is a ubiquitous virus that has a high rate of mutation. Vaccinations and antiviral medications must change with a similarly rapid pace to be effective against this capricious infection. At the inception of the mutation, influenza may produce some combinations of wild type surface proteins, wild type RNA, and mutated versions of both. To better aid understanding of these mutations and pave a path for combating mutated virions, we modify a mathematical model of intracellular replication dynamics created by Heldt et al. to include genetic mutations at varying times. We then measure the effect of the mutation time on the combination of wild type and mutated surface proteins and RNA.

1 Introduction

Influenza is a deadly and common virus. So much so that our attempts to prevent or treat the infection it causes are commonly prescribed to individuals every year during flu season. However, influenza has a particularly high rate of mutation [10, 11], making the drugs and vaccines that we make to fight and prevent it ineffective over short periods of time [3, 5, 9]. These mutations do not happen immediately after the virus has entered the cell, more likely during the eclipse phase of the virus life cycle. This leads to a mixture of both mutated and wild type surface proteins and RNA in progeny virions [3]. These chimerical virions can have mutated surface proteins and wild type RNA or vice versa.

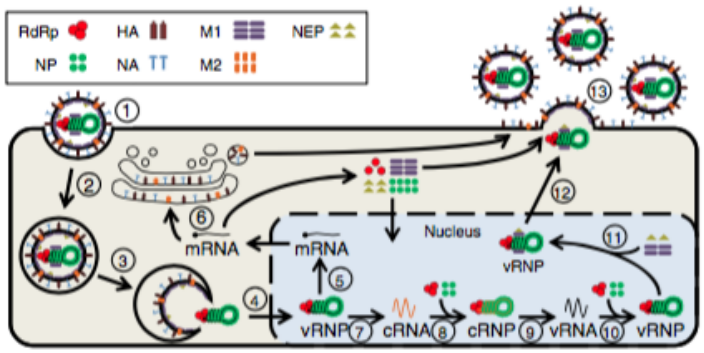

Figure 1: During its life cycle, a virion enters the cell through the cell membrane and deposits its genetic code into the nucleus. This genetic code is used to produce surface proteins and copies of the genetic code itself with materials inside the cell. These pieces are then packaged into progeny virions which exit the cell through the membrane.

Modeling studies have shown that chimerical virions behave differently to different medications [2, 3]. For instance, a wild type surface protein paired with a mutated RNA will still be affected by drugs whose mechanism of action targets the wild type surface protein. But a mutated surface protein paired with wild type RNA will make the wild type virus "immune" to these approaches. I will use an already established intracellular model of the virus life cycle [6] to determine the how much the amount of each kind of virus is affected by the time that the virus mutates.

2 Methods

2.1 Intracellular model

To better understand the nature of mutation in the influenza virus, we modified a model proposed by Heldt et al. [6] that describes the mechanisms involved in intracellular viral replication. This model consists of 25 ordinary differential equations that correspond to various chemical processes that take place during viral replication. A list of the equations from Heldt et al. can be found in appendix A. Pictured in Fig. 1 are the stages in a virus life cycle.

2.2 Adding a mutation

A mutation event was added such that any wild type RNA that had not yet been split into its cRNA and mRNA progeny would then have a mutated section of RNA affecting primarily the neuraminidase surface protein. This mutation affects two sections of progeny virus, the progeny neuraminidase surface protein and the viral RNA that would code future mutated surface proteins. These two mutations do not necessarily coincide, and in fact mismatches in mutated surface proteins and RNA sequences are a well established phenomenon. The mutated proteins and RNA along with their wild type counterparts were tracked as the mutation time was varied.

To incorporate the muatation, we modified the system by creating mutated counterparts for each equation that was implicated by the segment of RNA that codes the neuraminidase protein in the progeny virus or the production of the protein encapsulating the progeny virus. This mutation occurred at a predetermined time, after which all incoming vRNA would contain the mutation. The proportions of the various combinations of progeny wild type and mutated virus parts were measured as the mutation time was varied. The following are the equations added to simulate the mutated virus,

$$\frac{dVp_{cyt_{mut}}}{dt} = 8k^{fus}V^{en} - k^{imp}Vp_{cyt_{mut}} \tag{1}$$

$$\frac{dVp_{nuc_{mut}}}{dt} = k^{imp}Vp_{cyt_{mut}} + k^{bind}_{NP}P_{NP}R^{V}_{RdRp} - (k^{bind}_{M1}P_{M1} + k^{deg}_{RNP})Vp_{nuc_{mut}} \tag{2}$$

$$\frac{dR^{C}}{dt} = k^{syn}_{C}Vp_{nuc} + k^{syn}_{C}Vp_{nuc_{mut}} - k^{bind}_{RdRp}P_{RdRp}R^{C} - k^{deg}_{R}R^{C} \tag{3}$$

$$\frac{dVp_{nucM1_{mut}}}{dt} = k^{bind}_{M1}P_{M1}Vp_{nuc_{mut}} - (k^{exp}P_{NEP} + k^{deg}_{RNP})Vp_{nucM1_{mut}} \tag{4}$$

$$\frac{dVp_{cytM1_{mut}}}{dt} = k^{exp}P_{NEP}Vp_{nucM1_{mut}} - 8r^{Rel} - k^{deg}_{Rnp}Vp_{cytM1_{mut}} \tag{5}$$

$$\frac{RM6_{mut}}{dt} = \frac{k^{syn}_{M}}{L_6}\frac{Vp_{nuc_{mut}}}{8} - k^{deg}_{M}R^{M}_{6} \tag{6}$$

$$\frac{dP_{NA_{mut}}}{dt} = \frac{k^{syn}_{p}}{D_{rib}}R^{M}_{6_{mut}} - N_{P_{NA}}r^{Rel}, \tag{7}$$

$$\tag{8}$$

where Vp_{cyt} is the number of virions (unmutated unless denoted in the subscript) present in the cytoplasm of the cell. V^{en} is the number of virions fused to the cell membrane during endocytosis. V_{nuc} is similarly the number of virions within the nucleus of the cell. P_{NP}, P_{RdRp}, P_{M1}, P_{NEP} are the number of NP, RdRp, M1, and NEP proteins respectively. R^V_{RdRp} is the number of vRNP viral polymerase complexes. R^C is the number of cRNA molecules. Vp_{nucM1} and Vp_{cytM1} are the numbers of cytoplasmic and nuclear NEP-M1-vRNP complexes respectively. R^M_6 is the number 6 strand of mRNA. Further explanation of equations, rates, and variable names can be found in Heldt et al. [6].

2.3 Computational methods

We solved this system of equations using a version of the Gillespie algorithm [1]. This algorithm is often used in biochemical simulations for its low computational load compared to its accuracy in output [8]. The system of differential equations was split into 37 events corresponding to each term within an equation, the rates of which were given by the value of the terms. A dynamic uniform probability function was defined by the rates of possible events at a given time, each possible event was given a probability proportional to its time dependent rate so that more common events were more likely. Once an event is chosen from this distribution, a timestep over which to execute the event was chosen from an exponential distribution with a time constant equal to the sum of all rates. Because these rates are time dependent, the probability distribution for events and timesteps was also dynamic. This algorithm was implemented in python and run on a Linux-based system several times to smooth any noise introduced by the stochastic procedure.

3 Results

3.1 Model validation

Principally we want to show that the model worked as intended, several key elements of the viral life cycle were measured and graphed where both wild type and mutated parts appear (Fig. 2). As expected, the mutated versions do not appear until later in the infection

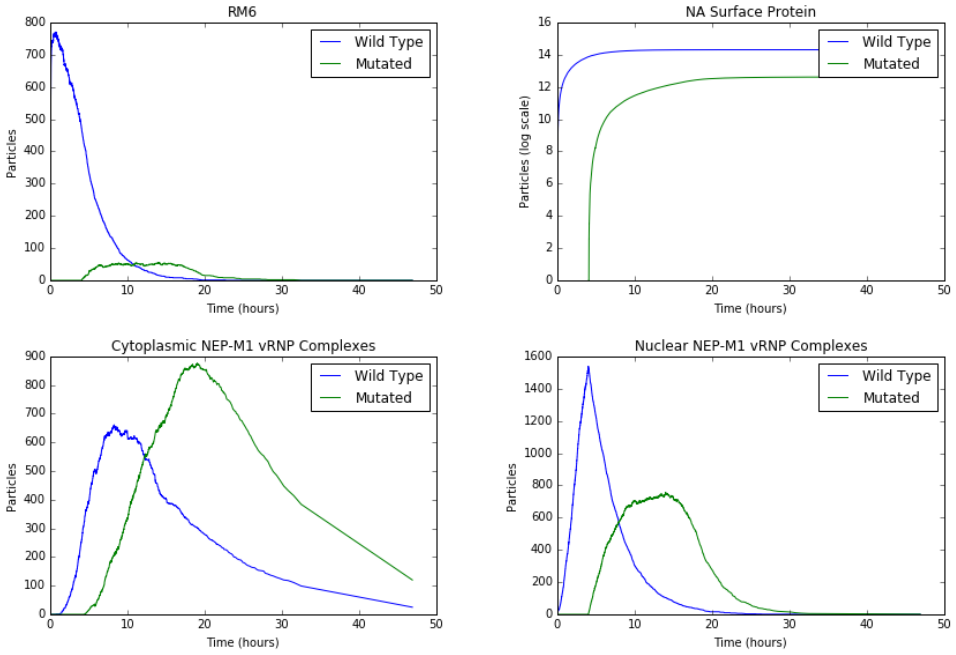

Figure 2: Graphs showing the change in number of particles both mutated and wild type over time during a simulation of mutation time at hour 4

and grow at a somewhat slower rate due to resources within the cell being used by wild type counterparts. This fact changes, as expected, as the time of mutation gets closer to the inception of the virus into the cell. Therefore, not only have we shown that the mutation event works as expected, but also that the implementation of the model was carried out correctly.

Further evidence of the veracity of our model can be gained by analyzing the changes in timestep over time (Fig. 3). Since the distribution for timestep is dynamic and dependent on the size of the sum of the rates, it is expected that when the most events are being executed that the timestep dip to its global minimum. The figure below shows the timestep as the simulation runs, and we observe the expected outcome. Between the hours of T+4:00 and T+10:00 the timestep shrinks significantly, which is also the time at which the most events occur. This corresponds biologically to the eclipse phase of the virus life cycle, when the virus orchestrates the mechanisms within the cell

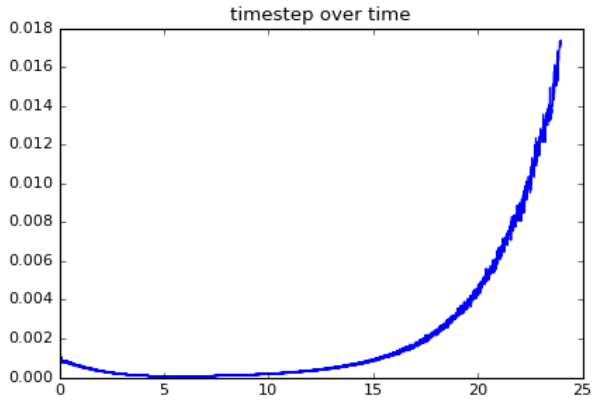

Figure 3: Timestep in relation to simulated time.

to produce parts that will be packaged as a progeny virion.

3.2 Time of mutation

The proportion of different combinations of wild type and mutated surface proteins and RNA are shown in Fig. 4. As one might hypothesize, the much later mutation times produced very little mutated progeny virus. This is to be expected as in the later stages of the viral life cycle, various necessary mechanisms provided by the host cell have been run dry. In addition, the majority of virus that has entered the cell has already passed the stage at which it could possibly mutate. Rather unexpectedly, the two chimerical virions were not equivalent in their proportions. Fully mutated virions only dominated in the runs with early mutation times, which were tracked closely by the virions with wild type surface proteins and mutated RNA. In runs where the mutation time was later, the cell still produced virions with mutated surface proteins, despite hardly any mutated RNA being produced. This leads us to believe that the limiting factor for a fully mutated progeny virus is the production of mutated progeny vRNP.

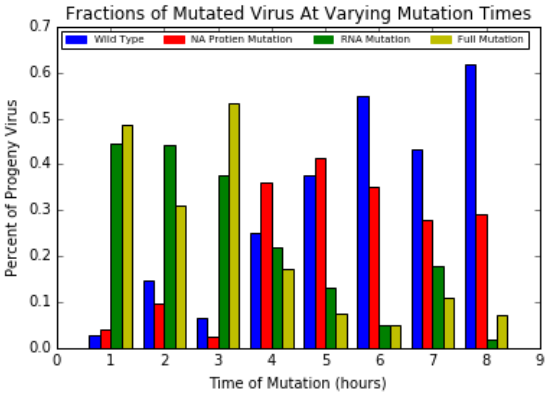

Figure 4: Relative proportions of the 4 possible combinations of mutated and wild type variants with respect to a change in mutation time.

4 Discussion

This research is part of a growing literature concerning chimerical virions. The computational method of measuring the proportions of these virions will be informative until empirical evidence can be gathered about their relative proportions. This may be difficult, as there exist biological limitations to measuring virion output in experimental settings, however there are some promising attempts to solve this problem using fluorescent labeling [7].

There are limitations to the model. This model does not take into account immune system response [4], which may have implications on the proportion of virus types if the immune system is more likely to target some types of virions or has a time dependent response. Another limitation is that the simulation is only of a single cell, the relationship between cells is likely important and will change the proportions of the chimerical virions. Further, the model depends on the accuracy of the constants for each of the equations, some of which have not been measured directly, but rather inferred from other measurements or found by fitting the model to relevant data [6].

In conclusion, further research may be guided by the findings in this research. An interesting relationship was discovered between the chimerical virions when varying time of mutation, and an unexpected asymmetry between them was discovered. The model may be im-

proved by implementing an immune system response and simulating multiple cells, despite this addition requiring a significant increase in computational power.

References

[1] A. A. CV Rao. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *J. Chem. Phys.*, 118(11):4999–5010, March 15 2013. doi: 10.1063/1.1545446.

[2] L. A. Deecke and H. M. Dobrovolny. Intermittent treatment of severe influenza. *J. Theor. Biol.*, 442:129–138, April 7 2018. doi: 10.1016/j.jtbi.2018.01.012.

[3] H. M. Dobrovolny and C. A. Beauchemin. Modelling the emergence of influenza drug resistance: The roles of surface proteins, the immune response and antiviral mechanisms. *PLoS One*, 12 (7):e0180582, 10 July 2017. doi: 10.1371/journal.pone.0180582.

[4] H. M. Dobrovolny, M. B. Reddy, M. A. Kamal, C. R. Rayner, and C. A. Beauchemin. Assessing mathematical models of influenza infections using features of the immune response. *PLoS One*, 8(2): e57088, 28 February 2013. doi: 10.1371/journal.pone.0057088.

[5] G. Dos Santos, E. Neumeier, and R. Bekkat-Berkani. Influenza: Can we cope better with the unpredictable? *Hum. Vaccin. Immunother.*, 12(3):699–708, 2016. doi: 10.1080/21645515.2015.1086047.

[6] F. S. Heldt, T. Frensing, and U. Reichl. Modelling the intracellular dynamics of influenza virus replication to understand the control of viral rna synthesis. *Journal of Virology*, 86(15):7806–7817, July 2012. doi: doi:10.1128/JVI.00080-12.

[7] J. Lee, S. R. Ahmed, S. Oh, J. Kim, T. Suzuki, K. Parmar, S. S. Park, J. Lee, and E. Y. Park. A plasmon-assisted fluoro-immunoassay using gold nanoparticle-decorated carbon nanotubes for monitoring the influenza virus. *Biosens. Bioelec.*, 64:311–317, February 15 2015. doi: 10.1016/j.bios.2014.09.021.

[8] H. Li and L. Petzold. Efficient parallelization of the stochastic simulation algorithm for chemicaly reacting systems on the graphics processing unit. *Intl. J. High Perform. Comp. Appl.*, 24 (2):107–116, 2010. doi: 10.1177/1094342009106066.

[9] A. S. Perelson, L. Rong, and F. G. Hayden. Combination antiviral therapy for influenza: Predictions from modeling of human infections. *J. Infect. Dis.*, 205:1642–1645, 1 June 2012. doi: 10.1093/infdis/jis265.

[10] R. Sanjuan and P. Domingo-Calap. Mechanisms of viral mutation. *Cell. Mol. Life Sci.*, 73(23):4433–4448, December 2016. doi: 10.1007/s00018-016-2299-6.

[11] M. Villa and M. Lassig. Fitness cost of reassortment in human influenza. *Plos Pathogens*, 13(11):e1006685, November 2017. doi: 10.1371/journal.ppat.1006685.

A Equations

$$\frac{dV^{\mathrm{Ex}}}{dt} = k_{\mathrm{hi}}^{\mathrm{Dis}} V_{\mathrm{hi}}^{\mathrm{Att}} + k_{\mathrm{lo}}^{\mathrm{Dis}} V_{\mathrm{lo}}^{\mathrm{Att}} - \left(k_{\mathrm{hi}}^{\mathrm{Att}} B_{\mathrm{hi}} + k_{\mathrm{lo}}^{\mathrm{Att}} B_{\mathrm{lo}}\right) V^{\mathrm{Ex}}, \tag{1}$$

$$\text{with } B_n = B_n^{\mathrm{tot}} - V_n^{\mathrm{Att}},\ n \in \{\mathrm{hi},\ \mathrm{lo}\}, \tag{2}$$

$$\frac{dV_{\mathrm{n}}^{\mathrm{Att}}}{dt} = k_{\mathrm{n}}^{\mathrm{Att}} B_n V^{\mathrm{Ex}} - \left(k_n^{\mathrm{Dis}} + k^{\mathrm{En}}\right) V_n^{\mathrm{Att}}, \tag{3}$$

$$\frac{dV^{\mathrm{En}}}{dt} = k^{\mathrm{En}}\left(V_{\mathrm{hi}}^{\mathrm{Att}} + V_{\mathrm{lo}}^{\mathrm{Att}}\right) - \left(k^{\mathrm{Fus}} + k_{\mathrm{Ven}}^{\mathrm{Deg}}\right) V^{\mathrm{En}},$$

$$\text{with } k_n^{\mathrm{Dis}} = \frac{k_n^{\mathrm{Att}}}{k^{\mathrm{Eq}}} \text{ and } k_{\mathrm{Ven}}^{\mathrm{Deg}} = \frac{1 - \mathrm{F}_{\mathrm{Fus}}}{\mathrm{F}_{\mathrm{Fus}}} k^{\mathrm{Fus}}, 0 < F_{\mathrm{Fus}} \leq 1, \tag{4}$$

$$\frac{dVp^{\mathrm{cyt}}}{dt} = 8k^{\mathrm{Fus}} V^{\mathrm{En}} - k^{\mathrm{Imp}} Vp^{\mathrm{cyt}}, \tag{5}$$

$$\frac{dVp^{\mathrm{nuc}}}{dt} = k^{\mathrm{Imp}} \mathrm{Vp}^{\mathrm{cyt}} + k_{\mathrm{NP}}^{\mathrm{Bind}} P_{\mathrm{NP}} R_{\mathrm{RdRp}}^{\mathrm{V}} - \left(k_{\mathrm{M1}}^{\mathrm{Bind}} P_{\mathrm{M1}} + k_{\mathrm{Rnp}}^{\mathrm{Deg}}\right) Vp^{\mathrm{nuc}}, \tag{6}$$

$$\frac{dR^{\mathrm{C}}}{dt} = k_{\mathrm{C}}^{\mathrm{Syn}} Vp^{\mathrm{nuc}} - k_{\mathrm{RdRp}}^{\mathrm{Bind}} P_{\mathrm{RdRp}} R^{\mathrm{C}} - k_{\mathrm{R}}^{\mathrm{Deg}} R^{\mathrm{C}}, \tag{7}$$

$$\frac{dR^{\mathrm{V}}}{dt} = k_{\mathrm{V}}^{\mathrm{Syn}} Cp - k_{\mathrm{RdRp}}^{\mathrm{Bind}} P_{\mathrm{RdRp}} R^{\mathrm{V}} - k_{\mathrm{R}}^{\mathrm{Deg}} R^{\mathrm{V}}, \tag{8}$$

$$\frac{dR_{\mathrm{RdRp}}^{\mathrm{V}}}{dt} = k_{\mathrm{RdRp}}^{\mathrm{Bind}} P_{\mathrm{RdRp}} R^{\mathrm{V}} - k_{\mathrm{NP}}^{\mathrm{Bind}} P_{\mathrm{NP}} R_{\mathrm{RdRp}}^{\mathrm{V}} - k_{\mathrm{RRdRp}}^{\mathrm{Deg}} R_{\mathrm{RdRp}}^{\mathrm{V}}, \tag{10}$$

$$\frac{dCp}{dt} = k_{\mathrm{NP}}^{\mathrm{Bind}} P_{\mathrm{NP}} R_{\mathrm{RdRp}}^{\mathrm{C}} - k_{\mathrm{Rnp}}^{\mathrm{Deg}} Cp, \tag{11}$$

$$\frac{dVp_{\mathrm{M1}}^{\mathrm{nuc}}}{dt} = k_{\mathrm{M1}}^{\mathrm{Bind}} P_{\mathrm{M1}} Vp^{\mathrm{nuc}} - \left(k^{\mathrm{Exp}} P_{\mathrm{NEP}} + k_{\mathrm{Rnp}}^{\mathrm{Deg}}\right) Vp_{\mathrm{M1}}^{\mathrm{nuc}}, \tag{12}$$

$$\frac{dVp_{\mathrm{M1}}^{\mathrm{cyt}}}{dt} = k^{\mathrm{Exp}} P_{\mathrm{NEP}} Vp_{\mathrm{M1}}^{\mathrm{nuc}} - 8\mathrm{r}^{\mathrm{Rel}} - k_{\mathrm{Rnp}}^{\mathrm{Deg}} Vp_{\mathrm{M1}}^{\mathrm{cyt}}. \tag{13}$$

$$\frac{dR_i^{\mathrm{M}}}{dt} = \frac{k_{\mathrm{M}}^{\mathrm{Syn}}}{L_i}\frac{V\mathrm{p}^{\mathrm{nuc}}}{8} - k_{\mathrm{M}}^{\mathrm{Deg}}R_i^{\mathrm{M}}, \quad i = 1, \ldots, 8, \tag{14}$$

$$\frac{dP_{\mathrm{PB1}}}{dt} = \frac{k_P^{\mathrm{Syn}}}{D_{\mathrm{Rib}}}R_2^{\mathrm{M}} - k^{\mathrm{RdRp}}\mathrm{P}_{\mathrm{pB1}}\mathrm{P}_{\mathrm{pB2}}\mathrm{P}_{\mathrm{pA}}, \tag{15}$$

$$\frac{dP_{\mathrm{PB2}}}{dt} = \frac{k_P^{\mathrm{Syn}}}{D_{\mathrm{Rib}}}R_1^{\mathrm{M}} - k^{\mathrm{RdRp}}\mathrm{P}_{\mathrm{pB1}}\mathrm{P}_{\mathrm{pB2}}\mathrm{P}_{\mathrm{pA}}, \tag{16}$$

$$\frac{dP_{\mathrm{PA}}}{dt} = \frac{k_P^{\mathrm{Syn}}}{D_{\mathrm{Rib}}}R_3^{\mathrm{M}} - k^{\mathrm{RdRp}}\mathrm{P}_{\mathrm{pB1}}\mathrm{P}_{\mathrm{pB2}}\mathrm{P}_{\mathrm{pA}}, \tag{17}$$

$$\frac{dP_{\mathrm{RdRp}}}{dt} = k^{\mathrm{RdRp}}P_{\mathrm{PB1}}P_{\mathrm{PB2}}P_{\mathrm{PA}} - k_{\mathrm{RdRp}}^{\mathrm{Bind}}P_{\mathrm{RdRp}}\left(R^{\mathrm{V}} + R^{\mathrm{C}}\right) - \left(N_{\mathrm{P}_{\mathrm{RdRp}}} - 8\right)r^{\mathrm{Rel}}, \tag{18}$$

$$\begin{aligned}\frac{dP_{\mathrm{NP}}}{dt} = {} & \frac{k_{\mathrm{P}}^{\mathrm{Syn}}}{D_{\mathrm{Rib}}}R_5^{\mathrm{M}} - \frac{L_V}{N_{\mathrm{NP}}^{\mathrm{Nuc}}}k_{\mathrm{NP}}^{\mathrm{Bind}}P_{\mathrm{NP}}\left(R_{\mathrm{RdRp}}^{\mathrm{V}} + R_{\mathrm{RdRp}}^{\mathrm{C}}\right) \\ & - \left(N_{P_{\mathrm{NP}}} - 8\frac{L_V}{N_{\mathrm{NP}}^{\mathrm{Nuc}}}\right)\mathrm{r}^{\mathrm{Rel}},\end{aligned} \tag{19}$$

$$\begin{aligned}\frac{dP_{\mathrm{M1}}}{dt} = {} & \frac{k_{\mathrm{P}}^{\mathrm{Syn}}}{D_{\mathrm{Rib}}}\left(1 - F_{\mathrm{Spl7}}\right)R_7^{\mathrm{M}} - \frac{L_{\mathrm{V}}}{N_{\mathrm{M1}}^{\mathrm{Nuc}}}k_{\mathrm{M1}}^{\mathrm{Bind}}P_{\mathrm{M1}}V\mathrm{p}^{\mathrm{nuc}} \\ & - \left(\mathrm{N}_{P_{\mathrm{M1}}} - 8\frac{L_{\mathrm{V}}}{N_{\mathrm{M1}}^{\mathrm{Nuc}}}\right)r^{\mathrm{Rel}},\end{aligned} \tag{20}$$

$$\frac{dP_{\mathrm{HA}}}{dt} = \frac{k_p^{\mathrm{Syn}}}{D_{\mathrm{Rib}}}R_4^{\mathrm{M}} - N_{P_{\mathrm{HA}}}r^{\mathrm{Rel}}, \tag{22}$$

$$\frac{dP_{\mathrm{NA}}}{dt} = \frac{k_p^{\mathrm{Syn}}}{D_{\mathrm{Rib}}}R_6^{\mathrm{M}} - N_{P_{\mathrm{NA}}}r^{\mathrm{Rel}}, \tag{23}$$

$$\frac{dP_{\mathrm{M2}}}{dt} = \frac{k_p^{\mathrm{Syn}}}{D_{\mathrm{Rib}}}F_{\mathrm{Spl7}}R_7^{\mathrm{M}} - N_{P_{\mathrm{M2}}}r^{\mathrm{Rel}}, \tag{24}$$

$$\frac{dV^{\mathrm{Rel}}}{dt} = r^{\mathrm{Rel}} = k^{\mathrm{Rel}}Vp_{\mathrm{M1}}^{\mathrm{cyt}}\prod_j \frac{P_j}{P_{\mathrm{j}} + K_{V^{\mathrm{Rel}}}N_{\mathrm{P_j}}} \quad \text{with } j \in \{\mathrm{RdRp, HA, NP, NA, M1, M2, NEP}\} \tag{25}$$

A list of variable names and their meanings along with tables including numerical values for rates can be found in Heldt et al. [6]