Karina Alexandra Pál

Bachelor in Chemical and Biochemical Engineering

Intracellular Stochastic Modelling of Influenza A Virus and DIP Replication

Dissertation submitted to obtain the degree of Master of Science in Chemical and Biochemical Engineering

Adviser:	Daniel Rüdiger, Master of Science, Max Planck Institute for Dynamics of Complex Technical Systems
Co-adviser:	Rui Manuel Freitas Oliveira, Associated Professor with Aggregation, Faculty of Science and Technology – NOVA University of Lisbon

Examination Committee

Chairperson:	Prof. Dr. Mário Fernando José Eusébio, Assistant Professor, NOVA University of Lisbon
Raporteur:	Dr. Rafael Sousa Costa, Postdoctoral Researcher, Instituto Superior Técnico
Member:	Prof. Dr. Rui Manuel Freitas Oliveira, Associated Professor with Aggregation, NOVA University of Lisbon

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

September, 2018

Karina Alexandra Pál

Bachelor in Chemical and Biochemical Engineering

Intracellular Stochastic Modelling of Influenza A Virus and DIP Replication

Dissertation submitted to obtain the degree of Master of Science in
Chemical and Biochemical Engineering

Adviser: Daniel Rüdiger, Master of Science,
Max Planck Institute for Dynamics of Complex Technical Systems

Co-advisers: Rui Manuel Freitas Oliveira, Associated Professor with Aggregation, Faculty of Science and Technology – NOVA University of Lisbon

Examination Committee

Chairperson: Prof. Dr. Mário Fernando José Eusébio, Assistant Professor, NOVA University of Lisbon
Raporteur: Dr. Rafael Sousa Costa, Postdoctoral Researcher, Instituto Superior Técnico
Member: Prof. Dr. Rui Manuel Freitas Oliveira, Associated Professor with Aggregation, NOVA University of Lisbon

September, 2018

Intracellular Stochastic Modelling of Influenza A Virus and DIP Replication

Copyright © Karina Alexandra Pál e Faculdade de Ciências e Tecnologia – Universidade NOVA de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade NOVA de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

The past, like the future, is indefinite
and exists only as a spectrum of possibilities.

Stephen Hawking

To lead of my acknowledgments, I would like to thank firstly to my adviser Daniel Rüdiger for giving me the opportunity to work in a field that I was very excited to learn more about and guiding me in the right direction through my research. He showed an amazing availability whenever I had any question or run into dead ends. I am gratefully for his valuable comments and corrections on this thesis, I learned so much with him and he did everything in his power to help me achieve my goal. In second, I am sure that professor Rui Oliveira contributions were also very important in providing me all the support before and during my Erasmus mobility, always showing a great concern about my progress.

I believe that Tanja, Sacha and my office partner Lukas, gave me an important feedback and valuable advises which helped me improving my thesis, also professor Udo Reichl and the Bioprocess Engineering Group were very helpful in giving me the opportunity to work at the MPI for Dynamics of Complex Technical Systems and providing me all the resources needed to pursue my research.

I also feel immensely grateful towards my parents, Aurelia and Ferenc, for supporting my decisions in life and providing me all the means to pursue my goals. A big share of gratitude to my grandmother, Măriuța, for all her words and advises that helped me not to give up and made me go chase my dreams. I would also like to thank my boyfriend, Gustavo, for this amazing 4 years and for all the patience during the more stressful times. We learned a lot together and I am very proud for what we achieved together, so far.

I would also like to praise my partners in crime, Cláudia and Suzana, for supporting me through this amazing journey in Magdeburg. Even during hard times, we were always there to ensure that each of us were laughing and that we could overcome the obstacles together. I am sure we will keep in touch and our friendship bound is for life. I would also want to thank to José, Nino, Lars, Clair and Weiwei for all the evenings spent drinking good German beer which helped me to take my mind off work and relax. I feel also appreciate for meeting Juliana and João, for having the possibility to talk in Portuguese and making me feel at home. I am so grateful for this experience, for allowing me to meet such amazing people and all the moments that we spent together that surly, I will keep them forever in my heart.

I also believe that my friends Filipa, David, Tai, Daniela, Daniel, Nídia, Sara and Inês deserve recognition for all the nights spent at the university while working on Project I and II. Their companionship helped me so much during the hardest times of the course and I do not know if it would be possible to accomplish everything without them. A big thank you to professor Mário Eusébio for the advises and support during this period.

Finally, I want to thank my life long friends Samantha, Gabriela, Beatriz and Diogo for always being there for me and accompanying me closely for such a long time.

Abstract

Defective interfering particles (DIPs) are mutated versions of viruses that are characterized by carrying internal deletions in their genome. These deletions are introduced randomly during virus replication and the truncated genomes interfere with the propagation of their standard virus (STV) leading to reduced infectious virus titers. Therefore, DIPs were recently proposed to be used for antiviral therapy which increased the demand for a reliable production process and a better understanding of the interference mechanism. The infection dynamics were analysed by a deterministic modelling approach, however, the impact of stochastic effects introduced by cell-to-cell variability, different coinfection scenarios and an independent genome segment replication remain largely elusive. Hence, we developed a stochastic model of influenza A virus and DIP replication which considers the influence of these random effects on STV and DIP release. We found that the viral nucleoprotein (NP), which is essential for encapsidation of the naked viral RNA (vRNA), is strongly affected by fluctuations and three distinct sub-populations emerged in our model. Furthermore, simulations performed with one DIP and one STV infecting the cell resulted in mostly non-productive simulations, mainly caused by failures during the endocytosis of particles and by the random degradation of vRNAs. Moreover, the optimal DIP production was achieved when STV enters nucleus first and the DIP entry is delayed between 1.5 and 3 hours. Lastly, we demonstrate that the implementation of a two-step packaging process, which separates the formation of genome complexes and the assembly of all required proteins for release, is crucial to achieve a substantial DIP advantage over STV production. Overall, our simulations suggest that a combination of various random effects influences the replication of STVs and DIPs inducing a broad distribution of progeny particle release. The stochastic model developed in this thesis provides an ideal basis for the analysis of these effects and their impact on DIP interference and production.

Keywords: Defective interfering particle, influenza A virus, interference mechanism, stochastic mathematical model, stochastic effects

Resumo

Partículas defeituosas interferentes (DIPs) são versões mutagénicas do vírus, sendo caracterizadas por apresentarem secções internas do seu genoma excluídas. Estas secções são geradas aleatoriamente durante a replicação do vírus e o genoma incompleto interfere com a propagação do seu vírus standard (STV), advindo numa redução da concentração de vírus infeciosos. Consequentemente, o uso de DIPs foi recentemente proposto para terapia antiviral, o que resultou num aumento da procura de um processo de produção viável e de um conhecimento mais aprofundado do mecanismo de interferência. A dinâmica de infeção foi analisada através de uma abordagem de modelação determinística, no entanto, o impacto de efeitos estocásticos causados pela variabilidade celular, diferentes cenários de coinfecção e replicação independente dos diferentes segmentos do genoma permanecem assim largamente elusivos. Portanto, desenvolveu-se um modelo estocástico da replicação do vírus influenza A e DIPs que considera o impacto de efeitos aleatórios na produção de STVs e DIPs. Foi descoberto que a nucleoproteina viral (NP), que é essencial para a encapsidação do RNA viral (vRNA), é fortemente afetado por flutuações e três subpopulações distintas advêm do nosso modelo. Além disso, simulações efetuadas em que um DIP e um STV infetam a célula, resultam maioritariamente em simulações não produtivas causadas geralmente por uma falha durante a endocitose de partículas e pela degradação aleatória de vRNAs. Mais adiante, a produção de DIP é otimizada quando o STV entra no núcleo primeiro e a entrada da DIP é atrasada entre 1.5 e 3 horas. Finalmente, demonstrou-se que a implementação de um processo de empacotamento com duas etapas, separando a formação de complexos genómicos e a montagem de todas as proteínas necessárias para a libertação de partículas, é crucial para conferir uma vantagem substancial à DIP em relação à produção de STV. Em geral, as simulações sugerem que as combinações de vários efeitos aleatórios influenciam a replicação de STVs e DIPs, resultando numa distribuição de partículas produzidas muito alargada. O modelo estocástico desenvolvido nesta tese consiste uma base ideal para a análise destes efeitos e do seu impacto na produção e interferência das DIPs.

Termos-chave: Partículas defeituosas interferentes, vírus influenza A, mecanismo de interferência, modelo matemático estocástico, efeitos estocásticos

Contents

Abstract vii

Resumo ix

List of Figures xiii

List of Tables xv

List of Equations xvii

List of Abbreviations xix

List of Symbols xxi

1 Introduction 1

2 Theoretical Background 3

2.1 Discovery of defective interfering particles 3

2.2 Influenza A virus structure and DIP *de novo* generation 3

2.3 Influenza A virus intracellular life cycle 5

2.4 DIP interference 9

2.5 Therapeutic potential of DIPs 10

2.6 Intracellular deterministic model of influenza A virus and DIP replication 10

2.7 Intracellular stochastic model of influenza A virus replication 11

2.8 Stochastic simulation algorithms 12

3 Models and Methods 15

3.1 Implementation of biochemical reactions 15

3.2 Model variations 24

3.3 Computation and simulation assumptions 27

3.4 Data analysis tools 29

4 Results 31
  4.1 DIP interference and comparison with the deterministic model 31
    4.1.1 High MOI and MODIP 31
    4.1.2 Low MOI and MODIP 37
    4.1.3 Impact of MOI and MODIP 42
  4.2 Non-productive simulations 45
    4.2.1 Effects of MOI and MODIP 45
    4.2.2 Filtering non-productive simulations 50
  4.3 Timing investigation 51
    4.3.1 Randomly generated delay 51
    4.3.2 Induced delay 54
  4.4 Model variations 59
    4.4.1 Minimum packaging model 59
    4.4.2 Replication model 63
5 Discussion 67
6 Conclusion 73
Bibliography 75
Appendix 83

List of Figures

Figure 2.1 | Influenza A virus particle, genome structure and encoding proteins. 4
Figure 2.2 | Structure of a defective interfering particle (DIP) and de novo generation mechanism of defective RNA. 5
Figure 2.3 | Diagram of the influenza A virus life cycle. 6
Figure 2.4 | Different hypothesis for the transition mechanism from transcription to replication. 7

Figure 3.1| DIP advantage over STV replication. 25
Figure 3.2 | Average converges with increased number of simulations. 28
Figure 3.3 | Averaged stochastic dynamics and distribution. 29

Figure 4.1 | Distribution of standard virus (STV) and DIP production at MOI/MODIP 10/10. 31
Figure 4.2 | Comparison of the stochastic and deterministic simulation. 32
Figure 4.3 | Correlation of the defective interfering vRNA and various vRNA segments and DIP production. 33
Figure 4.4 | Correlation of vRNA segment 5 with defective and functional segment 3 and its influence in STV and DIP production. 34
Figure 4.5 | Comparison of mRNA and protein dynamics for stochastic and deterministic model simulations at MOI/MODIP 10/10. 35
Figure 4.6 | Levels and distributions of HA and NP related to STV and DIP yield. 36
Figure 4.7 | Distribution of STV and DIP production at MOI/MODIP 1/1. 37
Figure 4.8 | Higher differences between the stochastic and deterministic simulation at MOI/MODIP 1/1. 38
Figure 4.9 | Correlation of vRNA segment 5 with defective and functional segment 3 and its influence in STV and DIP production. 39
Figure 4.10 | Comparison of mRNA and protein dynamics for stochastic and deterministic model simulations at MOI/MODIP 1/1. 40
Figure 4.11 | Distributions of RdRp and NP in MOI/MODIP 1/1 scenario. 41
Figure 4.12 | Levels of RdRp and NP related to STV and DIP yield. 42
Figure 4.13 | Average cumulative DIPs released with increasing MOI. 43

Figure 4.14 | Average cumulative DIP released and increased MODIP. 44
Figure 4.15 | vRNP ratios and impact on STV and DIP release. 45
Figure 4.16 | Factors which develop STV non-productive cells. 46
Figure 4.17 | Factors preventing DIP production. 47
Figure 4.18 | Detailed factors which prevent DIP production at different MOIs and a fixed MODIP of 1. 48
Figure 4.19 | Percentage of segment loss for each vRNA segment at MODIP 1 and different MOI. ... 48
Figure 4.20 | Detailed factors which prevent DIP production at different MODIPs and a fixed MOI of 1. 49
Figure 4.21 | Percentage of segment loss for each vRNA segment MOI 1 and different MODIP. 49
Figure 4.22 | Filtered results and comparison with the deterministic model at MOI/MODIP 1/1. 50
Figure 4.23 | Standard deviation and filtered results during infection at MOI/MODIP 1/1. 51
Figure 4.24 | Nucleus entry delay influence STV and DIP production. 52
Figure 4.25 | STV non-productive and 1 STV release simulations and nucleus entry delay. 53
Figure 4.26 | DIP release and non-productive simulations with nucleus entry delay. 54
Figure 4.27 | STV release and non-productive simulations with an induced delay. 55
Figure 4.28 | Distribution of DIP release with induced delay. 56
Figure 4.29 | DIP release results with compiled random and induced delays. 57
Figure 4.30 | DIP release when DIP entry is delayed for DI derived from segments 3 and 4. 58
Figure 4.31 | DIP release when STV entry is delayed for DI derived from segments 3 and 4. 59
Figure 4.32 | Distribution of fused particles at MOI/MODIP 10/10. 60
Figure 4.33 | STV and DIP release with different implementations of particle packaging. 61
Figure 4.34 | Release dynamics using different packaging approaches. 62
Figure 4.35 | Complex formation and release rates using different packaging approaches. 63
Figure 4.36 | STV and DIP release in the Replication Model. 64
Figure 4.37 | Effect of the defective cRNA replication advantage on particle release and genome segment levels. 65

List of Tables

Table 3.1 | Example of the implementation of DIP binding and detachment reactions in the stochastic model..28

Table 4.1 | Number of simulations which showed STV or DIP production despite losing (at least) one genome segment 50

Table A.1.1 | List of implemented parameters of the model 83
Table A.1.2 | List of additional implemented parameters in the Replication Model 85

Table A.2.1 | Number of simulations performed for different initial infection conditions. 86
Table A.2.2 | Number of simulations obtained for each random generated delay................................86
Table A.2.3 | Number of simulations performed for each induced delay and a DIP carrying a DI segment 3. 86
Table A.2.4 | Number of simulations performed for each induced delay and a DIP carrying a DI segment 4. 87

List of Equations

Equation 3.1 | Extracellular STV attachment and detachment to free binding sites 15
Equation 3.2 | Extracellular DIP attachment and detachment to free binding sites 15
Equation 3.3 | STV endocytosis 16
Equation 3.4 | DIP endocytosis 16
Equation 3.5 | STV fusion 16
Equation 3.6 | DIP fusion 16
Equation 3.7 | STV degradation in lysosomes 16
Equation 3.8 | DIP degradation in lysosomes 16
Equation 3.9 | STV nuclear import of vRNPs 17
Equation 3.10 | DIP nuclear import of vRNPs 17
Equation 3.11 | Synthesis of FL cRNAs 17
Equation 3.12 | Synthesis of DI cRNAs 17
Equation 3.13 | RdRp binding to FL cRNAs 18
Equation 3.14 | RdRp binding to DI cRNAs 18
Equation 3.15 | NP attachment to FL RdRp-cRNA complexes 18
Equation 3.16 | NP attachment to DI RdRp-cRNA complexes 18
Equation 3.17 | Synthesis of FL vRNAs 18
Equation 3.18 | Synthesis of DI vRNAs 18
Equation 3.19 | RdRp binding to FL vRNAs 18
Equation 3.20 | RdRp binding to DI vRNAs 18
Equation 3.21 | NP attachment to FL RdRp-vRNA complexes 18
Equation 3.22 | NP attachment to DI RdRp-vRNA complexes 18
Equation 3.23 | M1 binding to FL vRNPs 19
Equation 3.24 | M1 binding to DI vRNPs 19
Equation 3.25 | NEP attachment and nuclear export of FL vRNPs 19
Equation 3.26 | NEP attachment and nuclear export of DI vRNPs 19
Equation 3.27 | Transcription of mRNAs 20
Equation 3.28 | Formation of RdRp 20
Equation 3.29 | Synthesis of HA 20
Equation 3.30 | Synthesis of NP 20

Equation 3.31 | Synthesis of NA 20
Equation 3.32 | Synthesis of M1 21
Equation 3.33 | Synthesis of M2 21
Equation 3.34 | Synthesis of NEP 21
Equation 3.35 | Formation of vRNP-complexes containing eight FL vRNPs 21
Equation 3.36 | Formation of vRNP-complexes containing seven FL and one DI vRNPs 21
Equation 3.37 | Release of progeny STVs 22
Equation 3.38 | Release of progeny DIPs 22
Equation 3.39 | STV release rate 22
Equation 3.40 | DIP release rate 22
Equation 3.41 | Degradation of all molecules containing RNAs 23
Equation 3.42 | Degradation rates assuming the stabilization hypothesis 23
Equation 3.43 | Total number of FL vRNAs of each segment 23
Equation 3.44 | Total number of FL vRNA of segment 3 24
Equation 3.45 | Total number of DI vRNA 24
Equation 3.46 | DIP replication advantage on the DI cRNA synthesis 24
Equation 3.47 | STV complex formation rate 25
Equation 3.48 | DIP complex formation rate 25
Equation 3.49 | Modified STV complex formation rate 25
Equation 3.50 | Modified DIP complex formation rate 25
Equation 3.51 | Release of progeny STVs considering a single-step packaging reaction 26
Equation 3.52 | Release of progeny DIPs considering a single-step packaging reaction 26
Equation 3.53 | Modified STV release rate 27
Equation 3.54 | Modified DIP release rate 27

Equation 4.1 | Average of all FL vRNPs divided by the FL segment 3 ratio and DI vRNP segment 3 over the complementary FL segment ratio 44

List of Abbreviations

cRNA	complementary RNA
DI	defective interfering
DI RNA	defective interfering RNA
DIP	defective interfering particle
FL	full-length
HA	hemagglutinin
hpi	hours post infection
IAV	influenza A virus
M1	matrix protein 1
M2	matrix protein 2
MOI	multiplicity of infection
MODIP	multiplicity of defective interfering particles
mRNA	messenger viral RNA
NA	neuraminidase
NEP	nuclear export protein
NP	nucleoprotein
NS1	non-structural protein 1
ODE	ordinary differential equation
ORF	open reading frame
PA	polymerase acidic protein
PB1	polymerase basic protein 1
PB2	polymerase basic protein 2
RdRp	RNA-dependent RNA polymerase
RNA	ribonucleic acid
RT-qPCR	quantitative reverse transcription polymerase chain reaction
SSA	stochastic simulation algorithm
STV	standard virus
vRNA	viral genomic RNA
vRNP	viral ribonucleoprotein
VSV	vesicular stomatitis virus

List of Symbols

Symbol	Description	Unit
a_j	propensity function	cells·h^{-1}
B_{Hi}	number of free high-affinity binding sites	sites
B_{Lo}	number of free low-affinity binding sites	sites
B_{Hi}^{Tot}	total number of high-affinity binding sites	sites
B_{Lo}^{Tot}	total number of low-affinity binding sites	sites
Cp_{DI}	number of nuclear DI cRNPs	molecules·cell^{-1}
Cp_i	number of nuclear cRNPs of segment i	molecules·cell^{-1}
D_{Rib}	distance between two adjacent ribosomes on an mRNA	nucleotides
D_{Hi}^{Att}	number of DIPs attached to high-affinity binding sites	virions·cell^{-1}
D_{Lo}^{Att}	number of DIPs attached to low-affinity binding sites	virions·cell^{-1}
D^{Cyt}	defective complex of parental vRNPs in the cytoplasm	molecules·cell^{-1}
D_{Cplx}^{Cyt}	defective complex of progeny vRNPs in the cytoplasm	molecules·cell^{-1}
D^{En}	number of DIPs in endosomes	virions·cell^{-1}
D^{Ex}	number of DIPs in the extracellular medium	virions·cell^{-1}
D^{Rel}	number of progeny DIPs	virions·cell^{-1}
F_{Adv}	replication advantage of DI RNA	-
F_{Fus}	fraction of fusion-competent virions	-
F_{Spl7}	fraction of M2-encoding mRNAs	-
F_{Spl8}	fraction of NEP-encoding mRNAs	-
$K_{V^{Rel}}$	influence of protein concentration on virus release	virions
k_{Hi}^{Att}	attachment rate to high-affinity binding sites	site^{-1}·h^{-1}
k_{Lo}^{Att}	attachment rate to low-affinity binding sites	site^{-1}·h^{-1}
k_{M1}^{Bind}	binding rate of M1 to nuclear vRNPs	molecule^{-1}·h^{-1}
k_{NP}^{Bind}	binding rate of NP to RdRp-RNA complexes	molecule^{-1}·h^{-1}

k_{RdRp}^{Bind}	binding rate of RdRp-complexes to vRNA/cRNA	molecule^{-1}·h^{-1}
k^{Cplx}	formation constant of complexes containing eight vRNPs	molecule^{-7}·h^{-1}
k_{En}^{Deg}	degradation rate of virions in lysosomes	h^{-1}
k_{M}^{Deg}	degradation rate of mRNAs	h^{-1}
k_{R}^{Deg}	degradation rate of naked cRNA/vRNA	h^{-1}
k_{Rnp}^{Deg}	degradation rate of RNPs	h^{-1}
k_{RRdRp}^{Deg}	degradation rate of RdRp-RNA complexes	h^{-1}
k_{Hi}^{Dis}	detachment rate from high-affinity binding sites	h^{-1}
k_{Lo}^{Dis}	detachment rate from high-affinity binding sites	h^{-1}
k^{En}	endocytosis rate	h^{-1}
k_{Hi}^{Eq}	equilibrium constant of high-affinity binding sites	site^{-1}
k_{Lo}^{Eq}	equilibrium constant of low-affinity binding sites	site^{-1}
k^{Exp}	rate of NEP binding and nuclear export	molecule^{-1}·h^{-1}
k^{Fus}	fusion with endosomes rate	h^{-1}
k^{Imp}	nuclear import rate	h^{-1}
k^{Rel}	virus release constant	virions·molecule^{-1}·h^{-1}
k_{C}^{Syn}	cRNA synthesis rate	h^{-1}
k_{M}^{Syn}	mRNA synthesis rate	nucleotides·h^{-1}
k_{P}^{Syn}	protein synthesis rate	nucleotides·h^{-1}
k_{V}^{Syn}	vRNA synthesis rate	h^{-1}
L_{i}^{M}	length of the mRNA of segment i	nucleotides
L_{DI}^{V}	length of the vRNA and cRNA of DI segment	nucleotides
L_{i}^{V}	length of the vRNA and cRNA of segment i	nucleotides
N_{P_j}	number of proteins of type $j = \{RdRp, HA, NA, NP, M1, M2, NEP\}$ in a virus particle	molecules·virion^{-1}
N_{M1}^{Nuc}	number of nucleotides bound by one M1 molecule	nucleotides
N_{NP}^{Nuc}	number of nucleotides bound by one NP molecule	nucleotides
P_j	number of proteins of type $j = \{RdRp, HA, NA, NP, M1, M2, NEP\}$	molecules·cell^{-1}
R_j	chemical reaction channel j	-
R_{DI}^{C}	number of naked DI cRNAs	molecules·cell^{-1}
R_{i}^{C}	number of naked cRNAs of segment i	molecules·cell^{-1}
$R_{RdRp,DI}^{C}$	number of DI RdRp-cRNA complexes	molecules·cell^{-1}
$R_{RdRp,i}^{C}$	number of RdRp-cRNA complexes of segment i	molecules·cell^{-1}
r_{STV}^{Cplx}	virion complex formation rate	molecules·cell^{-1}·h^{-1}

r_{DIP}^{Cplx}	DIP complex formation rate	molecules·cell^{-1}·h^{-1}
R_i^M	number of mRNA of segment i	molecules·cell^{-1}
r_D^{Rel}	DIP release rate	virions·cell^{-1}·h^{-1}
r_V^{Rel}	virus release rate	virions·cell^{-1}·h^{-1}
r^{RdRp}	formation rate of polymerase complexes	molecules·cell^{-1}·h^{-1}
R_{DI}^V	number of naked DI vRNAs	molecules·cell^{-1}
R_i^V	number of naked vRNAs of segment i	molecules·cell^{-1}
$R_{RdRp,DI}^V$	number of DI RdRp-vRNA complexes	molecules·cell^{-1}
$R_{RdRp,i}^V$	number of RdRp-vRNA complexes of segment i	molecules·cell^{-1}
$R_{Tot,DI}^V$	total number of DI vRNAs in a cell	molecules·cell^{-1}
$R_{Tot,i}^V$	total number of vRNAs of segment i in a cell	molecules·cell^{-1}
S_i	chemical specie i	-
t	time	h
τ	time step during which multiple reactions can occur simultaneously	h
ν_{ij}	state-change matrix	-
V_{Hi}^{Att}	number of virions attached to high-affinity binding sites	virions·cell^{-1}
V_{Lo}^{Att}	number of virions attached to low-affinity binding sites	virions·cell^{-1}
V^{Cyt}	complex of eight parental vRNPs in the cytoplasm	molecules·cell^{-1}
V_{Cplx}^{Cyt}	complex of eight progeny vRNPs in the cytoplasm	molecules·cell^{-1}
V^{En}	number of virions in endosomes	virions·cell^{-1}
V^{Ex}	number of virions in the extracellular medium	virions·cell^{-1}
V^{Rel}	number of progeny virions	virions·cell^{-1}
$Vp_{M1,DI}^{Cyt}$	cytoplasmatic DI M1-NEP-vRNP complexes	molecules·cell^{-1}
$Vp_{M1,i}^{Cyt}$	cytoplasmatic M1-NEP-vRNP complexes of segment i	molecules·cell^{-1}
Vp_{DI}^{Nuc}	number of nuclear DI vRNPs	molecules·cell^{-1}
Vp_i^{Nuc}	number of nuclear vRNPs of segment i	molecules·cell^{-1}
$Vp_{M1,DI}^{Nuc}$	number of nuclear DI M1-vRNP complexes	molecules·cell^{-1}
$Vp_{M1,i}^{Nuc}$	number of nuclear M1-vRNP complexes of segment i	molecules·cell^{-1}
x	realization of the state vector	-
X	state vector	-

1 Introduction

Influenza A viruses (IAVs) are intracellular pathogens that infect cells and take over the biosynthetic machinery and cell resources to spread the infection by producing progeny viruses. They infect an extensive number of species, e.g. poultry, wild birds, pigs, horses, dogs, sea mammals and humans [1]. An influenza infection in humans causes the flu which is a contamination of the epithelial cells of the upper respiratory tract characterized by symptoms as high fever, dry caught, headache and rhinitis [2]. The seasonal infection in healthy individuals is usually not severe and patients recover after one or two weeks of treatment. However, the elderly, the young and individuals with compromised immune system are most susceptible to this contagious disease which leads to an increased mortality among these risk groups [2]. The World Health Organization estimates that IAV annually causes up to 5 million cases of severe illness and 250 000 to 500 000 deaths [1]. Due to its segmented genome comprising eight single-stranded RNAs, antigenic drift and shift can easily occur during replication resulting in new and more potent IAV strains [3]. Besides the seasonal epidemics, specific mutations which may include a combination of viruses from different host species are responsible for global pandemics which occur every 50-60 years [4]. The most disastrous outburst was the “Spanish flu” in 1918 with an estimated number of 50 million deaths [5]. Moreover, during the most recent pandemic of 2009 22 million cases were reported worldwide and the estimated number of casualties rose up to 203 000 [6].

These dramatic events increased the demand for research in the field of IAV infection with the aim of developing treatment and prevention strategies. Currently various antiviral drugs are available to treat IAV infections, however, the most effective method to prevent severe illness and propagation is vaccination. Usually, researchers are focused on infectious influenza virus particles, i.e. virus particles which are responsible for infecting cells, producing progeny virions and spreading the disease. However, it has been found that in both IAV infection and vaccine production, the majority of progeny particles are non-infectious [7,8]. There are different types of non-infectious particles, but we focused on the study of defective interfering particles (DIPs) which are characterized by carrying an internal deletion in at least one of their genome segments. Since their coding sequence is lacking a part of the genetic information, they are unable to produce all the proteins required for their propagation. Consequently, their replication depends on the coinfection with a complete functional virus that will provide the missing protein(s) [9]. Several experiments show that DIPs can impair the replication and production of IAV [7,10]. Moreover, it has been suggested that these defective particles can also impact other properties of the virus, e.g. its evolution and pathogenicity [11,12]. Since DIPs can interfere with the virus replication and considerably reduce the production of infectious virions, they have been proposed as a potential antiviral agent [13,14]. Recent studies showed that the administration of DIPs in mice and ferrets protects them from severe illness and death [15,16]. Furthermore, it has been suggested that DIP production can potentially overcome the limitations of conventional vaccination methods [13].

In order to find novel and efficient antiviral strategies, an extensive understanding of the complex steps of the viral life cycle is crucial. However, the inherent biological mechanism of DIP interference on IAV replication is not completely understood. Systems biology approaches, which comprise the computational and mathematical modelling of complex biochemical processes, can support the elucidation of the intracellular interactions during DIP replication. These approaches have a special importance for virology since the resources and reactants used in viral experiments can be very expensive. Mathematical models can overcome this burden by reducing costs associated with this research and provide a prediction of the system dynamics in different initial infection conditions. However, is important to notice that the model needs to be supported by experimental data to achieve a reliable description. Therefore, it is essential to validate the model predictions with laboratorial experiments. In the last decades, theoretical studies of intracellular IAV replication have been used for process optimization of vaccine production and the developing new treatment methods [17]. Frequently, these mathematical models are developed assuming that the system dynamics can be described with a deterministic approach. However, stochastic effects have a major impact in systems with a low number of molecules [18]. Since a single virus can infect a cell and replicate, this process is highly susceptible to stochastic fluctuations which are caused by the random nature of biochemical reactions [19]. Such random effects impact virus replication resulting in a wide-spread distribution of virus yields and a large cell-to-cell heterogeneity. These random scenarios can be simulated and analysed using a stochastic modelling approach [20].

2 Theoretical Background

2.1 Discovery of defective interfering particles

Cells infected by IAV release infectious and non-infectious particles. Moreover, the majority of progeny particles are non-infectious and were observed for the first time in 1944 by Friedewald and Pickels during sedimentation experiments using high speed centrifugation [21]. Then, 10 years later von Magnus suggested that these particles might impair influenza virus replication since he observed a reduction in the ratio of infectious over non-infectious virus particles after successive passages in embryonated chicken eggs at high multiplicity of infection (MOI) [21]. In 1970, Huang and Baltimore coined the term “defective interfering particles” (DIPs) to describe these non-infectious virus particles. DIPs do not encode for all viral proteins due to an internal deletion in at least one of their genome segments (“defective”). Therefore, they depend on a coinfection with a completely functional standard virus (STV) that provides the missing resources for replication [22]. Additionally, DIPs impair STV propagation, as shown by von Magnus, by interfering with the regular virus replication (“interfering”). The exact mechanisms of this interference are not fully understood, however, an advantage of the DIP at the replication and/or packaging level was suggested by previous studies [10,23].

2.2 Influenza A virus structure and DIP *de novo* generation

IAV is a member of the *Orthomyxoviridae* family and contains a segmented genome which consists of eight single-stranded viral RNAs (vRNAs) of negative polarity (Figure 2.1A). The genome segments are present inside the spherical virus particle as viral ribonucleoprotein complexes (vRNPs). These complexes include the vRNA which is associated to the polymerase complex (RdRp) and multiple copies of the nucleoprotein (NP) (Figure 2.1C) [24]. Each segment has a double-helical structure and the polymerase is attached to both 5’ and 3’ ends of the vRNA [25]. The NP encapsidation stabilizes the vRNA which prevents degradation processes in the host cell nucleus [26]. The eight genome vRNPs form a “7+1” configuration inside the virus particle: a central segment is surrounded by the other seven vRNPs [27].

Each genome segment encodes for at least one protein which is essential to virus propagation (Figure 2.1B) [28]. The vRNP segments 1 to 3 encode for three protein sub-units which form the RdRp complex: polymerase basic proteins 1 and 2 (PB1 and PB2) and the polymerase acidic protein (PA). NP, which is responsible for stabilizing the vRNA, is derived from segment 5. The virus envelope surface contains the glycoproteins hemagglutinin (HA) and neuraminidase (NA) which are encoded by segments 4 and 6, respectively. Furthermore, the matrix protein 2 (M2) is also located on the lipid membrane surface. M1 forms a layer underneath the virus envelope and is also associated to vRNPs. Both proteins are transcribed from spliced viral mRNA of segment 7. The shortest genome segment 8 encodes for the nuclear export protein (NEP) and the non-structural protein 1 (NS1). Other viral proteins were observed in the IAV particles, however, they are not essential for virus replication as they were only expressed under laboratory experiments and in a few virus strains [29].

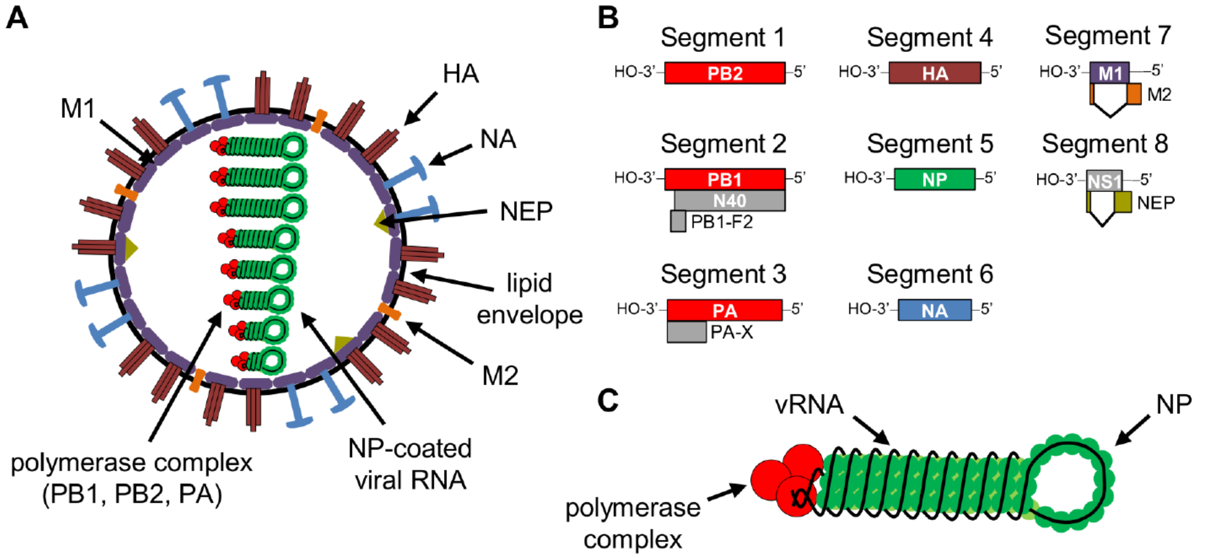


Figure 2.1 | Influenza A virus particle, genome structure and encoding proteins. (A) Structure and proteins of an influenza A virus particle. The represented proteins are: PB – polymerase basic protein, PA – polymerase acidic protein, HA – hemagglutinin, NP – nucleoprotein, NA – neuraminidase, M – matrix protein, NEP – nuclear export protein. (B) Diagram of different genome segments encoding viral proteins. The boxes represent the encoded proteins and the lines at the end of each box are the non-coding regions. The V-shapes indicate the introns of the spliced mRNAs of segment 7 and 8. NS1 – non-structural protein 1. (C) Structure of an influenza A virus ribonucleoprotein (vRNP). Figure taken from the PhD dissertation of Heldt [1].

DIPs carry at least one defective interfering (DI) segment which contains internal deletions in the coding sequence of its genome (Figure 2.2A). These deletions can vary in size and affect different segments [8,12]. However, it has been found that deletions on segments 1 to 3, which encode for the three RdRp subunits, are most common [30,31]. The average size of deletions is between 300 to 500 nucleotides (nt), but DI vRNAs can lack more than 80% of their original functional segment [12,30]. Consequently, DIPs are incapable to replicate on their own and require a coinfection with a STV which provides the missing protein(s). Since the DI segment has the 3' and 5' promoters which enable the polymerase attachment, DI vRNA replication is possible (Figure 2.2B) [32]. Furthermore, transcription of DI mRNA can occur and truncated versions of proteins, which usually lost their function, can be synthesized. Due to the double-helical structure of vRNPs, it has been suggested that DIPs might be generated by an erroneous translocation of the viral polymerase during replication [8,12]. This can be caused by the premature dissociation of the RdRp from the template and reattachment further on the sequence space (Figure 2.2C). However, the molecular mechanism of the DIP *de novo* generation is still largely elusive. Concerning this matter, our research group is currently modelling and developing different hypothesis to further understand how DIPs are generated.

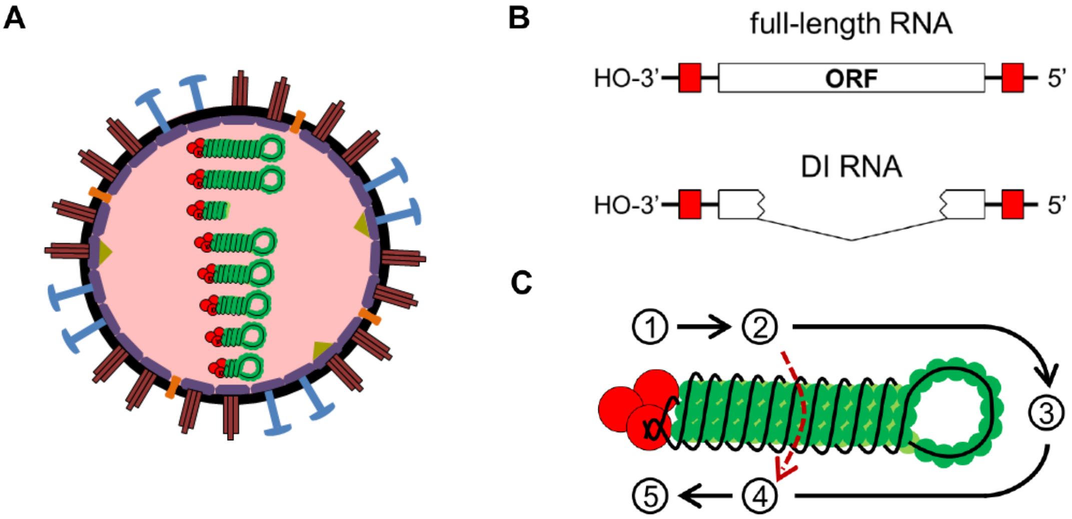


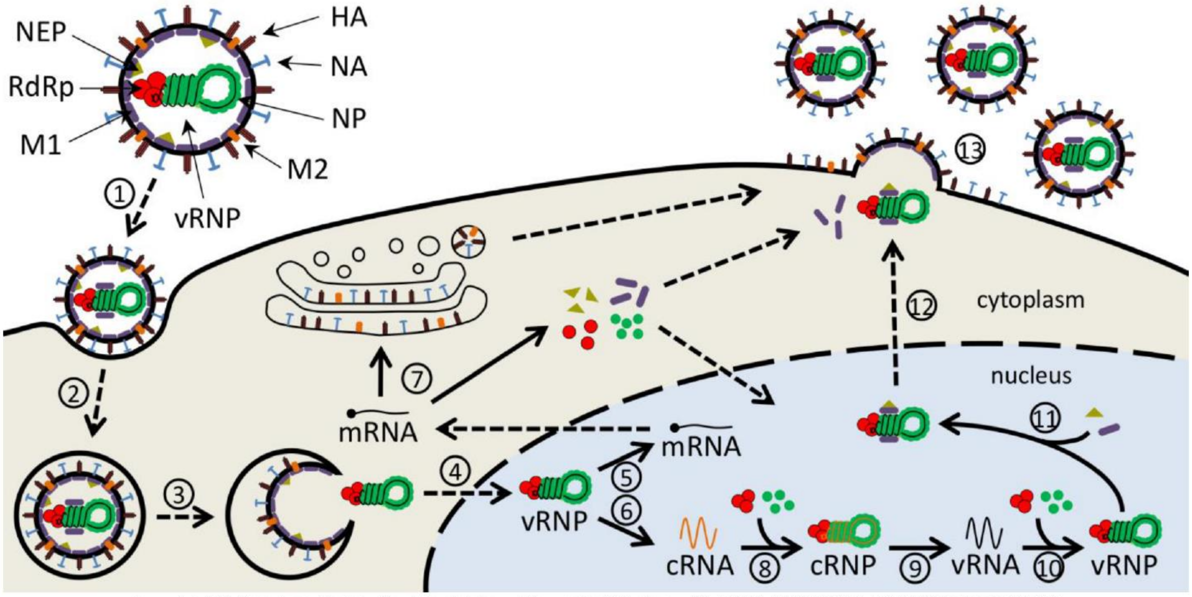
Figure 2.2 | Structure of a defective interfering particle (DIP) and *de novo* generation mechanism of defective RNA. (A) Morphology of a DIP with a defective segment 3 (B) Structure of a full-length (FL) and defective interfering (DI) RNA. The functional RNA has an open reading frame (ORF) represented by the straight lines. The defective RNA carries an internal deletion in its segment which is indicated by the V-shape line. The red boxes represent the terminal promoter sequences. (C) Potential mechanism for DI RNA *de novo* generation. The polymerase normal pathway along the template is represented by the sequential numbers and black arrows. The translocation of the polymerase with dissociation and reattachment at number 2 and 4, respectively deletes the internal coding sequence and originate DIPs. This path is indicated by the red dashed line. Figure adapted from the PhD dissertation of Heldt [1].

2.3 Influenza A virus intracellular life cycle

The deterministic intracellular model of IAV replication was developed by Heldt *et al.* and describes the IAV life cycle in a single mammalian cell using ordinary differential equations (ODEs) [33]. The detailed description of this model is provided by Frank Stefan Heldt as part of his PhD dissertation [1]. The IAV life cycle comprises the virus entry into the host cell, the vRNA replication and synthesis of all viral proteins and the release of progeny virions into the extracellular space. We assumed in our model that the core mechanisms for STV and DIP replication are very similar. In this section we explain the essential steps of the viral life cycle implemented in the original model (Figure 2.3) and in the next section we will highlight the adjustments for DIP replication.

Virus entry

The IAV entry into the host cell is initiated by the interaction of specific viral envelope proteins with the receptor binding sites of the cell: HA protein attaches to neuraminic acids (sialic acids) on the host cell surface which promotes virus entry [34]. An endosome is formed on the cell surface which induces virus entry through a receptor-mediated endocytosis. The virus goes through endocytosis until the acidification in late endosomes which enables the virus envelope to fuse with the endosomal membrane [34,35]. This reaction results in viral uncoating, i.e. protons enter the virus particle through M2 ion channels causing the vRNPs to detach from M1 proteins which promotes the release of vRNPs into the host cell cytoplasm. During endocytosis, the virus can either successfully fuse with the viral envelope and release its vRNPs or fail to fuse and the virus particle will eventually be degraded. The fusion process in late endosomes is the only step during cell entry in which the virus can be degraded in the model, which can prevent further replication. Note that other biological processes can cause virus degradation, e.g. fail to attach the cell surface or cytoplasmatic transport, however such mechanisms were neglected in the model.



Copyright © American Society for Microbiology, [Journal of Virology, 86, 2012, 7806-7817, 10.1128/JVI.00080-12]

Figure 2.3 | Diagram of the influenza A virus life cycle. To simplify the figure, only one of the eight vRNPs is represented and non-structural proteins were omitted. Solid arrows indicate synthesis or protein attachment and dashed arrows represent transport processes. The different life cycle steps are indicated by the numbers: 1 – attachment, 2 – endocytosis, 3 – fusion in late endosomes, 4 – nuclear import, 5 – transcription, 6 – replication (cRNA synthesis), 7 – protein translation, 8 – cRNA encapsidation, 9 – replication (vRNA synthesis), 10 – vRNA encapsidation, 11 – M1 and NEP binding, 12 – nuclear export, 13 – virus assembly and budding. Figure taken from the PhD dissertation of Heldt [1].

Nuclear import of vRNPs

After the viral fusion the vRNPs travel through the cytoplasm. Since the virus does not possess all the enzymes and resources necessary for its replication, it requires to take control over the biosynthesis machinery in a host cell’s nucleus to promote its propagation. Experimental data suggest that the eight vRNPs are transported together across the cytoplasm and only separate when they reach the karyoplasm [36]. The model does not directly consider cell compartments, however, it contains an inactive cytoplasmic state of vRNPs which is converted into separated nuclear vRNP segments when the genomes reach the nucleus. The nuclear import of vRNPs occurs via an active transport mechanism mediated by transport receptors located on the nucleus membrane [37,38]. Inside the nucleus each vRNP segment behaves as an independent functional structure [36].

Viral replication

Once inside the nucleus, the eight vRNP segments are used as templates to produce viral mRNA and the complementary RNA (cRNA). However, the mechanism which coordinates the synthesis of both molecules is still under debate and different hypothesis were proposed in the past decades [25]. The most accepted theories are the switching and the stabilization hypothesis.

The switching hypothesis suggests that before cRNA synthesis, an initial round of viral protein translation occurs which promotes the accumulation of soluble NP, i.e. free NP that is not attached to vRNAs. This hypothesis proposed that NP switches the vRNP activity from transcription to replication (Figure 2.4A) [39]. The theory is supported by laboratory experiments which show that the synthesis of cRNA depends on NP [40] and a study with temperature-sensitive NP mutants which revealed a reduced cRNA synthesis and normal viral mRNA production [41]. However, other experimental data showed that over expressed NP did not increase cRNA replication [42]. By contrast, the stabilization hypothesis suggests that both viral mRNA and cRNA are synthesised early on after vRNPs nuclear import. This theory describes that due to the attachment of RdRp and NP the nascent cRNA forms cRNP which prevents its degradation by cellular nucleases, i.e. cRNA does not accumulate unless these proteins stabilize it (Figure 2.4B) [43]. This hypothesis is supported by *in vitro* experiments which showed that both transcription and replication occur early on and in the absence of free NP [44]. The vRNA synthesis proceeds from the positive-stranded cRNP which was formed previously. In a similar way, the vRNA needs to be stabilized by the viral polymerase and NP such that vRNPs can be formed in the host nucleus [45]. Is important to notice that in the model all intermediate molecules (cRNA, vRNA, cRNP and vRNP) can be degraded by nucleases during replication, however, the stabilized RNPs are degraded significantly slower.

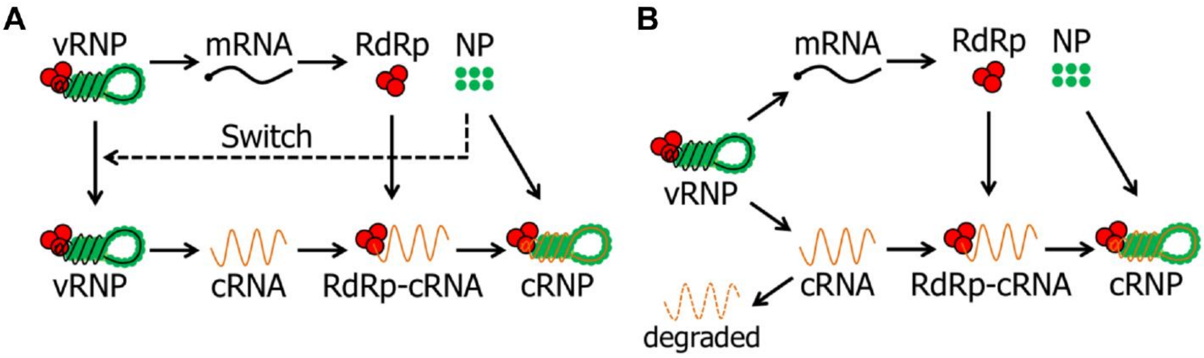


Figure 2.4 | Different hypothesis for the transition mechanism from transcription to replication. (A) Switching hypothesis – the transcription of mRNA by the polymerase occurs early in infection. Accumulation of NP switches the polymerase toward replication. (B) Stabilization hypothesis – vRNPs engage both transcription and replication early in infection. cRNA requires stabilization of viral polymerase and NP to prevent its degradation by cellular nucleases. Figure taken from the PhD dissertation of Heldt [1].

Transcription and protein synthesis

Experimental data showed that in infections with IAV the number of viral mRNA is negatively correlated to its length, i.e. shorter mRNAs are in higher abundance [46]. The transcribed viral mRNAs are then transported to the cytoplasm where the protein translation will take place. The viral mRNA needs to compete for transcripts and resources of the cell. To achieve this, several mechanisms are triggered that enables the viral mRNA to have preferential access to these resources and take over the cell translation machinery [47]. The synthesised proteins can engage two different pathways: they can either enter the cell nucleus or be transported to the plasma membrane [37,38]. RdRp and NP enter the nucleus to stabilize the newly produced vRNA and cRNA. Furthermore, a fraction of M1 and NEP takes the first path, and the remaining fraction takes the second route since they are involved in the nuclear export of vRNPs as well as in the process of virus particle release. Finally, the surface proteins HA, NA and M2 are processed in the endoplasmic reticulum and Golgi complex and transported to the plasma membrane.

Nuclear export of vRNPs

After replication, the newly synthesised vRNPs can either be used as template for viral mRNA and cRNA production or be exported from the nucleus to be assembled into progeny virions. It has been suggested that the nuclear vRNPs can be divided into two types: the active vRNPs which participate in viral mRNA and cRNA synthesis and the inactive ones that leave the nucleus and produce new viral particles [48]. Laboratory experiments showed that overexpression of M1 protein inhibit the viral mRNA transcription suggesting that this protein is a the mediator for vRNP inactivation [49,50]. There is further evidence that M1 binds to NP [49] and also to RNA [51] to promote the nuclear export of vRNPs. Additional studies illustrate the importance of M1: experiments with a fluorescence marker for NP, which can be used to localize vRNPs, showed that cells that lack M1 or were treated with antibodies which retain M1 in the cytoplasm, lead to an accumulation of vRNPs inside the nucleus [52]. Furthermore, another study, in which an M1 production inhibitor was applied, resulted in the nuclear retention of vRNPs [53]. It was also suggested that NEP participates in the export process: experimental work revealed that no vRNP was exported from the nucleus after injection of anti-NEP antibodies [54]. However, it has also been observed that even in complete absence of NEP nuclear export still occurs [55]. This results indicates that NEP is not a determinant factor for vRNP nuclear export or at least suggests that this protein is not required in high quantities [39].

Viral packaging and release

After replication and protein synthesis, the intracellular IAV continues its life cycle with the processes of virus packaging and finally the release from the host cell. To form a functional virus particle, one copy of each vRNP genome segment is required. However, whether the eight vRNP segments are assembled by a random mechanism or in a segment-specific packaging has been debated for several years [56]. Experimental studies showed that the genome segments form a “7+1” configuration where seven segments display a ring around a central vRNP [27,57]. Further experiments indicated that this complex configuration might be mediated by RNA-RNA interactions between the IAV genome segments [58,59]. In addition, it has been observed that DIPs compete for their full-length (FL) segments in a segment-specific way during packaging [7,10]. These experimental results provide strong evidence that the viral genome assembly is controlled by a specific packaging mechanism. The virus assembly includes the aggregation of vRNP segments and the packaging of the structural proteins. It has been suggested that the M1 protein has a very important role during the virus assembly since observations showed it attaches to vRNPs, forms a layer underneath the virus envelope and interacts with HA and NA facilitating an association with the cell membrane [60]. The virus particle release is initiated by the formation of a bud which is mediated by the interaction of HA, NA and M2 with the cell membrane. Then, the bud is extended and the vRNPs segments are incorporated into the progeny particle. In the final steps M2 promotes the membrane curvature and the separation of the budding particle [61]. Once in the cell exterior, NA removes the sialic acids from the viral envelope which prevents HA from binding to other progeny particles.

2.4 DIP interference

Since influenza virus DIPs carry an internal deletion in at least one of their genome segments, they do not possess the complete coding sequence to synthesise all proteins. Consequently, they depend on coinfections with STVs that provide the missing protein(s) to form progeny particles. During these coinfections, DI RNA interferes with the STV replication which results in a reduction of the infectious virus titer and mainly progeny DIPs are produced [11,13]. However, the molecular mechanism inducing this interference is still not completely understood [12].

Several experiments showed a preferential amplification of the DI RNAs over the FL genome segments [7,62] which indicates that the interference can emerge during RNA replication. cRNA synthesis has been suggested to be the step comprising the replication advantage [63]. In this regard, it has been proposed that the reduced length of the DI vRNA induces this advantage, because synthesis processes could occur faster [12,13]. Since the viral polymerase synthesizes a constant number of nucleotides per unit time, shorter RNAs can be generated in greater abundance if the necessary resources are available. This theory is in agreement with experimental data from a dual luciferase reporter assay for RNA replication [64]. In this experiment, the shorter of the two luciferase-encoding influenza virus-like RNAs showed a higher interference potential on luciferase expression when compared to the longer reporter segment. Although shorter DI segments are assumed to be synthesised faster, large deletions can also disrupt terminal packaging signals [13]. This indicates that there may exist an optimal length of DI RNA at which DIPs achieve the maximum replication advantage and are still efficiently packaged into progeny particles. However, it has been observed that some DI RNAs do not accumulate to high levels in coinfected cells [7] which indicates that RNA length is not the only determinant factor of interference.

Other interference mechanism has been suggested to be the competition for limited viral and/or cellular resources [13]. Experimental studies support this assumption, because the competition between DI RNA and FL segments for viral polymerases has been observed in vesicular stomatitis virus (VSV) and influenza virus. Furthermore, different DI segments showed a more efficient packaging which results in an increased production of DIPs to the detriment of STV release [10,23]. Besides the competition for resources, it has also been hypothesised that the proteins synthesised from the DI RNA could contribute to the interference mechanism. Since the DI segment has the 3' and 5' promoters, the polymerase can transcribe DI mRNAs and synthesise truncated versions of proteins [62]. However, these proteins are shorter than the functional ones, are therefore very likely to lose their function and do not show an increased interference potential [12,62].

2.5 Therapeutic potential of DIPs

In recent years, DIPs have been suggested to be used in antiviral therapies due to their interference with STV replication which efficiently reduces the infectious virus yield [13,14]. Experimental results revealed that the administration of DIPs prevented mice and ferrets from developing lethal infections caused by different IAV strains [15,16]. More precisely, it reduced fever, weight loss, respiratory symptoms and the infectious load. Additionally, it was shown that the administration of DIPs between 24 and 48h post virus infection completely prevented clinical disease and death in mice [15]. Furthermore, it was observed that coinfections with STV and DIPs activate the immune response inducing anybody production [16]. Hence, research results suggest that DIPs have a major potential to be used for the prevention of IAV infections and as antiviral medication, especially when the infectious strain is unknown or resistant to other antiviral drugs [15].

2.6 Intracellular deterministic model of influenza A virus and DIP replication

A deterministic intracellular model of DIP replication was developed by Laske *et al.* [9] which constitutes an extension from the previously implemented deterministic model of the IAV life cycle [33]. This model provided novel insights about the factors which influence DIP production focussing in the replication advantage, coinfection timing and DI RNA originated from different genome segments. The model comprises a DIP carrying a DI vRNA which has a replication advantage at the cRNA synthesis level due to its reduced length.

Using the deterministic model of DIP replication, Laske *et al.* evaluated the effect of the length dependent replication advantage on STV and DIP production. The obtained results suggested that there is an optimal DI RNA synthesis rate which results in a maximum DIP production. Model simulations showed that when the replication advantage increases above its optimum, the production of DI segments consumes an increasing amount of proteins which leads to the depletion of NP.

Consequently, the DIP release is reduced due to protein limitation. The authors concluded that at the optimum replication advantage, a balance between the FL and DI RNA synthesis is established such that the pool of viral resources would not be depleted.

Additionally, the authors evaluated the impact of different time delays of successive confections with STV and DIP. The results showed that DIP produced progeny particles only if DIP infection was delayed no longer than 3 hours after STV infection. It has been suggested that when the STV replication already progressed too far, DIP replication does not occur. In this case, the high levels of M1 and NEP accumulated in the cytoplasm promote the attachment of these proteins to the DI vRNP which become inactive. To compile these findings, simulations with different DI segments were performed. The model results showed that a defect in segment 3 provides an advantage in DIP replication compared to a DIPs carrying a defective segment 4 when STV coinfection is delayed more than 3 hours. Furthermore, simulations with DI segments 5 showed to be less productive than a DI segment 3 for all tested delays. The authors concluded that defects in segments which encode for proteins of the polymerase complex can provide an advantage in DIP replication. Hence, these DIPs can overcome DIPs with other defective segments after several passages and emerge as the dominant species, explaining why most DI RNAs of IAV are originated from segments 1 to 3.

2.7 Intracellular stochastic model of influenza A virus replication

Cell-to-cell variability is a phenomenon usually observed in nature and most of these fluctuations are caused by the random nature of biochemical reactions [19]. By contrast to the deterministic model, which considers that the reactants in the system are continuous variables, the stochastic approach defines them as discrete numbers of molecules that can be changed by random events. Stochastic models can provide valuable information about how the system dynamics respond to the random effects inherent to biological processes [65]. Furthermore, for the same initial conditions the deterministic model provides one definitive result, however, stochastic model simulations result in a distribution that reflects the variance introduced by the randomness of biochemical reactions [3]. The stochastic effects are more pronounced in systems with a low number of molecules [18]. This is the case in virus infections since one single viral particle can be sufficient to infect an entire organism. However, stochastic simulations are usually more computationally demanding since the credibility of the stochastic results is related to the number of simulations performed [66].

A stochastic intracellular model of IAV replication was implemented by Heldt & Dorl [3,67] which is a new approach to the previous deterministic model of the IAV life cycle [33]. Model simulation results showed that the randomness of biochemical reactions, i.e. stochastic effects, have a significant impact on IAV replication at the single-cell level. The experiments performed in this study include the single-cell analysis of viral replication dynamics using real-time RT-qPCR and the observed results were complemented with the stochastic intracellular model. The authors focussed on different sources of noise affecting virus replication, genome segmentation, infections at low MOI and stochastic effects during virus entry.

The stochastic model of IAV replication showed that one of the major sources of variability in IAV infection is the segmented genome since each segment is affected by stochastic effects independently. Furthermore, simulations indicated that the viral replication process, which is assumed to be autocatalytic, amplifies heterogeneity in the levels of viral RNA. The authors suggested that the genome segmentation combined with the autocatalytic synthesis of vRNA from cRNA and vice versa, are the cause for the large fluctuations detected in the model for RNA and proteins levels which affect virus production.

The model predicted that nearly 93% of simulations performed at MOI 1 did not release any progeny virion. In half of these simulation runs the virus failed to fuse its envelope with the endosomal membrane. Another reason for this high percentage of non-productive infections is segment loss which is caused by the degradation of vRNA in the nucleus by cellular nucleases. If the vRNA of one or more genome segments is completely degraded before replication, the virus lacks part of its genome and, consequently, is unable to produce any viral particle. Due to these observations, the authors concluded that fusion failure and segment loss events are the major steps in virus infection that induce the cell-to-cell heterogeneity observed in IAV production.

2.8 Stochastic simulation algorithms

The stochastic chemical kinetics considers that in a well-stirred system N chemical species $\{S_1, S_2, \dots, S_N\}$ interact through M chemical reactions $\{R_1, R_2, \dots, R_M\}$. $X_i(t)$ denotes the number of molecules of species S_i in the system at time t. The goal of a stochastic simulation is to estimate the state vector $X(t) \equiv (X_1(t), X_2(t), \dots, X_N(t))$ knowing that the system was in state $X(t_0) = x_0$ at some initial time t_0 [20]. The changes in the species populations are caused by the chemical reactions. In this context, stochastic simulations use state-change vectors and propensity functions to characterize each reaction channel R_j. The change in the system is not given by ODEs as in the deterministic model, but rather using a state change vector which is represented as: $\nu_j \equiv (\nu_{1j}, \nu_{2j}, \dots, \nu_{2j})$, where ν_{ij} is the change in the molecular population of S_i induced by a single R_j reaction. If the system is in state x and one R_j reaction occurs, the system will jump to state $x + \nu_j$. Furthermore, the reaction constants are described in the form of propensity functions rather than reaction rates as in the deterministic approach. The propensity function is represented as: $a_j(x)\,dt$, which is the probability, given $X(t) = x$, that one R_j reaction will occur in the next infinitesimal time interval $[t, t + dt]$ [20]. The time to the next reaction occurring is an exponentially distributed random variable with average $1/a_0(x)$ and $a_0(x) = \sum a_j(x)$, which represents the sum of all propensity functions.

These concepts are included in the chemical master equation (CME) which determines the probability that each species will have a specific molecular population at a given future time point [68]. However, an analytical solution for the CME is usually quite complicated to obtain even for simple systems. To approximate a solution, Gillespie proposed a Monte Carlo procedure to simulate time trajectories of the molecular populations described by the CME. This procedure was named stochastic simulation algorithm (SSA) and has been used as the basis for stochastic simulations until today [69,70].

The SSA is usually preferred over other iteration methods due to its accurate results and simple coding [71]. However, it simulates every single reaction separately which makes it very inefficient for systems with a high number of molecules, as is the case for IAV infection. During later stages of infection, this problem becomes even more challenging as the amount of interacting molecules increases rapidly from a low number to millions of molecules. As the population of some species increases, the values of the propensity functions increase substantially [3]. Consequently, $a_0(x)$ achieves higher values and the time until the next reaction is on average reduced which results in more simulations per time unit. To improve computational speed, Gillespie proposed an extension to the SSA named Tau-leaping in 2001 [72]. This algorithm approximates the result of the SSA by advancing the system by a time step τ during which multiple reaction events can occur simultaneously. Hence, computational efficiency can be increased by this method as it allows to aggregate a high number of individual reactions in a single computational step. However, the leap condition has to be considered: τ has to be small enough that none of the states X_i in the system change by an amount above a certain threshold, which is normally defined as 3%.

As mentioned before, the stochastic algorithm to simulate influenza virus infection was previously developed by Heldt & Dorl and we opted to use this algorithm as a basis for our model [3]. The SSA implementation followed the steps described in the original work of Gillespie [70] and the Tau-leaping procedure was based on the work of Cao *et al.* [73], which has been indicated as the most efficient method for explicit tau-leaping [20]. The step-by-step explanation of the algorithm procedure can be consulted in the Bachelor thesis of Sebastian Dorl [3].

3 Models and Methods

3.1 Implementation of biochemical reactions

Our stochastic model of DIP replication is based on a previous stochastic model of the intracellular life cycle of IAV developed by Heldt & Dorl [3,67]. We extended this model by implementing the mechanism of DIP interference described in the deterministic model of DIP replication implemented by Laske *et al.* [9]. For most of our simulations, we established a DI RNA derived from segment 3, which encodes for the polymerase acidic protein (PA). This specific segment was chosen, because it is one of the polymerase-encoding segments which are the most abundant DI segments found in IAV production [30,31]. Furthermore, the previous deterministic DIP model also established a DI RNA for segment 3 and to enable the comparison between both models we opted to apply the same modification in our implementation. In the following, we provide a complete list of the biochemical equations established in our model. For a comparison between different DI segments, we also applied a DI RNA derived from segment 4 which encodes for HA.

Virus entry

The extracellular STVs (V^{Ex}) and DIPs (D^{Ex}) bind to free binding sites, i.e. sialic acids, on the cell surface (B_n),

$$V^{Ex} + B_n \underset{k_n^{Dis}}{\overset{k_n^{Att}}{\rightleftarrows}} V_n^{Att} \tag{3.1}$$

$$D^{Ex} + B_n \underset{k_n^{Dis}}{\overset{k_n^{Att}}{\rightleftarrows}} D_n^{Att} \tag{3.2}$$

$$with \quad B_n = B_n^{Tot} - V_n^{Att} - D_n^{Att} \quad and \quad k_n^{Dis} = \frac{k_n^{Att}}{k_n^{Eq}}, \quad n \in \{\, Hi, Lo \,\}$$

V_n^{Att} and D_n^{Att} denote the STVs and DIPs, respectively, which are attached to binding sites of type n. Experimental data showed that two types of binding sites can be distinguished: high affinity sites

$(n = Hi)$ which represents the attachment to sialic acids and low affinity $(n = Lo)$ that correspond to binding sites with less specific interactions [35]. The number of free binding sites is calculated by subtracting the number of attached particles from the total number of binding sites (B_n^{Tot}). The particles can attach to these sites with rate k_n^{Att} or dissociate from them with rate k_n^{Dis}, which is calculated from the equilibrium constant k_n^{Eq}.

After attachment to the cell surface, the particles are imported by receptor-mediated endocytosis,

$$V_n^{Att} \xrightarrow{k^{En}} V^{En} + B_n \tag{3.3}$$

$$D_n^{Att} \xrightarrow{k^{En}} D^{En} + B_n \tag{3.4}$$

V^{En} and D^{En} represent the STVs and DIPs in endosomes and k^{En} is the endocytosis rate. Since our model assumes a fast recycling of receptors, the binding sites become available instantaneously when particles begin endocytosis [67].

The particles in endosomes can either fuse with the endosomal membrane or be degraded in lysosomes,

$$V^{En} \xrightarrow{k^{Fus}} V^{Cyt} \tag{3.5}$$

$$D^{En} \xrightarrow{k^{Fus}} D^{Cyt} \tag{3.6}$$

$$V^{En} \xrightarrow{k_{En}^{Deg}} \emptyset \tag{3.7}$$

$$D^{En} \xrightarrow{k_{En}^{Deg}} \emptyset \tag{3.8}$$

$$with \quad k_{En}^{Deg} = \frac{1 - F_{Fus}}{F_{Fus}} k^{Fus}, \quad and \quad 0 < F_{Fus} \leq 1$$

k^{Fus} and k_{En}^{Deg} are the rates of fusion and degradation, respectively. The F_{Fus} represents the fraction of fusion-competent particles, which is derived from experiments showing that half of the infecting virions fail to reach the nucleus [74]. After fusion, STVs and DIPs release their genome into the cytoplasm. Experimental data showed that upon fusion, the eight genome segments of IAV travel together through the cytoplasm until they reach the nucleus [36]. Hence, we assumed that both particles deliver a complex which includes all eight genome segments. These complexes that carry the genomes of either an STV or DIP are denoted as V^{Cyt} and D^{Cyt} either is they are derived from STVs or DIPs, respectively.

Nuclear import of vRNPs

When the vRNP complex enters the nucleus, the segments separate and act as independent functional units for replication and transcription.

$$V^{Cyt} \xrightarrow{k^{Imp}} Vp_1^{Nuc} + Vp_2^{Nuc} + Vp_3^{Nuc} + \cdots + Vp_8^{Nuc} \tag{3.9}$$

$$D^{Cyt} \xrightarrow{k^{Imp}} Vp_1^{Nuc} + Vp_2^{Nuc} + Vp_4^{Nuc} + \cdots + Vp_8^{Nuc} + Vp_{DI}^{Nuc} \tag{3.10}$$

The STV complex splits into its eight FL vRNPs $\left(Vp_i^{Nuc}\right)$ with $i = \{1, \dots, 8\}$, and DIP complex separates its seven FL segments and a DI vRNP of segment 3 (Vp_{DI}^{Nuc}). k^{Imp} is the nuclear import rate.

Viral replication

Once inside the nucleus, the vRNPs are used as template for replication.

$$Vp_i^{Nuc} \xrightarrow{k_C^{Syn}} Vp_i^{Nuc} + R_i^C \tag{3.11}$$

$$Vp_{DI}^{Nuc} \xrightarrow{(F_{Adv}+1)k_C^{Syn}} Vp_{DI}^{Nuc} + R_{DI}^C \tag{3.12}$$

$$with \quad F_{Adv} = \left(\frac{L_3^V}{L_{DI}^V} - 1\right), \quad i = \{1, \dots, 8\}$$

To replicate the viral genome, the cRNA of each FL segment $\left(R_i^C\right)$ is synthesised with rate $\left(k_C^{syn}\right)$. Based on experimental data, we assumed that the DI cRNA $\left(R_{DI}^C\right)$ has a replication advantage over the FL segments [63]. This advantage is represented as an advantage factor (F_{Adv}) and its implementation follows the description of the previous deterministic model of DIP replication [9]. The value of this factor is derived from the difference in length of segment 3 vRNA (L_3^V) and of the DI vRNA (L_{DI}^V).

Based on the stabilization hypothesis, the newly synthesised cRNA needs to be stabilized by the binding of polymerase complexes $\left(P_{RdRp}\right)$ and NP (P_{NP}) to prevent its degradation by cellular nucleases [43].

$$R_i^C + P_{RdRp} \xrightarrow{k_{RdRp}^{Bind}} R_{RdRp,i}^C \tag{3.13}$$

$$R_{DI}^C + P_{RdRp} \xrightarrow{k_{RdRp}^{Bind}} R_{RdRp,DI}^C \tag{3.14}$$

$$R_{RdRp,i}^C + \left[L_i^V / N_{NP}^{Nuc}\right] P_{NP} \xrightarrow{k_{NP}^{Bind}} Cp_i \tag{3.15}$$

$$R_{RdRp,DI}^C + [L_{DI}^V / N_{NP}^{Nuc}] P_{NP} \xrightarrow{k_{NP}^{Bind}} Cp_{DI} \tag{3.16}$$

The RdRp attaches to FL and DI cRNA with rate k_{RdRp}^{Bind} and forms FL and DI RdRp-cRNA complexes ($R_{RdRp,i}^C$ and $R_{RdRp,DI}^C$ respectively). Subsequently, multiple molecules of NP bind to these complexes with rate k_{NP}^{Bind} and form FL and DI cRNPs (Cp_i and Cp_{DI} respectively). Since NP needs to cover the complete vRNA (see Figure 2.1C), the number of NP molecules in one cRNP is calculated based on the length of the segment (L_i^V or L_{DI}^V) and the number of nucleotides bound by one NP molecule (N_{NP}^{Nuc}).

The second step of replication uses the newly formed cRNP as a template,

$$Cp_i \xrightarrow{k_V^{Syn}} Cp_i + R_i^V \tag{3.17}$$

$$Cp_{DI} \xrightarrow{k_V^{Syn}} Cp_{DI} + R_{DI}^V \tag{3.18}$$

the FL and DI vRNAs (R_i^V and R_{DI}^V , respectively) are synthesised with rate k_V^{Syn}.

Similar to the encapsidation of cRNAs, the progeny vRNAs are stabilized by the attachment of polymerase complexes $\left(P_{RdRp}\right)$ and NP (P_{NP}),

$$R_i^V + P_{RdRp} \xrightarrow{k_{RdRp}^{Bind}} R_{RdRp,i}^V \tag{3.19}$$

$$R_{DI}^V + P_{RdRp} \xrightarrow{k_{RdRp}^{Bind}} R_{RdRp,DI}^V \tag{3.20}$$

$$R_{RdRp,i}^V + \left[L_i^V / N_{NP}^{Nuc}\right] P_{NP} \xrightarrow{k_{NP}^{Bind}} Vp_i^{Nuc} \tag{3.21}$$

$$R_{RdRp,DI}^V + [L_{DI}^V / N_{NP}^{Nuc}] P_{NP} \xrightarrow{k_{NP}^{Bind}} Vp_{DI}^{Nuc} \tag{3.22}$$

where $R_{RdRp,i}^V$ and $R_{RdRp,DI}^V$ denote the FL and DI RdRp-vRNA complexes, respectively. The binding of NP induces the formation of progeny FL and DI vRNP segments (Vp_i^{Nuc} and Vp_{DI}^{Nuc} respectively).

Nuclear export of vRNPs

The newly produced vRNPs can either be used as templates for viral mRNA and cRNA synthesis or be exported from nucleus to be included into new progeny particles. Our model describes nuclear export as a two-step process. First, the M1 protein (P_{M1}) binds to the vRNP segments and forms either functional $\left(Vp_{M1,i}^{Nuc}\right)$ or defective $\left(Vp_{M1,DI}^{Nuc}\right)$ M1-vRNP complexes. As a preparation for nuclear export, M1 binding is responsible for the inactivation of vRNPs which prevents viral mRNA and cRNA synthesis [49,50].

$$Vp_i^{Nuc} + \left[L_i^V/N_{M1}^{Nuc}\right]P_{M1} \xrightarrow{k_{M1}^{Bind}} Vp_{M1,i}^{Nuc} \tag{3.23}$$

$$Vp_{DI}^{Nuc} + [L_{DI}^V/N_{M1}^{Nuc}]P_{M1} \xrightarrow{k_{M1}^{Bind}} Vp_{M1,DI}^{Nuc} \tag{3.24}$$

k_{M1}^{Bind} denotes the M1 binding rate. The number of M1 proteins necessary to cover the vRNP is calculated based on the length of the segment (L_i^V or L_{DI}^V) and the number of nucleotides bound by one M1 molecule (N_{M1}^{Nuc}).

In the final step of nuclear export, NEP binds to the inactivated vRNPs and induces their transport to the cytoplasm.

$$Vp_{M1,i}^{Nuc} + P_{NEP} \xrightarrow{k^{Exp}} Vp_{M1,i}^{Cyt} \tag{3.25}$$

$$Vp_{M1,DI}^{Nuc} + P_{NEP} \xrightarrow{k^{Exp}} Vp_{M1,DI}^{Cyt} \tag{3.26}$$

$Vp_{M1,i}^{Cyt}$ and $Vp_{M1,DI}^{Cyt}$ represent the NEP-M1-vRNP complex in the cytoplasm derived from a FL or DI segment, respectively. k^{Exp} is the export rate. It has been suggested that NEP is not required in stochiometric quantities [39]. Therefore, we established that one molecule is sufficient to start the nuclear export. This transport process was assumed to occur fast.

Transcription and proteins synthesis

According to the stabilization hypothesis, the progeny vRNPs in the nucleus start to replicate and transcribe viral mRNA of their respective FL genome segment $\left(R_i^M\right)$ [44]. Experimental evidence indicates that truncated versions of proteins can be generated from the DI segment [62], however they are likely to have lost their function. Therefore, and to improve computational performance, the transcription of viral mRNAs from defective segments was not considered in our model.

$$Vp_i^{Nuc} \xrightarrow{k_M^{Syn}/L_i^M} Vp_i^{Nuc} + R_i^M \tag{3.27}$$

k_M^{Syn} denotes the transcription rate. In experiments a negative correlation between the IAV mRNA and their length was observed, showing that shorter mRNAs are more abundant [46]. Hence, we established that the transcription of viral mRNA is length-dependent $\left(L_i^M\right)$. Subsequently, viral mRNA is exported to the cytoplasm. This step is not explicitly implemented in our model since mRNA export occurs fast [75].

Once the viral mRNA is in the cytoplasm, protein translation is initiated. The viral mRNAs of segment 1 – 3 encode for the three subunits (PA, PB1 and PB2) of the viral polymerase complex (P_{RdRp}).

$$R_1^M + R_2^M + R_3^M \xrightarrow{r^{RdRp}} R_1^M + R_2^M + R_3^M + P_{RdRp} \tag{3.28}$$

$$with \quad r^{RdRp} = \frac{k_P^{Syn}}{D_{Rib}} min(R_1^M, R_2^M, R_3^M)$$

During the implementation of the previous stochastic model of IAV replication, the authors identified the synthesis of each individual polymerase subunit and their assembly as the most time-consuming step during simulation [3]. Hence, to reduce the time of each simulation, the synthesis of RdRp is performed as a single-step reaction and its synthesis is limited by the least abundant of the three viral mRNAs encoding for the RdRp subunits. Since multiple ribosomes can bind to a single mRNA, the protein synthesis in the model is assumed to be proportional to the protein synthesis rate $\left(k_P^{Syn}\right)$ over the distance between two adjacent ribosomes on a viral mRNA (D_{Rib}).

The viral mRNAs of segments 4 – 6 encode for HA (P_{HA}), NP (P_{NP}) and NA (P_{NA}) proteins, respectively.

$$R_4^M \xrightarrow{k_P^{Syn}/D_{Rib}} R_4^M + P_{HA} \tag{3.29}$$

$$R_5^M \xrightarrow{k_P^{Syn}/D_{Rib}} R_5^M + P_{NP} \tag{3.30}$$

$$R_6^M \xrightarrow{k_P^{Syn}/D_{Rib}} R_6^M + P_{NA} \tag{3.31}$$

The synthesis of these proteins occurs as described above.

The M1, M2 and NEP are transcribed from spliced viral mRNAs of segment 7 and 8.

$$R_7^M \xrightarrow{k_P^{Syn}(1-F_{Spl7})/D_{Rib}} R_7^M + P_{M1} \tag{3.32}$$

$$R_7^M \xrightarrow{k_P^{Syn}F_{Spl7}/D_{Rib}} R_7^M + P_{M2} \tag{3.33}$$

$$R_8^M \xrightarrow{k_P^{Syn}F_{Spl8}/D_{Rib}} R_8^M + P_{NEP} \tag{3.34}$$

F_{Spl7} is the fraction of viral mRNA of segment 7 which encodes for M1 protein. Furthermore, F_{Spl8} represents the fraction of viral mRNA of segment 8 encoding for NEP. To reduce computational costs, the synthesis of non-structural proteins, e.g. NS1 which is derived from the spliced viral mRNA of segment 8, was neglected in our model.

Viral packaging and release

The particle assembly step in our model is based on previous experimental work which suggests a segment-specific packaging mechanism [57]. After the nuclear export, the progeny FL and DI vRNPs in the cytoplasm ($Vp_{M1,1}^{Cyt}$ and $Vp_{M1,DI}^{Cyt}$, respectively) form complexes that contain one copy of each genome segment.

$$Vp_{M1,1}^{Cyt} + Vp_{M1,2}^{Cyt} + Vp_{M1,3}^{Cyt} + \cdots + Vp_{M1,8}^{Cyt} \xrightarrow{k^{Cplx}} V_{Cplx}^{Cyt} \tag{3.35}$$

$$Vp_{M1,1}^{Cyt} + Vp_{M1,2}^{Cyt} + Vp_{M1,4}^{Cyt} + \cdots + Vp_{M1,8}^{Cyt} + Vp_{M1,DI}^{Cyt} \xrightarrow{k^{Cplx}} D_{Cplx}^{Cyt} \tag{3.36}$$

These complexes are formed with rate k^{Cplx}. The vRNP complexes can either contain the FL segment 3 $\left(V_{Cplx}^{Cyt}\right)$ or the DI segment $\left(D_{Cplx}^{Cyt}\right)$. Both segments compete to be incorporated into progeny particles.

After the assembly of the vRNP-complexes, the final step of release is the packaging of viral proteins.

$$V_{Cplx}^{Cyt} + N_{P_{HA}}P_{HA} + N_{P_{NA}}P_{NA} + N_{P_{M2}}P_{M2} + \left(N_{P_{M1}} - \sum_{i}\left[L_i^V/N_{M1}^{Nuc}\right]\right)P_{M1} \xrightarrow{r_V^{Rel}} V^{Rel} \quad (3.37)$$

$$D_{Cplx}^{Cyt} + N_{P_{HA}}P_{HA} + N_{P_{NA}}P_{NA} + N_{P_{M2}}P_{M2} + \left(N_{P_{M1}} - \sum_{k}\left[L_k^V/N_{M1}^{Nuc}\right] - \left[L_{DI}^V/N_{M1}^{Nuc}\right]\right)P_{M1} \xrightarrow{r_D^{Rel}} D^{Rel} \quad (3.38)$$

$$with \quad i = \{1,2,3 \ldots, 8\} \quad and \quad k = \{1,2,4, \ldots, 8\}$$

N_{P_j} with $j = \{HA, NA, M1, M2\}$ represents the number of each protein inside one particle. However, M1 is also included in the vRNPs. Therefore, we subtracted the amount of protein in a complete set of eight vRNPs from the total protein level inside a particle. The release of progeny STVs $\left(V^{Rel}\right)$ and DIPs $\left(D^{Rel}\right)$ occurs with rates r_V^{Rel} and r_D^{Rel}, respectively.

The release rates are determined by the product of the specific release rate k^{Rel}, the abundance of either STV V_{Cplx}^{Cyt} or DIP D_{Cplx}^{Cyt} vRNP-complexes and the protein content. The release rate was multiplied by eight to obtain a similar release rate as in the original model of IAV life cycle, where the virus assembly was determined by the total amount of vRNPs inside the cytoplasm [67], instead of a complex containing all eight genome segments.

$$r_V^{Rel} = 8k^{Rel}V_{Cplx}^{Cyt}\prod_j \frac{P_j}{P_j + K_{V^{Rel}}N_{P_j}} \quad (3.39)$$

$$r_D^{Rel} = 8k^{Rel}D_{Cplx}^{Cyt}\prod_j \frac{P_j}{P_j + K_{V^{Rel}}N_{P_j}} \quad (3.40)$$

$$with \quad j \in \{HA, NA, M1, M2\}$$

The influence of proteins during release is considered as a Michaelis-Menten-like mechanism, where $K_{V^{Rel}}$ represents the number of virus particles for which viral components must be present to reach half the maximum release rate.

Regarding the reactions described above, all molecules containing RNAs (with exception of V^{Cyt}) are affected by degradation.

$$R_i^V \xrightarrow{k_R^{Deg}} \emptyset, \quad R_i^C \xrightarrow{k_R^{Deg}} \emptyset, \quad R_i^M \xrightarrow{k_M^{Deg}} \emptyset,$$
$$R_{DI}^V \xrightarrow{k_R^{Deg}} \emptyset, \quad R_{DI}^C \xrightarrow{k_R^{Deg}} \emptyset,$$
$$R_{RdRp,i}^V \xrightarrow{k_{RRdRp}^{Deg}} \emptyset, \quad R_{RdRp,i}^C \xrightarrow{k_{RRdRp}^{Deg}} \emptyset, \quad Vp_i^{Nuc} \xrightarrow{k_{Rnp}^{Deg}} \emptyset,$$
$$R_{RdRp,DI}^V \xrightarrow{k_{RRdRp}^{Deg}} \emptyset, \quad R_{RdRp,DI}^C \xrightarrow{k_{RRdRp}^{Deg}} \emptyset, \quad Vp_{DI}^{Nuc} \xrightarrow{k_{Rnp}^{Deg}} \emptyset,$$
$$Cp_i \xrightarrow{k_{Rnp}^{Deg}} \emptyset, \quad Vp_{M1,i}^{Nuc} \xrightarrow{k_{Rnp}^{Deg}} \emptyset, \quad Vp_{M1,i}^{Cyt} \xrightarrow{k_{Rnp}^{Deg}} \emptyset,$$
$$Cp_{DI} \xrightarrow{k_{Rnp}^{Deg}} \emptyset, \quad Vp_{M1,DI}^{Nuc} \xrightarrow{k_{Rnp}^{Deg}} \emptyset, \quad Vp_{M1,DI}^{Cyt} \xrightarrow{k_{Rnp}^{Deg}} \emptyset \tag{3.41}$$

k_R^{Deg} denotes for the degradation rate of naked cRNA and vRNA. k_M^{Deg} is the degradation rate of viral mRNA. Moreover, the degradation rate of the RdRp-cRNA and RdRp-vRNA complexes is represented by k_{RRdRp}^{Deg}. Lastly, k_{Rnp}^{Deg} is the degradation rate of RNPs. According to the stabilization hypothesis, the RdRp and NP binding to the RNAs protects them from degradation. Therefore, is assumed that:

$$k_R^{Deg} > k_{RRdRp}^{Deg} > k_{Rnp}^{Deg} \tag{3.42}$$

To determine the total number of FL vRNAs of each segment $\left(R_{Tot,k}^V\right)$, we summed up the overall amount of complexes that contain vRNA in the system:

$$R_{Tot,k}^V = V_{Hi}^{Att} + V_{Lo}^{Att} + V^{En} + V^{Cyt} + R_k^V + R_{RdRp,k}^V + Vp_k^{Nuc} + Vp_{M1,k}^{Nuc} + Vp_{M1,k}^{Cyt} + V_{Cplx}^{Cyt} + D_{Cplx}^{Cyt} \tag{3.43}$$

$$with \quad k = \{1,2,4,\dots,8\}$$

Since the FL vRNA of segment 3 is not included in the DIP complex particle, the total number of vRNAs of segment 3 $\left(R_{Tot,3}^{V}\right)$ was calculated as follows:

$$R_{Tot,3}^{V} = V_{Hi}^{Att} + V_{Lo}^{Att} + V^{En} + V^{Cyt} + R_{3}^{V} + R_{RdRp,3}^{V} + Vp_{3}^{Nuc} + Vp_{M1,3}^{Nuc} + Vp_{M1,3}^{Cyt} + V_{Cplx}^{Cyt} \tag{3.44}$$

Finally, the total DI vRNA abundance in the system was calculated as described below:

$$R_{Tot,DI}^{V} = D_{Hi}^{Att} + D_{Lo}^{Att} + D^{En} + D^{Cyt} + R_{DI}^{V} + R_{RdRp,DI}^{V} + Vp_{DI}^{Nuc} + Vp_{M1,DI}^{Nuc} + Vp_{M1,DI}^{Cyt} + DI_{Cplx}^{Cyt} \tag{3.45}$$

3.2 Model variations

Experimental data suggests that the DIP possesses an advantage over STV production [63], however, is still elusive in which part of the replication process this emerges. We implemented this advantage in our model based on the previous deterministic model of DIP replication which assumes that due to their shorter length, DI cRNAs segments replicate faster than the FL cRNAs [9]. Hence, the factor advantage (F_{Adv}) is represented by the ratio between the length of segment 3 vRNA (L_{3}^{V}) and the length of the DI vRNA (L_{DI}^{V}). In our model, we used a value of 3.65 for F_{Adv} . Is important to notice that the length dependent replication was only assumed for the DI RNAs since, so far, no link has been found between the length of the FL segments and its abundance inside the cell [46,75].

$$k_{C,DI}^{Syn} = (F_{Adv} + 1) k_{C,FL}^{Syn}, \qquad with\ \ F_{Adv} = \left(\frac{L_{3}^{V}}{L_{DI}^{V}} - 1\right) = 3.65 \tag{3.46}$$

The cRNA synthesis rate of the DI segment $\left(k_{C,DI}^{Syn}\right)$ is proportional to the product of the advantage factor and the replication rate of the FL cRNA segments $\left(k_{C,FL}^{Syn}\right)$. This assumption establishes a significantly higher replication of the DI cRNA compared to its complementary FL segment (Figure 3.1). The resulting increase of DI segment levels induces a preferential packaging of DI vRNPs (Figure 3.1). As explained before, the packaging stage includes two steps: first the eight vRNPs in the cytoplasm form complexes comprising one copy of each genome segment. Then, the vRNP complexes assemble proteins and are released from the cell. The complex formation rate for STV $\left(r_{STV}^{Cplx}\right)$ and DIP $\left(r_{DIP}^{Cplx}\right)$ is proportional to the abundance of the cytoplasmic vRNP segment 3 $\left(Vp_{M1,3}^{cyt}\right)$ and DI vRNP $\left(Vp_{M1,DI}^{cyt}\right)$, respectively. Since the DI segment is present in higher levels, more DIP complexes can be formed which promotes release.

$$r_{STV}^{Cplx} \propto k^{Cplx} Vp_{M1,3}^{Cyt} \prod_k Vp_{M1,k}^{Cyt} \tag{3.47}$$

$$r_{DIP}^{Cplx} \propto k^{Cplx} Vp_{M1,DI}^{Cyt} \prod_k Vp_{M1,k}^{Cyt} \tag{3.48}$$

$$with\ \ k = \{1,2,4,\dots,8\}$$

k^{Cplx} is the complex formation constant. Furthermore, the complex formation depends also on the product of the remaining FL segments $\left(\prod Vp_{M1,k}^{Cyt}\right)$, which assumes that the complex is formed almost instantaneously.

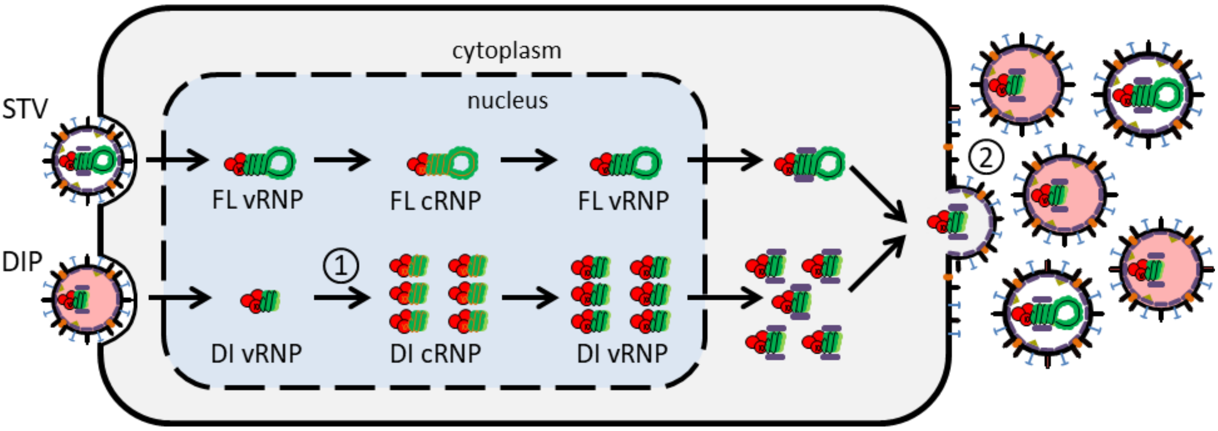


Figure 3.1| DIP advantage over STV replication. The implemented advantage factor results in an increased replication of DI cRNPs (1) and consequently higher numbers of DI vRNPs. This induces a preferential packaging of the DI segment (2) and leads to higher DIP release compared to STVs. Figure adapted from Laske *et al*. [9].

The assumptions explained below, were implemented in our model. However, in the last section of the results chapter we established changes regarding the packaging step in the model. Furthermore, we evaluated how these model variations impact the release dynamics and compared them with the original one. Firstly, we considered that the complex formation rate is considerably high and may generate unrealistic results. Consequently, we modified this parameter by assuming that its value is proportional to the least abundant vRNP in the cytoplasm $\left(min\left(Vp_{M1,k}^{Cyt}\right)\right)$. The modified model was named *Minimum Packaging Model*.

$$r_{STV}^{Cplx} \propto k^{Cplx} Vp_{M1,3}^{Cyt} min\left(Vp_{M1,k}^{Cyt}\right) \tag{3.49}$$

$$r_{DIP}^{Cplx} \propto k^{Cplx} Vp_{M1,DI}^{Cyt} \min\left(Vp_{M1,k}^{Cyt}\right) \tag{3.50}$$

$$with\ k = \{1,2,4,\dots,8\}$$

The previously implemented stochastic model of IAV replication established that the virus assembly occurs in a single reaction step that combines the eight genome segments and all structural proteins. This assumption was implemented in the previous model since the detailed molecular mechanism of virus packaging is still elusive. Hence, we were interested to test how this assumption would impact the DIP production. To that end, we disregarded the complex formation step in our model and implemented the single-step packaging reaction. Since this modified model does not include the complex formation step during packaging, but still comprises the replication advantage at the DI cRNA synthesis level, this variation of our original model was named *Replication Model.* The implemented biochemical reaction for the release of STV and DIP are represented below:

$$\begin{aligned} &Vp_{M1,1}^{Cyt} + Vp_{M1,2}^{Cyt} + Vp_{M1,3}^{Cyt} + \cdots + N_{P_{HA}} P_{HA} + N_{P_{NA}} P_{NA} + N_{P_{M2}} P_{M2} \\ &+ \left(N_{P_{RdRp}} - 8\right) P_{RdRp} + \left(N_{P_{NP}} - \sum_i [L_i^V / N_{NP}^{Nuc}]\right) P_{NP} \\ &+ \left(N_{P_{M1}} - \sum_i \left[L_i^V / N_{M1}^{Nuc}\right]\right) P_{M1} + (N_{P_{NEP}} - 8) P_{NEP} \xrightarrow{r_V^{Rel}} V^{Rel} \end{aligned} \tag{3.51}$$

$$\begin{aligned} &Vp_{M1,1}^{Cyt} + Vp_{M1,2}^{Cyt} + Vp_{M1,4}^{Cyt} + \cdots + Vp_{M1,DI}^{Cyt} + N_{P_{HA}} P_{HA} + N_{P_{NA}} P_{NA} + N_{P_{M2}} P_{M2} \\ &+ \left(N_{P_{RdRp}} - 8\right) P_{RdRp} + \left(N_{P_{NP}} - \sum_i [L_i^V / N_{NP}^{Nuc}] - [L_{DI}^V / N_{NP}^{Nuc}]\right) P_{NP} \\ &+ \left(N_{P_{M1}} - \sum_i \left[L_i^V / N_{M1}^{Nuc}\right] - [L_{DI}^V / N_{M1}^{Nuc}]\right) P_{M1} + (N_{P_{NEP}} - 8) P_{NEP} \xrightarrow{r_D^{Rel}} D^{Rel} \end{aligned} \tag{3.52}$$

In the one-step packaging model, we assumed that STV is released $\left(V^{Rel}\right)$ if the eight FL M1-vRNPs molecules for each segment $\left(Vp_{M1,i}^{Cyt}\right)$ and the all proteins required for assembly are available in the cytoplasm. The DIPs $\left(D^{Rel}\right)$ are produced when the seven FL M1-vRNPs, the DI M1-vRNP $\left(Vp_{M1,DI}^{Cyt}\right)$ and the proteins are present in the cytoplasm. N_{P_j} stands for the number of each protein inside one particle with $j = \{RdRp, HA, NA, NP, M1, M2, NEP\}$. The amount of required RdRp and NEP was determined by the total number of proteins inside a virus particle minus the proteins that are already included in the vRNPs. In a similar way, the calculation for NP and M1 was accounting for the number of each protein in a complete set of vRNPs depending on the length of each segment.

The release rate of STVs $\left(r_V^{Rel}\right)$ is proportional to the amount of FL vRNP-M1 of segment 3 in the cytoplasm $\left(Vp_{M1,3}^{Cyt}\right)$. By contrast, the release rate of DIPs $\left(r_D^{Rel}\right)$ depends on the abundance of the DI vRNP-M1 molecule $\left(Vp_{M1,DI}^{Cyt}\right)$. Furthermore, this model variation considers that all the viral proteins are necessary for particle release including free polymerases, NP and NEP which are not bound to vRNPs. This assumption considers the total number of each protein required per particle $\left(N_{P_j}\right)$.

$$r_V^{Rel} = k^{Rel} Vp_{M1,3}^{Cyt} \prod_j \frac{P_j}{P_j + K_{V^{Rel}} N_{P_j}} \tag{3.53}$$

$$r_D^{Rel} = k^{Rel} Vp_{M1,DI}^{Cyt} \prod_j \frac{P_j}{P_j + K_{V^{Rel}} N_{P_j}} \tag{3.54}$$

$$with \quad j \in \{RdRp, HA, NA, NP, M1, M2, NEP\}$$

3.3 Computation and simulation assumptions

Our stochastic model of DIP replication followed the previously implemented stochastic model of the IAV life cycle developed by Heldt & Dorl [3,67]. For model simulation, they used the SSA described in the original work of Gillespie [70]. Furthermore, to improve computational performance, an approximation method for the SSA, i.e. the Tau-leaping procedure, was applied. This approach was developed by Cao *et al.* and enables to calculate a high number of individual reactions in a single computational step [73].

In the model, the propensity function $a_j(x)\ dt$, i.e. the probability that reaction R_j will occur in the next infinitesimal time step, was assumed to be equivalent to the reaction rates. Furthermore, the state-change matrix ν, where ν_{ij} represents the change in the molecular population of S_i induced by a single R_j reaction, is equivalent to the stoichiometric coefficients. However, in the reactions of NP and M1 binding to the RNA segments, these values are calculated by dividing the length of the respective segment over the number of nucleotides bound by one NP or M1 molecule $\left(L_i^V / N_{NP}^{Nuc}\right.$ or $\left.L_i^V / N_{M1}^{Nuc}\right)$. Since stochastic models considers the state variables as discrete, this result was rounded down to the next smaller integer. An example for the establishment of the propensity function and state change matrix for the DIP binding and detachment reaction can be observe on Table 3.1.

The model was implemented in the integrated development environment Visual Studio 2017 (version 15.0 Community edition) developed by Microsoft using C++ coding language. All the simulations were performed on a Linux-based system. The model result files were analysed with MATLAB software (version 9.2.0.538062 R2017a). All the parameters implemented in the model can be consulted in the appendix (Table A.1.1 and additional parameters for the *Replication Model* in Table A.1.2).

Each stochastic simulation can be considered as an individual cell. We performed and analysed simulations for different initial infection conditions, which vary in the number of initial extracellular STVs and DIPs. We established that the representation of these conditions would follow as MOI/MODIP. In that notation, MOI stands for multiplicity of infection and is determined by the number of infectious STVs

$\left(V^{Ex}(t=0)\right)$, and MODIP is the multiplicity of DIPs and comprises the number of extracellular DIPs $\left(D^{Ex}(t=0)\right)$.

Table 3.1 | Example of the implementation of DIP binding and detachment reactions in the stochastic model. The propensity functions correspond to the reaction rates. The state change matrix is constructed considering the stoichiometric coefficients for each species in every reaction.

Reaction: DIP binding and detachment from high and low affinity sites						
$D^{Ex} + B_n \underset{k_n^{Dis}}{\overset{k_n^{Att}}{\rightleftarrows}} D_n^{Att}, \qquad n \in \{\, Hi, Lo \,\}$						
Propensity Function		**State Change Matrix, $\boldsymbol{\nu}$**				
j	$a_j(x)$	D^{Ex}	B_{Hi}	B_{Lo}	D_{Hi}^{Att}	D_{Lo}^{Att}
r_{Hi}^{Att}	$k_{Hi}^{Att} D^{Ex} B_{Hi}$	-1	-1	-	+1	-
r_{Lo}^{Att}	$k_{Hi}^{Att} D^{Ex} B_{Lo}$	-1	-	-1	-	+1
r_{Hi}^{Dis}	$k_{Hi}^{Dis} D_{Hi}^{Att}$	+1	+1	-	-1	-
r_{Lo}^{Dis}	$k_{Lo}^{Dis} D_{Lo}^{Att}$	+1	-	+1	-	-1

The property which determines the validity of the stochastic simulation datasets is the number of individual simulations performed for each initial infection condition. This is particularly important to achieve representative averages of the cell populations. To evaluate how many simulations are needed to obtain an accurate result we analysed how the average value of DIP release changes with an increasing number of simulations (Figure 3.2). We observe that after 1000 simulations the average DIP production changes only by 2% and we achieve a stable average. Therefore, we regarded 1000 simulations as sufficient to analyse our stochastic data for each scenario.

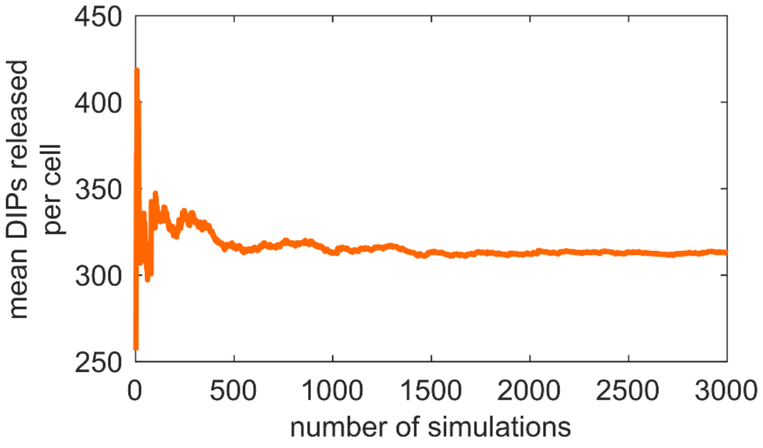


Figure 3.2 | Average converges with increased number of simulations. The average DIP released was calculated while increasing the number of simulations at MOI/MODIP 10/10. After 1000 simulations performed, the mean value changes only by 2% and it achieves a stable average.

Simulations in which genome segment 7 (encoding M1 and M2) is lost or the level of M1 protein is very low lead to the prevention of vRNP export from the nucleus. This results in an exponential increase in the RNA segment levels and high computational simulation times. Hence, we implemented a criterion that leads to the abortion of such simulations to significantly reduce computational effort. To that end, we established that if the number of reactions, $R > 4.8\text{x}10^{10}$, then the simulation would be stopped, and the results discarded. Since in some scenarios this criterion is not sufficient to prevent runs with an exorbitant accumulation of RNA, we implemented an additional criterion. When a total of $1\text{x}10^{8}$ reactions occurred in the simulation, the algorithm calculates the average number of vRNPs inside the nucleus $\left(\frac{\sum Vp_i^{Nuc}}{8}\right)$. Next, after $1\text{x}10^{9}$ reactions were performed, the same average is determined again. In a simulation without vRNA accumulation the export of vRNPs from the nucleus would have already been initiated at this point, which would lead to a decreased value. However, if the average number of vRNPs inside the nucleus more than doubled between these two steps, we stopped the simulation, because this suggests that vRNA is still accumulating in the nucleus. The simulation runs that meet one of the two abortion criteria are discarded from all further analysis.

3.4 Data analysis tools

When we apply a set of initial infection conditions, the deterministic model generates only one possible outcome. By contrast, the stochastic model considers that randomness is involved in the biochemical reactions. Therefore, in each simulation a different sequence of reactions occurs producing different results for the same initial conditions. One of our main goals is to compare our stochastic model of DIP replication with the previously developed deterministic approach.

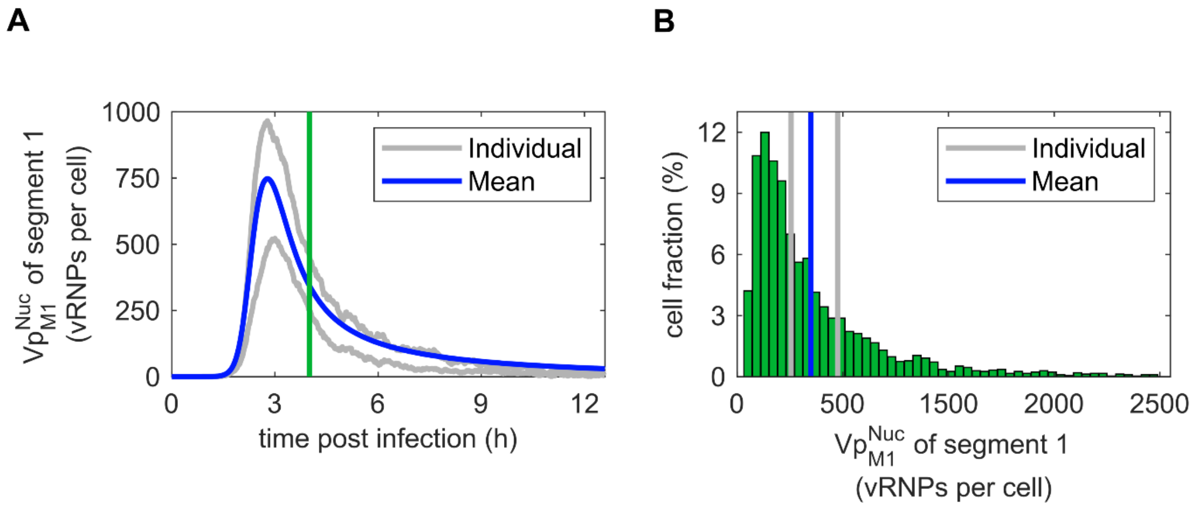


Figure 3.3 | Averaged stochastic dynamics and distribution. (A) Dynamics of Vp_{M1}^{Nuc} of segment 1 until 12 hpi at MOI/MODIP 10/10. Two individual stochastic simulations are represented. The average is calculated for each time point over 3060 performed simulations. (B) Distribution of the same molecule at 4 hpi (represented by the green line in A). The lines are the individual and average results showed in A at the same time point.

To illustrate the molecule dynamics of the two approaches, we calculated the average of all stochastic simulations for each time point (Figure 3.3A) and compared it with the deterministic result. For each time-point (green line in Figure 3.3A corresponds to 4 hours post infection (hpi)), the individual stochastic results will deviate from each other which results in a distribution that describes the variance introduced by the intrinsic stochasticity of the system (Figure 3.3B).

4 Results

4.1 DIP interference and comparison with the deterministic model

4.1.1 High MOI and MODIP

To evaluate how the randomness of events affects the IAV replication in the presence of DIPs we compared the results of the stochastic model that we developed with the deterministic model previously implemented. We performed a total of 3060 simulations for the initial infection condition of MOI/MODIP 10/10. Our first main goal was to analyse the heterogeneity of DIP replication, so we started by observing the distributions of STV and DIP production. At MOI/MODIP 10/10 our simulations reveal that STV have a distribution towards low productivity with 60% of simulations being non-productive (Figure 4.1A) and 39% producing up to 15 virions (Figure 4.1B). Regarding DIP production, only 0.5% of the simulations are non-productive and results show a wider distribution (Figure 4.1C).

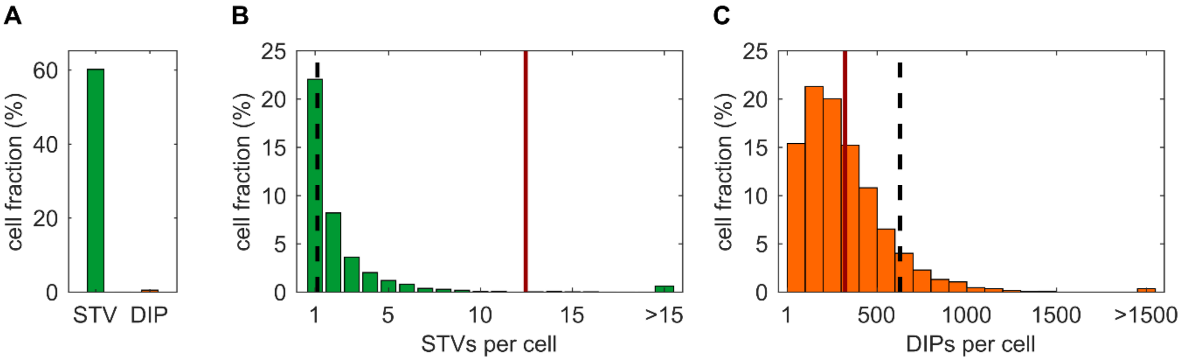


Figure 4.1 | Distribution of standard virus (STV) and DIP production at MOI/MODIP 10/10. (A) Percentage of non-productive STV and DIP cells. (B,C) Distribution of total number of released STVs (B) and DIPs (C) until 12 hpi in stochastic simulations. The lines correspond to the mean value for the stochastic model (solid line) and the deterministic result (dashed line).

The average virion release is 12 ± 295 STVs per cell. Although this result indicates that STV release has a wide-spread distribution, upon closer examination we identified six simulations with a range between 1076 and 10423 STV per cell that substantially increase the average and standard deviation. The simulations that result in high STV release values have one event in common, which is the failure of DIP fusion. This allows the STV to replicate without interference. When we exclude these outliers, we obtain an average of 5 ± 5 STVs per cell which is closer to the deterministic result.

Both DIP and STV release distributions show that the data is shifted away from the right side of the histogram, so we can assume that the distributions tend to be skewed to the right. The DIP production results (Figure 4.1C) closely follow a log-normal distribution with an average of 312 ± 295 DIPs per cell which is lower than the deterministic value. In 99% of the simulations the amount of produced DIPs varies between 1 and 1500 particles indicating that DIP release is highly susceptible to fluctuate due to stochastic effects.

Overall when we compare the average molecule dynamics in the stochastic simulations with the deterministic results we observe that the differences are more significant for STV release. As explained before, due to simulations that allow the STV to replicate without interference, virions production differ with a factor of 10 from the deterministic solution (Figure 4.2A). DIP production differs with only with a factor of 2 (Figure 4.2B). In agreement with the deterministic model, the implementation of a DIP replication advantage causes the DI cRNA to replicate faster than the corresponding FL segment. Consequently, levels of DI vRNA are almost three orders of magnitude higher than its FL counterpart (Figure 4.2C). Furthermore, in our model vRNA and cRNA levels (Figure 4.2D) are very similar with the deterministic result that differ with factor ranges between 1.2 and 1.5. Although the intermediate

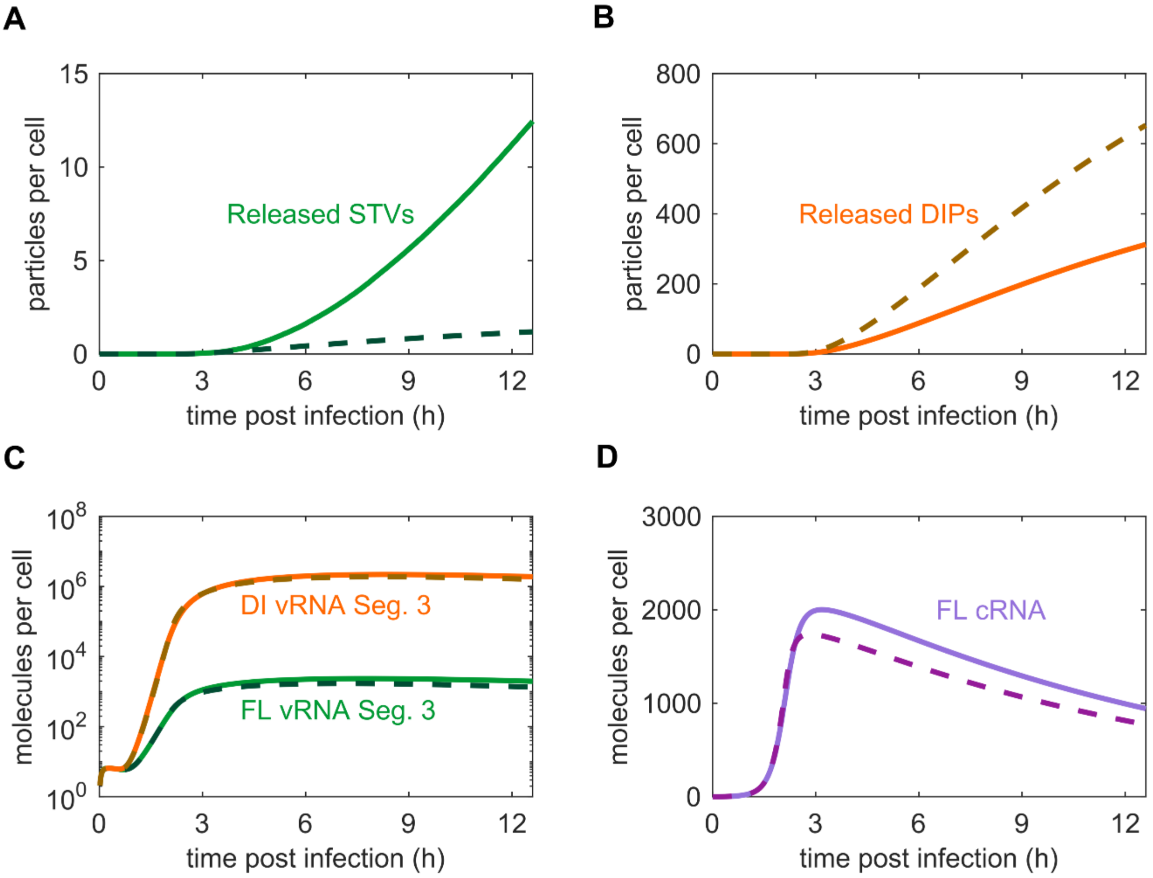


Figure 4.2 | Comparison of the stochastic and deterministic simulation. (A,B) Total cumulative released STV (A) and DIP (B). (C) Level of total DI and FL vRNA of segment 3. (D) Dynamics of total FL cRNA. The lines represent the mean stochastic results of all simulations performed at MOI/MODIP 10/10 (solid line) and the deterministic result (dashed line).

molecules in our model show very similar dynamics with the deterministic solution, this difference became more pronounced during release.

Our model considers all genome segments individually: cRNA and vRNA of the eight FL segments and the DI segment replicate independently, which means that each segment level inside a cell can vary significantly. Consequently, the absence or low level for one segment necessary for replication, can impair release. By contrast, in the deterministic model a complete set of vRNP reaches the nucleus and the segments replicate simultaneously maintaining similar vRNA levels during the infection. Since in the stochastic model the abundance of individual vRNA segments can be substantially different, we questioned how the DI segment correlates with other FL segments. The factors that affect vRNA abundance can be summarized in two major categories [1]: (I) extrinsic noise such as cell stage, cell size or protein content, which affects all individual vRNA levels evenly and increases the vRNA intersegment correlation and (II) intrinsic noise, as for example, the independent segment synthesis, is related with the randomness of biochemical reactions and affects the vRNA levels differently which will lead to a decrease in their correlation. To elucidate how this correlation affects DIP production, we split

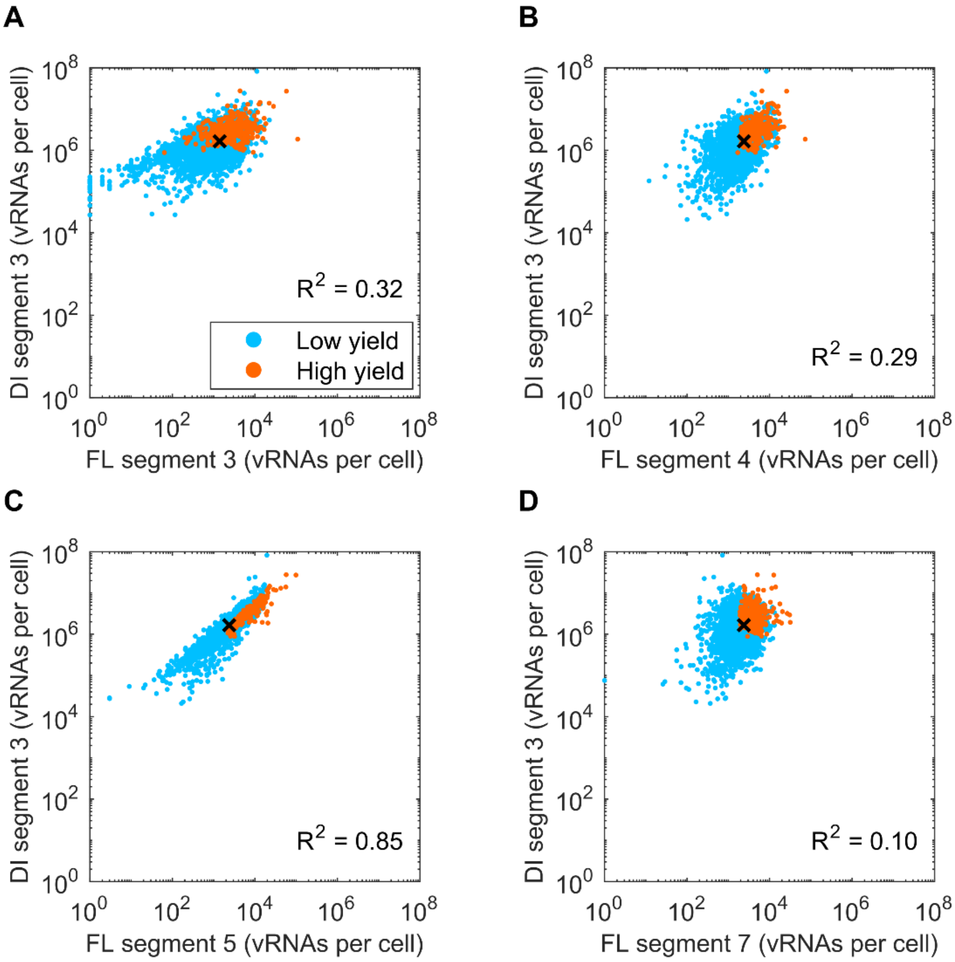


Figure 4.3 | Correlation of the defective interfering vRNA and various vRNA segments and DIP production. DI segment 3 correlated with its corresponding FL segment (A), FL segment 4 (B), FL segment 5 (C) and FL segment 7 (D). Simulations performed at MOI/MODIP 10/10. The represented vRNA levels are at 12 hpi and the black X indicate the deterministic solution. Only cells that produce DIPs are represented. High productive cells were defined as the cells that are in the top 10% of all productive cells.

the cells into high and low DIP yield: the high DIP productive cells were defined as the cells that are in the top 10% of the all productive cells.

Although high DIP yield cells contain elevated levels of DI and FL vRNA, we can observe that low levels of FL segment 3 support high DIP production since DIP do not incorporate this vRNA segment in its genome (Figure 4.3A). Most of the FL segments have a weak correlation with the DI segment and show lower limit of 10^3 vRNAs per cell for high yield cells (Figure 4.3B,D) which reveals that intrinsic factors influence DIP productivity. By contrast, DI segment has a strong correlation with the FL segment 5 (Figure 4.3C), that encodes for NP protein which is essential for packing the free vRNA into vRNP. High yield cells have a narrow distribution indicating that DIP production is also affected by extrinsic factors.

Since FL segment 5 correlates strongly with the DI segment 3 we wondered if the same dependency was observed when we correlate FL segment 5 with other FL segments and if it affects STV production as well. We divided the STV productive cells in high and low yield: the STV high

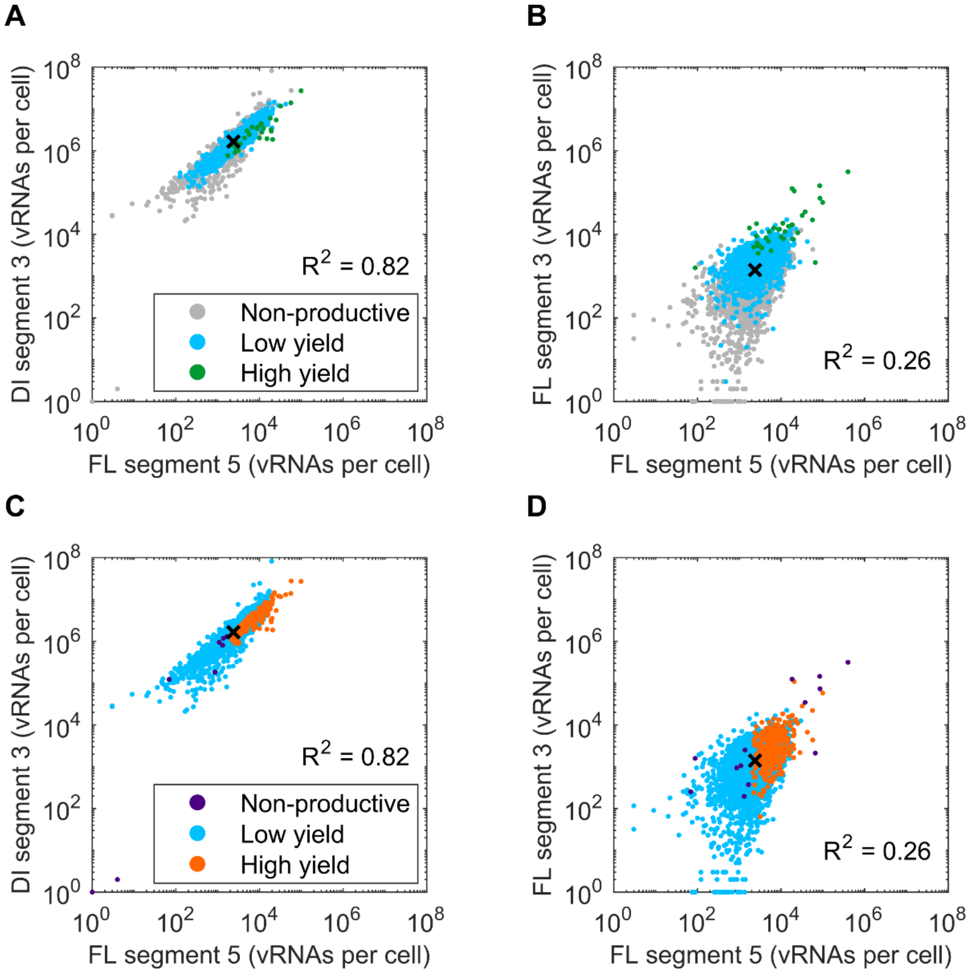


Figure 4.4 | Correlation of vRNA segment 5 with defective and functional segment 3 and its influence in STV and DIP production. STV production and correlation of FL segment 5 with DI segment 3 (A) and FL segment 3 (B). DIP production and correlation of FL segment 5 with DI segment 3 (C) and FL segment 3 (D). Simulations performed at MOI/MODIP 10/10. The represented vRNA levels are at 12 hpi and the black X indicate the deterministic solution. High STV yield and high DIP yield cells were defined as the cells that are in the top 3% and top 10% of all productive cells, respectively.

productive cells were defined as the cells that are the top 3% of the all productive cells that result in a range between 9 and 10423 released STV. As known before, overall the simulations show that there are more cells that do not release STV than cells not producing DIPs and that the DI segment 3 level is much higher than its complementary FL segment due to the DIP replication advantage. We can observe that the correlation coefficient between FL segment 5 and DI segment 3 is higher than with the FL segment 3 (Figure 4.4A,B). This weak correlation with FL segment 3 shows that intrinsic factors have a role in STV release (Figure 4.4B). When the level of FL segment 5 is lower than 10^2 vRNAs per cell STV do not produce but in the same scenario DIP can still produce low amounts of progeny particles (Figure 4.4C). Furthermore, the lower limit at 10^3 vRNAs per cell for high DIP yield cells shows that DIP production can be determinate by the level of FL segment 5 (Figure 4.4D).

Another limiting factor for DIP replication should be the availability of resources since its genome lacks information to encode one protein that is essential for its replication, requiring a coinfection with a STV that will provide the missing protein. We analysed the total viral mRNA dynamics which is responsible for protein production and the two proteins that we consider to be the most important for replication: the viral polymerase complex (RdRp) as the implemented DIP has a defective segment 3 whose complementary FL encodes for one of the three subunits of the polymerase, and NP because as shown previously the level of vRNA segment 5 impacts DIP production.

Our stochastic model captures the dynamics of total viral mRNA levels although it starts do differ from the deterministic simulation with a factor of 1.5 after 2 hpi (Figure 4.5A). As our model reveals a 1.5 factor difference as well for vRNA dynamics, this deviation is propagated to the mRNA levels due to transcription. The stochastic result captures the RdRp and NP level of the deterministic model simulation during early infection but after approximately 2 hpi both proteins begin to differ (Figure 4.5B). Between 2 and 7 hpi RdRp shows a slightly lower level than deterministic model because we assumed that the polymerase complex formation rate is proportional to the least abundant of the three mRNA that encode for RdRp subunits (segment 1 – 3). In the majority of simulations segment 3 is the last abundant segment since the DIP genome contributes to the abundance of all FL segments with

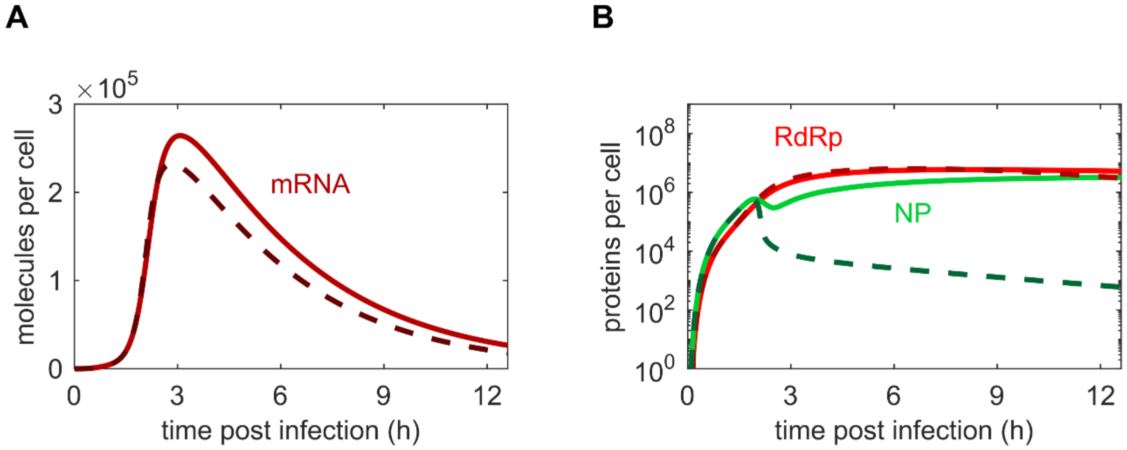


Figure 4.5 | Comparison of viral mRNA and protein dynamics for stochastic and deterministic model simulations at MOI/MODIP 10/10. (A) Total viral mRNA of all eight segments. (B) Total unbound RdRp (red line) and NP (green line). The lines represent the mean stochastic results of all simulations performed at MOI/MODIP 10/10 (solid line) and the deterministic result (dashed line).

exception of segment 3. Later in infection we can observe that our model has increased levels of unbound RdRp that differ with a factor of 1.8 from the deterministic simulation.

NP represents by far the highest deviation between the stochastic and the deterministic model regarding all molecules in the system at 12 hpi. In addition, since all intermediates begin to slightly differ after NP begins to significantly diverge, we suspect that NP is causing other molecules dynamics differences. This raised the question about what causes this large discrepancy and how the NP level is distributed.

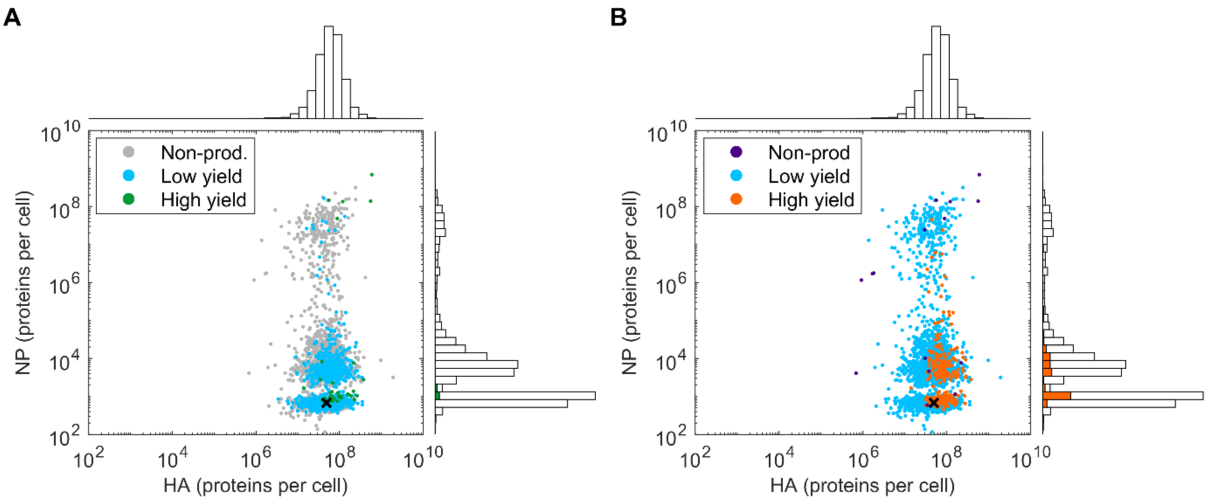


Figure 4.6 | Levels and distributions of HA and NP related to STV and DIP yield. (A) Protein distributions and STV production. (B) Protein distributions and DIP production. Results at 12 hpi with initial conditions MOI/MODIP 10/10. The black X indicates the deterministic solution.

When we compare the histograms of HA and NP we observe that HA has a log-normal distribution, but NP distribution reveals to be much more spread out. The levels of NP range over six orders of magnitude and three subpopulations with distinct characteristics emerge (Figure 4.6A,B). The NP distribution shows its highest peak when levels at 12 hpi are low and in this range, NP has a narrow distribution. We observe that DIP will strongly replicate and STV also displays a small cluster of high productive cells. In this area the low level of NP indicates that the increased FL and DI vRNP replication consumes almost all free NP available. The second NP subpopulation displays a more spread distribution with range between 10^3 and 10^6 proteins per cell. In this region cells continue to show high DIP yields while STV production is lowered. Lastly, when NP levels are above 10^6 proteins per cell, it shows an approximately uniform distribution and the STV mainly do not produce any progeny virions although DIP are still produced in low amounts. It is also important to notice that when levels of NP are high we can find few simulations that result in STV replication without DIP interference.

Although the peak of NP distribution occurs at low protein levels, which result in higher yields for both STV and DIP in agreement with the deterministic model, most of our simulations show higher NP levels. The simulations that show very high levels of NP may occur due to fusion failure events or the loss of at least one genome segment. Such imbalance in replication, which is not accounted for in the deterministic model, would lead to an accumulation of NP, because less vRNA would need to be stabilized by the protein. Furthermore, we observe that there are more cells that produce high amounts of DIPs when at least low amounts of STV is released. The cause for that could be that STV release indicates a balanced replication of all segments which increases the availability of resources for DIP release. However, DIPs are also released when the cells do not produce any STV, but their yield is reduced.

4.1.2 Low MOI and MODIP

So far, the stochastic simulation results are in good agreement with the dynamics described by the deterministic model considering high MOI and MODIP conditions. Most properties, e.g. the average amount of vRNA molecules, capture deterministic averages well despite large distributions. However, some molecules, i.e. NP and released STV particles, show significant differences that could be used as starting points for further analysis. In this section we want to evaluate how the stochastic effects impact DIP production for low initial infection conditions, so we performed a total of 6352 simulations for MOI/MODIP 1/1. We were able to perform a higher number of simulations for this scenario due to the reduced amount of infecting virus particles. This decreased the number of molecule interactions significantly leading to an overall reduction in the necessary computation time.

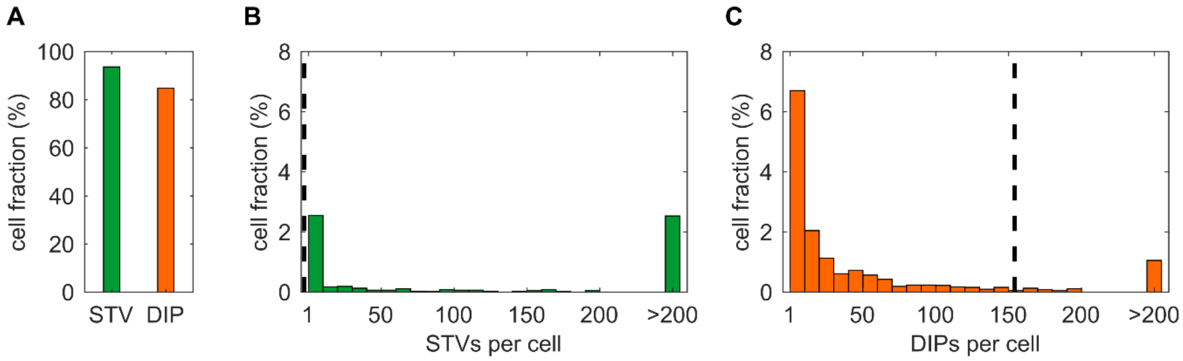


Figure 4.7 | Distribution of STV and DIP production at MOI/MODIP 1/1. (A) Percentage of non-productive STV and DIP cells. (B,C) Distribution of total number of released STVs (B) and DIPs (C) until 12 hpi in stochastic simulations. The lines correspond to the mean value for the stochastic model (solid line) and the deterministic result (dashed line).

In low MOI and MODIP conditions our simulations reveal that STV and DIP show a trend towards low productivity with 94% and 85% of the cells being non-productive, respectively (Figure 4.7A). Both releases have a wide-spread distribution which indicates that the production of virus particles is highly variable in this scenario. The STV production distribution shows a wider spread when compared with the high MOI and MODIP result with a maximum release of 15401 STV per cell. The average virus release is 75 ± 693 STV per cell which is considerably higher than the average deterministic result. As explained before the high average and standard deviation values are caused by the 2.5% of simulations that produce more than 200 STV per cell (Figure 4.7B). The average DIP release amounts to 8 ± 59

DIP per cell which is very low compared with the deterministic simulation that on average produces 171 DIP per cell. The high percentage of cells that do not produce DIPs in low infection conditions shifts the average release to a very low value. In contrast to STV production, the maximum number of DIP release is just 3350 DIPs per cell and only 1% of simulations have a release higher than 200 DIPs per cell (Figure 4.7C).

Overall the differences between the stochastic and deterministic model are more pronounced at low MOI/MODIP conditions. Release of both STVs and DIPs shows strong deviations from the deterministic results at 12 hpi (Figure 4.8A,B). Furthermore, the higher differences are also observed for FL vRNA segment 3 and total FL cRNA (Figure 4.8C,D) which will influence the STV production and increase its deviation from the deterministic simulation. This higher discrepancy between the stochastic and deterministic model is caused by the imbalanced replication which results in an elevated percentage of DIP and STV non-productive cells. These non-productive simulations emerge mainly due to stochastic factors, e.g. failure of virus fusion, which occurs when either STV or DIP fail to enter the cell, and genome segment loss which means that vRNA is degraded in the nucleus. In general, when DIPs fail to fuse, the STV can replicate without interference and potentially release high numbers of virions. In contrast, when the STV fails to fuse, the DIP cannot replicate by itself, which greatly reduced

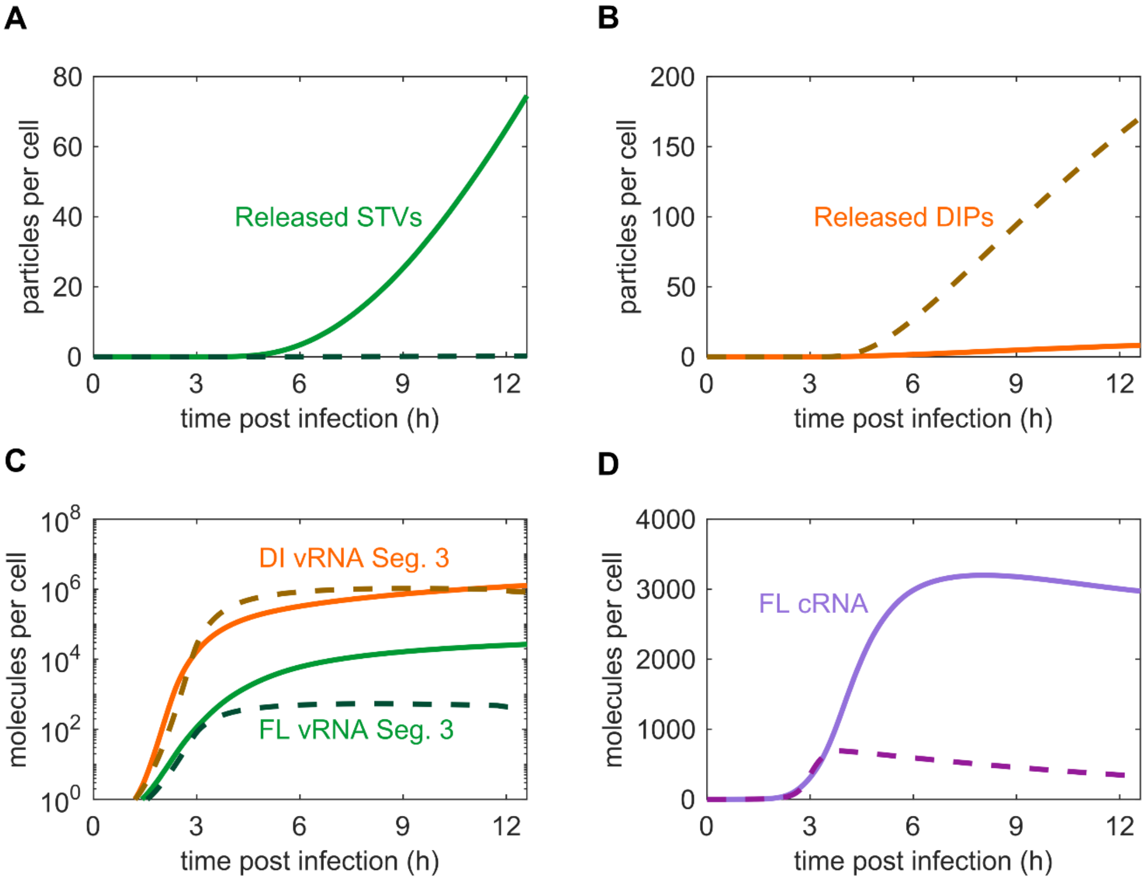


Figure 4.8 | Higher differences between the stochastic and deterministic simulation at MOI/MODIP 1/1. (A,B) Total cumulative released STV (A) and DIPs (B). (C) Levels of total DI and FL vRNA of segment 3. (D) Dynamics of total FL cRNA. The lines represent the mean stochastic results of all simulations (solid line) and the deterministic result (dashed line).

the average DIP release in the low MOI/MODIP infection scenario. This topic will be discussed in more detail in the next chapter.

Once again, we decided to analyse the correlation of DI and FL segment 3 with FL segment 5 to investigate how these impact STV and DIP replication. Our hypothesis that cells showing high STV yields are generated in the absence of DI genome segments was confirmed by the stochastic simulation results. We observed that all simulations with high STV yields had a DI segment 3 level equal to zero (due to logarithmic scale these data cannot be observed in Figure 4.9A). Furthermore, the correlation coefficients between the segments are significantly reduced compared with the high MOI and MODIP scenario. This difference can be explained by the increased influence of stochastic effects that highlight the importance of intrinsic noise at low initial conditions. Besides, FL segment 3 distribution has the most disperse data comprising almost eight orders of magnitude (Figure 4.9B). As observed before, high DIP yield cells show a strong correlation between DI segment 3 and FL segment 5 which indicates that extrinsic factors also affect DIP replication (Figure 4.9C) and DIPs can still strongly replicate when FL segment 3 levels are low (Figure 4.9D).

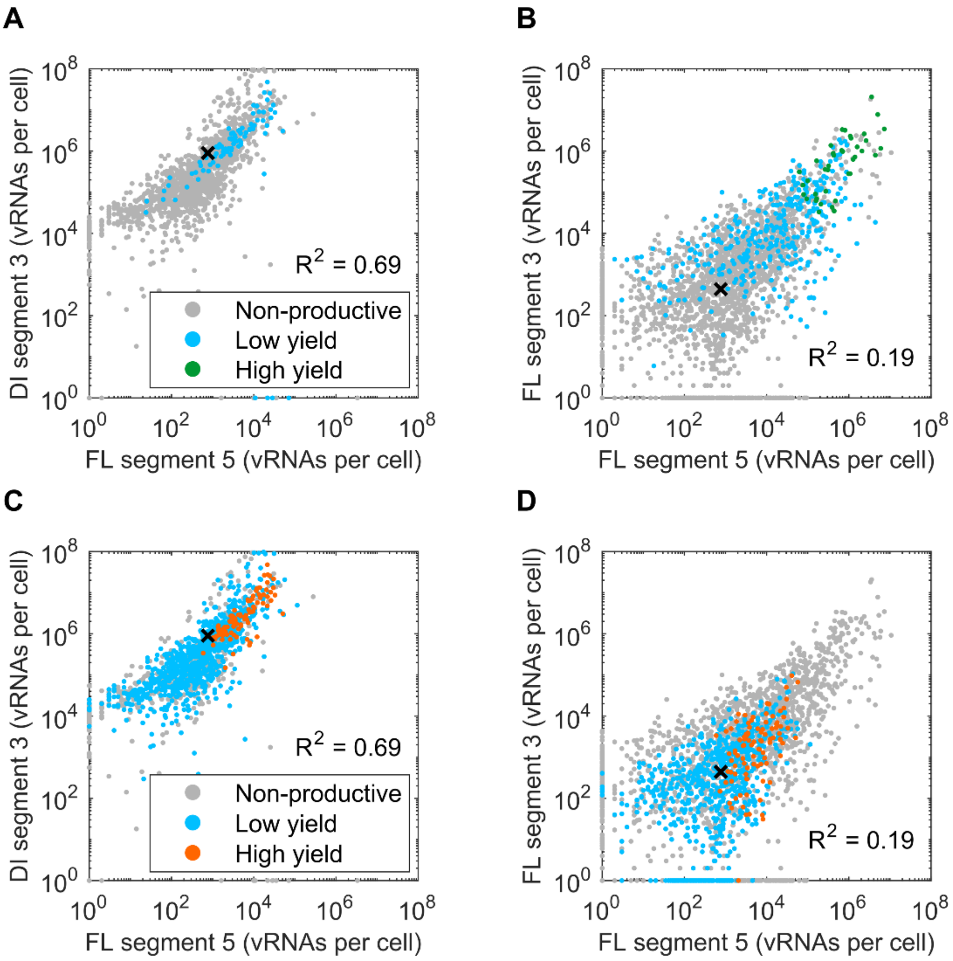


Figure 4.9 | Correlation of vRNA segment 5 with defective and functional segment 3 and its influence in STV and DIP production. STV production and correlation of FL segment 5 with DI segment 3 (A) and FL segment 3 (B). DIP production and correlation of FL segment 5 with DI segment 3 (C) and FL segment 3 (D). Simulations performed at MOI/MODIP 1/1. The represented vRNA levels are at 12 hpi and the black X indicate the deterministic solution. High STV yield and high DIP yield cells were defined as the cells that are in the top 10% of all productive cells.

In the deterministic model enveloped particles can be degraded or fuse with the membrane, both with roughly the same rate. This leads to only ~50% of the initially provided particle to reach the nucleus. The same dynamics was also implemented in the stochastic model, although, as the the deterministic model approaches the problem as a continuous system, the results are very different. In the deterministic simulation of a MOI/MODIP 1/1 scenario, half of each particle will successfully fuse and induce progeny particle production. It is obvious that this does not happen in real biological systems, as one complete viral particle is necessary to infect a single cell. In this context, the stochastic model is more realistic by supporting four possible fusion scenarios: STV fuses alone, DIP fuses alone, both fuse or both fail fusing. In the scenario that STV fuses alone, viral replication occurs without interference leading to higher amounts of viral cRNA and viral mRNA (Figure 4.8D and Figure 4.10A).

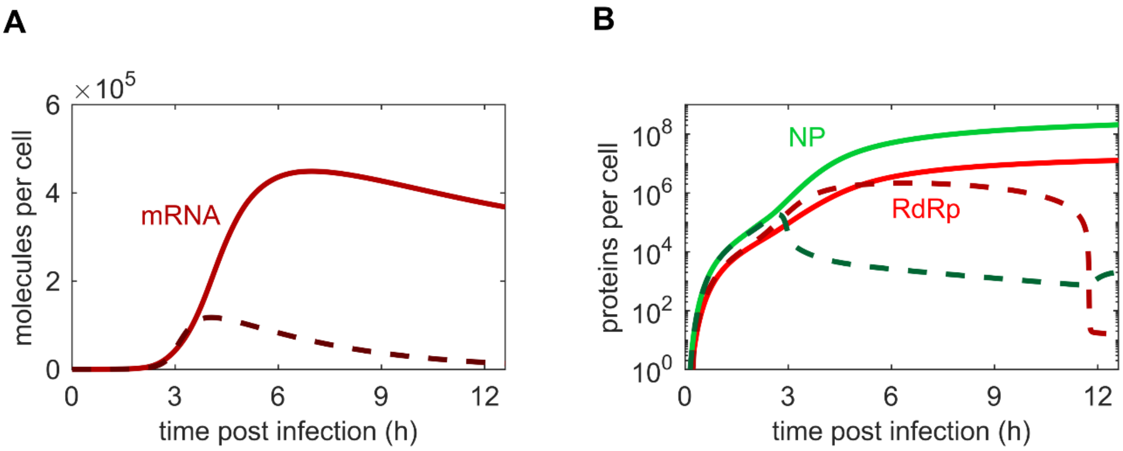


Figure 4.10 | Comparison of viral mRNA and protein dynamics for stochastic and deterministic model simulations at MOI/MODIP 1/1. (A) Total viral mRNA of all eight segments. (B) Total unbound RdRp (red line) and NP (green line). The lines represent the mean stochastic results of all simulations performed at MOI/MODIP 1/1 (solid line) and the deterministic result (dashed line).

Regarding protein levels, for MOI/MODIP 10/10 our model captured the RdRp dynamics of the deterministic model, although NP was the molecule that showed the highest difference in our simulations. Up to 2 hpi, in a low initial infection condition scenario, our protein dynamics have very similar levels with the deterministic result. Then, beginning at approximately 2 hpi, protein levels start to deviate (Figure 4.10B). Although the overall difference between the stochastic and deterministic model during the progression of infection is higher for NP, at 12 hpi RdRp deviation is more pronounced.

To understand the reason why our average RdRp and NP dynamics are so different from the deterministic simulation we observed the distributions of both proteins. Our simulations reveal that 60% and 24% of simulations do not produce any RdRp and NP, respectively (Figure 4.11A). The higher percentage of cells that do not produce RdRp is caused by the fusion failures of STV which prevent the production of the polymerase complex because the DIP has a defective segment 3. NP can be produced if either STV or DIP or both enter the cell since the implemented DIP has a functional segment 5 that enables to synthesise NP protein in the absence of STV coinfection.

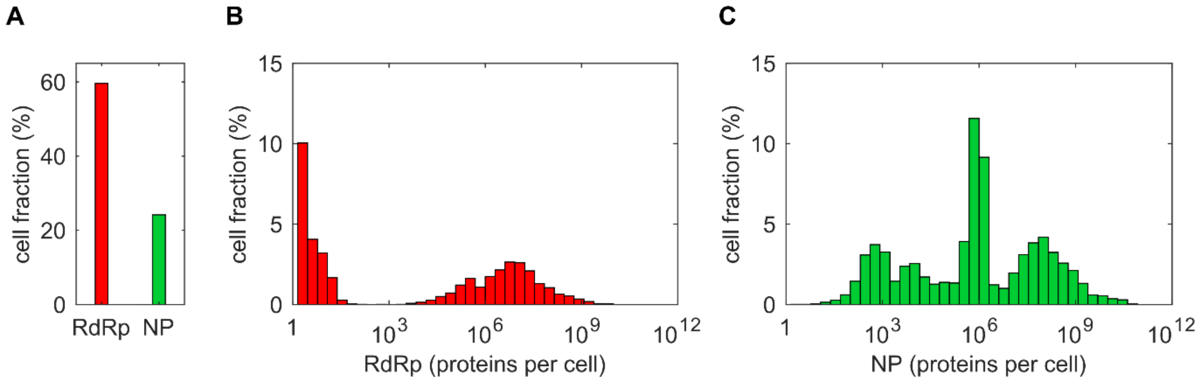


Figure 4.11 | Distributions of RdRp and NP in a MOI/MODIP 1/1 scenario. (A) Percentage of cells that do not produce RdRp and NP. (B,C) Distribution of unbound RdRp (B) and NP(C) at 12 hpi.

RdRp distribution is spread over ten orders of magnitude and consists of two subpopulations (Figure 4.11B): the first cluster has a narrow distribution between 1 and 10^2 proteins per cell and shows a trend towards lower protein levels. The second subpopulation has a wider distribution with lower and upper limits of 10^3 and 10^{10} proteins per cell.

When we analyse the NP levels at 12 hpi we observe a wide-spread distribution that comprises eleven orders of magnitude and three distinct subpopulations emerge (Figure 4.11C). The cell subpopulation located between 10^5 and 10^7 proteins per cell has a narrow distribution and the overall NP peak is in the centre of this distribution at approximately 10^6 proteins per cell. The more dispersed subpopulations are observed at ranges between 1-10^5 and 10^7-10^{11} proteins per cell. Although, as we explained in the previous section, in cells showing high DIP yields and low levels of NP at 12 hpi were observed because vRNA uses most of the protein to form vRNP. In summary, the NP protein has an extra layer of variability and its disperse distribution can be caused either by segment 5 loss or by increased vRNP replication.

The different fusion failure and segment loss events increase the heterogeneity in protein levels and consequently affect the STV and DIP production. To analyse how the different RdRp and NP levels impact progeny release we combined both proteins levels and virus particles production in the same graph (Figure 4.12).

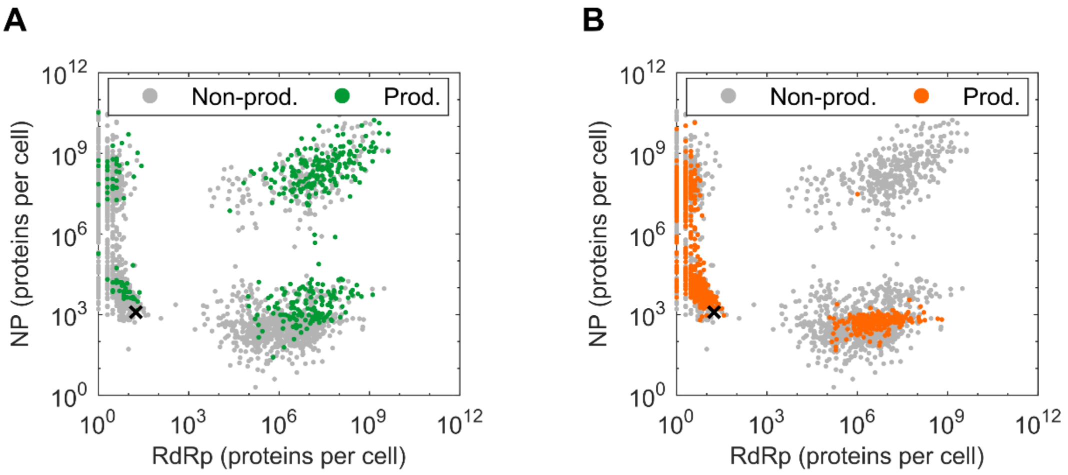


Figure 4.12 | Levels of RdRp and NP related to STV and DIP yield. (A) Protein levels and STV production. (B) Protein levels and DIP production. Results simulated at 12 hpi with initial conditions of MOI/MODIP 1/1. The black X indicates the deterministic solution.

We observe that the deterministic result is in an area where RdRp and NP levels are low, which indicates that this region corresponds to the scenario in which both DIP and STV successfully fused. When RdRp level is low at 12 hpi DIP and STV release progeny, although we observe that DIP has more productive cells in this region. For higher RdRp levels ranged between 10^{3} and 10^{10} proteins per cell we observe two distinct subpopulations: when NP levels are higher than 10^{6} proteins per cell STV replicates without interference due to DIPs that fail to enter the nucleus. In contrast, when NP levels are low we observe two scenarios: if NP is lower than 10^{3} proteins per cell DIP releases progeny particles, by contrast if NP level is higher than 10^{3} STV replicates alone. In between we have a common region approximately at 10^{3} where both DIP and STV replicate. Interestingly, we showed before in Figure 4.11C that the peak of NP distribution as around 10^{6} proteins per cell, although Figure 4.12 only shows few simulations located in that region (due to logarithmic scale). These simulations correspond to scenarios in which STV fails to fuse and this prevents RdRp production. Due to the lack of viral polymerases neither STVs nor DIPs produce progeny in this scenario.

4.1.3 Impact of MOI and MODIP

After we analysed in detail how the stochastic effects impact the STV and DIP replication for high and low MOI and MODIP we investigated how different combinations of initial condition would affect DIP production. To that end, we performed a total of 19180 simulations for MOI and MODIP 1, 2, 5 and 10. The number of simulations performed for each combination of initial conditions can be consulted in the appendix (Table A.2.1). Initially, we fixed the MODIP at 1 and increased the MOI. Then, we fixed the MOI at 1 and subsequently increased the MODIP. We compared the average DIP production from our stochastic simulations with the deterministic result.

We observe higher DIP releases with increasing MOI for both models. Furthermore, the differences between the average stochastic simulations and the deterministic solution regarding DIP production become less pronounced when MOI is higher (Figure 4.13A). As shown before, for low initial infection conditions stochastic effects are more pronounced and will affect the outcome of the system. In addition, our data reveal that higher MOIs increase the average DIP release and its standard deviation (Figure 4.13B). This result suggests that higher number of initial STV particles infecting a single cell increases the potential for DIP to release higher amounts of progeny.

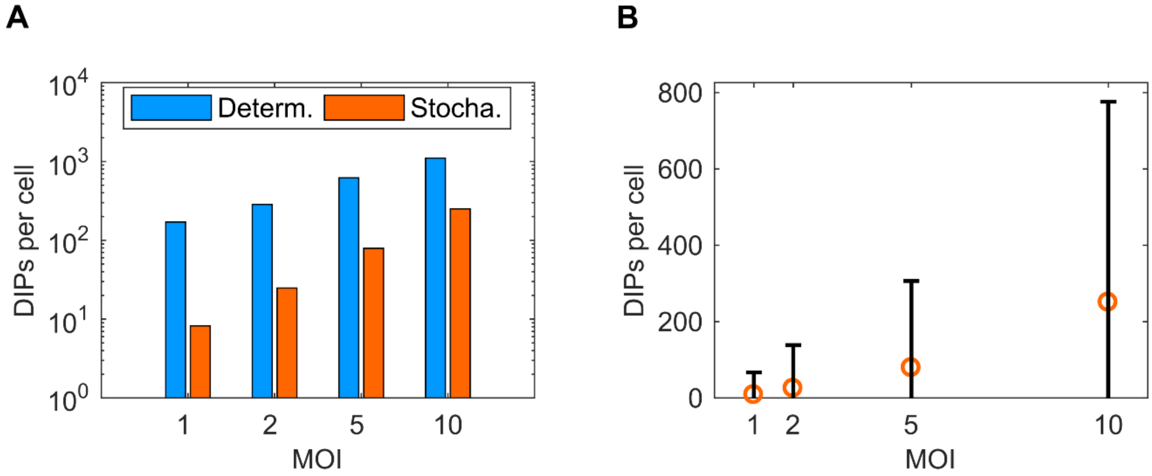


Figure 4.13 | Average cumulative DIPs released with increasing MOI. (A) Average cumulative DIP production of stochastic simulations and comparison with the deterministic solution. (B) Average cumulative DIP production and standard deviation. Represented results at 12 hpi and different MOIs.

When we maintain a constant MOI of 1 and increase the MODIP we observe that the differences between the average stochastic simulations and the deterministic solution are slightly more pronounced compared with the increased MOI scenario (comparing Figure 4.13A and Figure 4.14A). When the MODIP is increased, DIP production in the deterministic model barely changes. In contrast, the average stochastic simulations show a larger increase in DIP productivity (Figure 4.14A). On average, the stochastic model shows that an increased MODIP can benefit DIP production and its standard deviation does not vary significantly (Figure 4.14B).

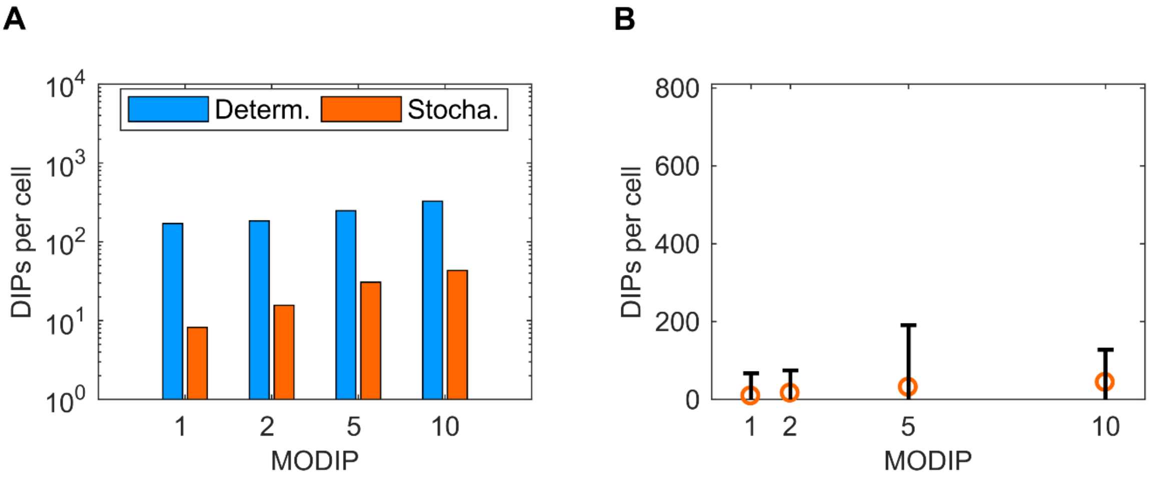


Figure 4.14 | Average cumulative DIP released and increased MODIP. (A) Average cumulative DIP production of stochastic simulations and comparison with the deterministic solution. (B) Average cumulative DIP production and standard deviation. Represented results at 12 hpi and different MODIPs.

One of the sources that causes variability in replications is the different vRNP levels inside the cell nucleus, so we defined two ratios and evaluated how it influences STV and DIP replication:

$$\frac{mean(Vp^{Nuc})}{FL\ Vp^{Nuc}\ Segment\ 3} \qquad \frac{DI\ Vp^{Nuc} Segment\ 3}{FL\ Vp^{Nuc} Segment\ 3} \tag{4.1}$$

The first ratio represents the disadvantage of the FL segment that corresponds to the DI segment and is calculated by dividing the average of all functional vRNP by the FL vRNP segment 3. The second ratio represents the statistical advantage of DI segment 3 over its complementary FL segment. These ratios were quantified in our simulations at 3 hpi, which represent a time at which viral replication has on average already started. In Figure 4.15 these simulations are represented by their STV and DIP release at 12 hpi. As this evaluation resulted in very disperse data clouds (Figure 4.15A,C), we decided to group STV and DIP release in subclasses of 50 particles per cell and for each we calculated the average of the respective ratio (Figure 4.15B,D).

We observe that higher STV production is achieved when the mean of FL vRNP over the FL vRNP segment 3 ratio is around 1 (Figure 4.15A,B). This shows that the level of FL segment 3 should be similar to the average vRNP level to increase STV production. If we focus on DIP replication lower values for the ratio of DI segment 3 over its complementary FL segment are beneficial for DIP production (Figure 4.15C,D). This result indicates that higher levels of FL vRNP segment 3 increases DIP release. However, high levels of DI vRNP seem to result in a disadvantageous self-inference.

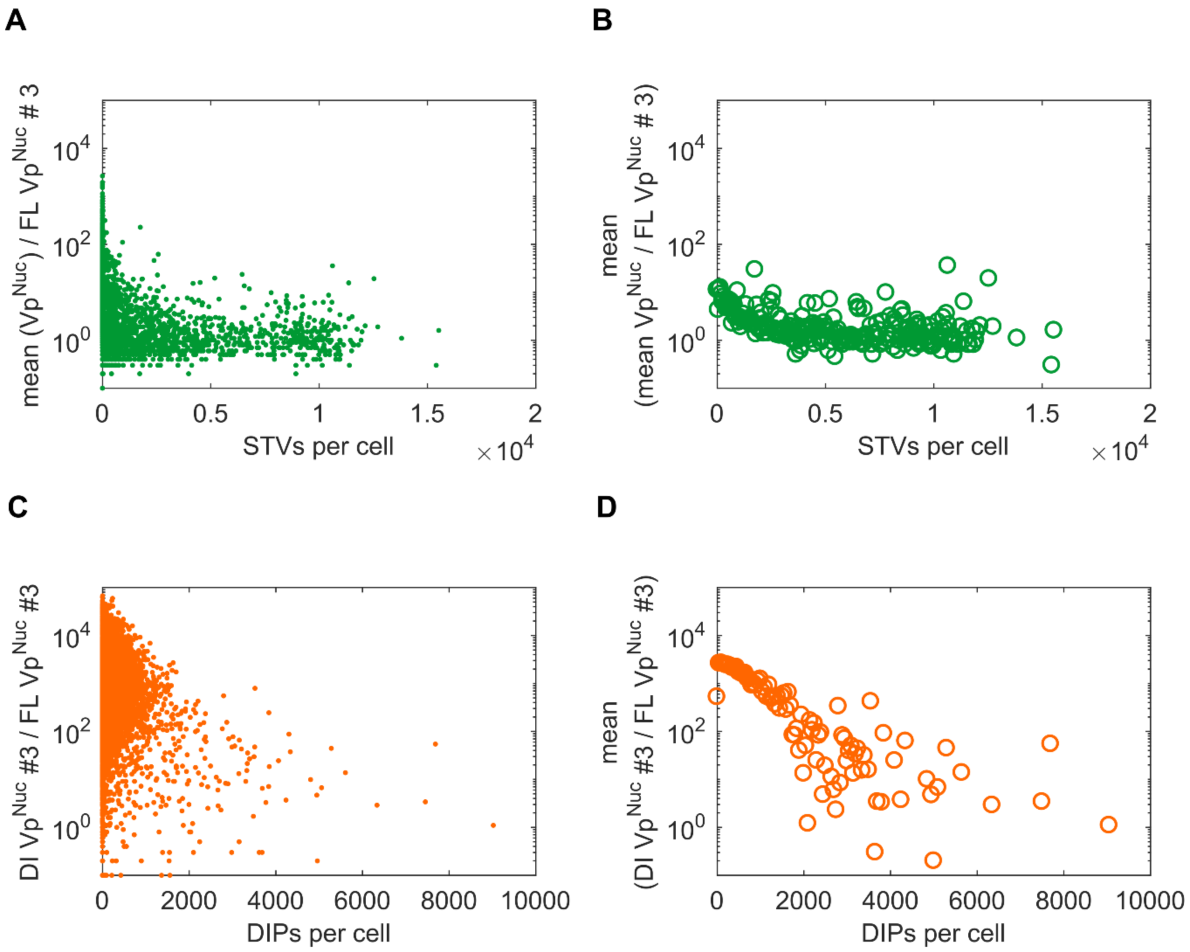


Figure 4.15 | vRNP ratios and impact on STV and DIP release. STV disadvantage ratio of the average of all functional vRNP over the FL vRNP segment 3 inside the nucleus and STV production (A) and the average of the respective ratio (B). Statistical DIP advantage ratio of DI segment 3 over its complementary FL segment inside the nucleus and DIP production (C) and the average of the respective ratio (D). Ratios are quantified at 3hpi and virus particle production at 12 hpi.

4.2 Non-productive simulations

4.2.1 Effects of MOI and MODIP

As shown in the previous section, in a MOI/MODIP 1/1 scenario most of the simulations do not achieve production of STV or DIPs. The two main reasons that induce non-productive cells are: (I) fusion failure which occurs when either STV or DIP fail to fuse their envelope with the endosomal membrane preventing them to enter the nucleus and (II) segment loss of at least one genome segment due to degradation of vRNA by cellular nucleases. In this section we will study how often these events occur for different combinations of MOI/MODIP and how this impacts STV and DIP replication. For the first comparison, we maintained the MODIP at 1 and increased the MOI. Then, we fixed the MOI at 1 and increased the MODIP.

First, we analysed the STV production. When we increased the MOI, we observe a 70% decrease in the number of non-productive cells (Figure 4.16A). Moreover, the impact of fusion failures and segment loss decreased, since higher MOIs lead to more STVs successfully entering the nucleus, which also support the replication of vRNPs. By contrast, co-infections with MOI 1 and higher MODIPs show that the percentage of unproductive simulations barely changes (slightly increases from 94% to 96%), highlighting a strong DIP interference due to its high factor advantage (Figure 4.16B). Furthermore, since the functional genome segments contained in the DIP support the replication of these vRNPs, the segment loss probability is decreasing.

It is important to notice that for both scenarios the impact of other reasons preventing STV release increases with higher initial MOI or MODIP conditions. These other reasons contemplate, e.g., an imbalanced replication of the genome segments or DIP interference. As explained before, since our model considers that all genome segments can replicate independently, higher availability of vRNPs for replication can increase the differences of independent segments levels during replication and impair STV release. Furthermore, increasing MODIP can result in a considerably higher DIP interference which impairs STV replication.

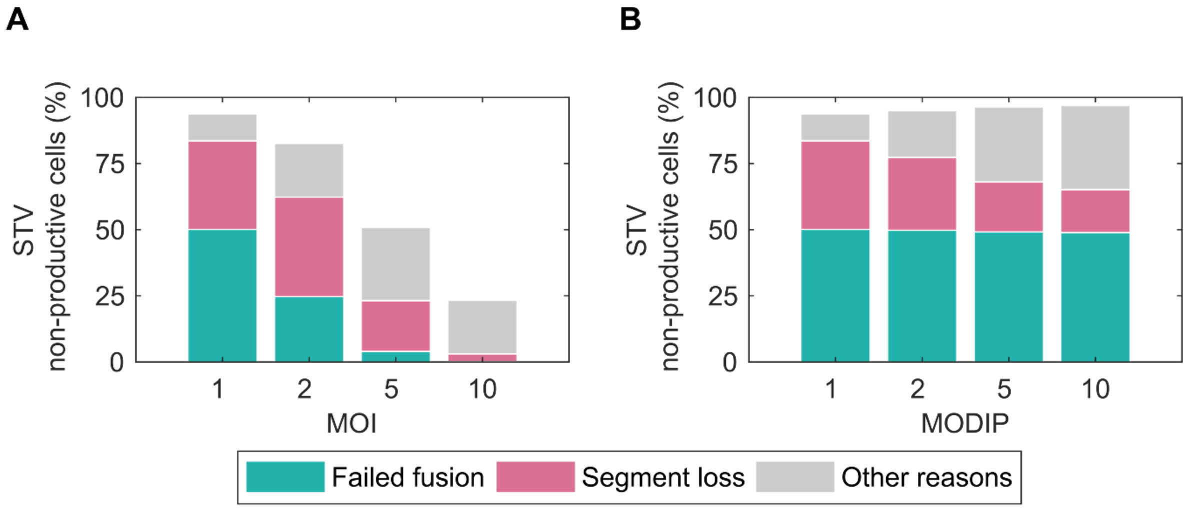


Figure 4.16 | Factors which develop STV non-productive cells. (A) Percentage of STV non-productive cells at different MOIs. (B) Percentage of STV non-productive cells at different MODIPs.

When we focus on cells that do not produce DIPs, our simulations reveal that the increase of either the MOI or the MODIP lead to very similar scenarios. (Figure 4.17A,B). Increasing either MOI or MODIP will lower the number of cells that do not produce DIPs. If the MOI is increased, more resources (vRNPs) are available for DIP replication. If the MODIP is increased, the chances of the DI vRNP segment to be synthesised for DIP production is higher. Furthermore, when MOI or MODIP is increased to 10, we observe a 58% and 54% reduction in the non-productive cells, respectively (when we compare the results for MOI or MODIP of 1 and 10). The failed fusion and segment loss probability decreased as well.

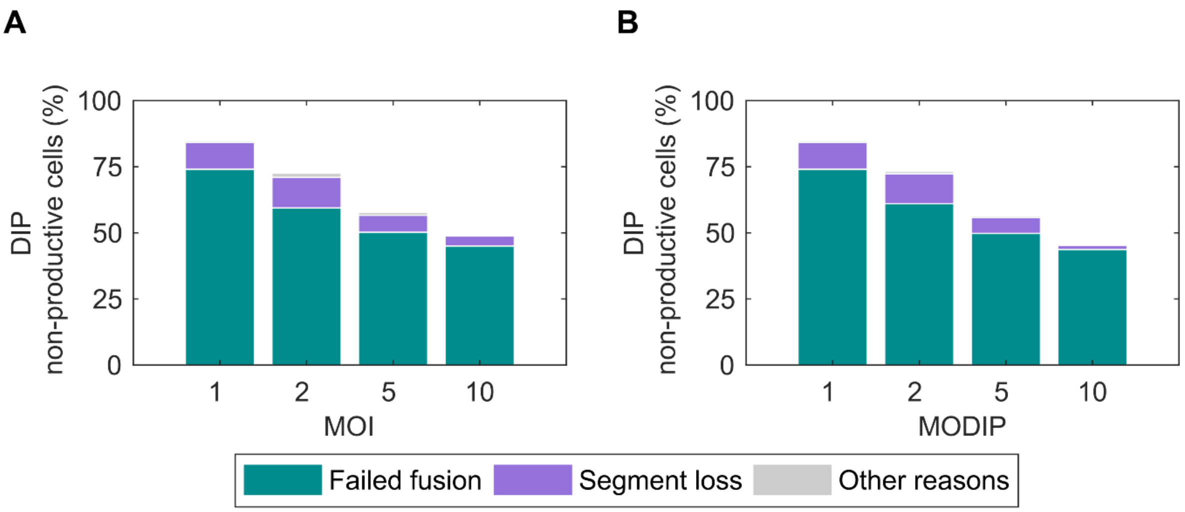


Figure 4.17 | Factors preventing DIP production. (A) Percentage of DIP non-productive cells at different MOIs with a fixed MODIP of 1. (B) Percentage of DIP non-productive cells at different MODIPs with a fixed MOI of 1.

As seen before, DIP replication requires that both DIP and STV successfully fuse and enter the cell. By contrast, STV propagation depends solely on its fusion although DIP can interfere and lower STV release. Regarding segment loss, as shown before, the DI segment has a replication advantage causing faster replication, which lowers the probability of DI segment to be degraded compared to a regular FL vRNA segment. Overall our simulation results expose these factors and we can conclude that failed virus particle fusion strongly affects DIP replication and segment loss has a bigger impact on most STV replication.

Since DIP replication depends on both STV and DIP, the reasons that cause cells to not produce DIP are much more complex. We wanted to analyse these factors in more detail, so we split the cells that do not release DIPs in different subcategories. The cells that showed failed fusions were grouped into simulations in which either DIP or STV fails to fuse or both fail to fuse simultaneously. The segment loss category was sub divided depending on if the DI or at least one FL segment got degraded.

When we maintain a constant MODIP of 1 and increase the MOI, more particles are available to infect a cell. Accordingly, the probability of simultaneous fusion failure of STV and DIP is reduced (Figure 4.18A). Furthermore, the number of simulations in which STV fails to fuse alone decreases as well. In the same scenario, the FL segment loss probability decreases due to higher levels of functional vRNA which prevents its degradation (Figure 4.18B). Additionally, higher MOI increases the probability of DI segment loss. Although DIP itself has a replication advantage, when only one DIP enters a cell and the MOI is higher, STV can overcome this advantage. High MOI allow STV to replicate faster and consequently can win the competition for the viral polymerase complexes, which in turn increase the FL vRNP levels even more. Furthermore, these higher levels of FL vRNAs can deplete the available NP, which is essential for vRNA stabilization. Without stabilization, DI segments have a higher chance to be degraded.

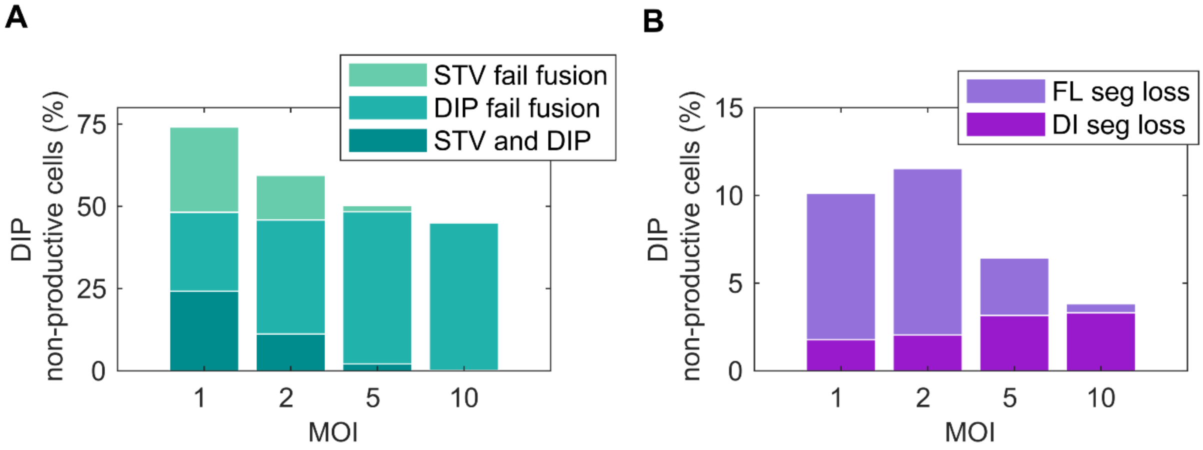


Figure 4.18 | Detailed factors which prevent DIP production at different MOIs and a fixed MODIP of 1. (A) Percentage of factors which cause failed fusions. (B) Percentage of factors which cause segment loss.

Interestingly, when the MOI is increased from 1 to 2 the amount of FL segment loss events increased. This raised the question about what percentage of segment loss it is observed for each genome segment. To that end, we calculated the percentage of each segment that was degraded over the number of all segment loss simulations. As explained previously, our data shows that when the MOI is increased, most of the FL vRNA segments have a lower degradation probability (Figure 4.19). The DI segment, however, shows an almost constant segment loss probability with increased MOI. The FL segment 3 has considerably higher segment loss susceptibility than the others FL segments.

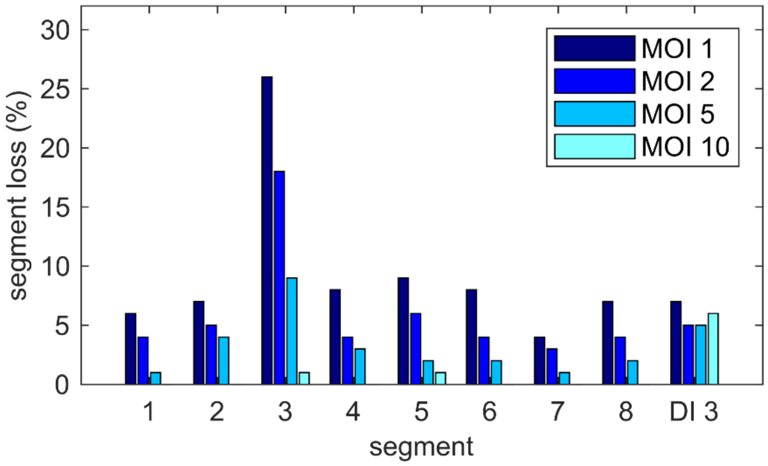


Figure 4.19 | Percentage of segment loss for each vRNA segment at MODIP 1 and different MOI.

When we fix the MOI at 1 and increase the MODIP, as we have only one STV to infect a cell, the number of simulations in which only STV fails to fuse increases (Figure 4.20A). By contrast, the number of simulations in which all DIPs fail to fuse decreased. Additionally, we observe a reduction in the number of events in which STV and DIP simultaneously fail to enter the cell at higher MODIP. As expected, higher levels of DI vRNA will prevent loss of the defective segment (Figure 4.20B). Furthermore, the FL segment loss is reduced as well when MODIP is increased since DIP contributes with its functional genome segment to increase the overall FL vRNA levels lowering their degradation probability with exception to FL segment 3.

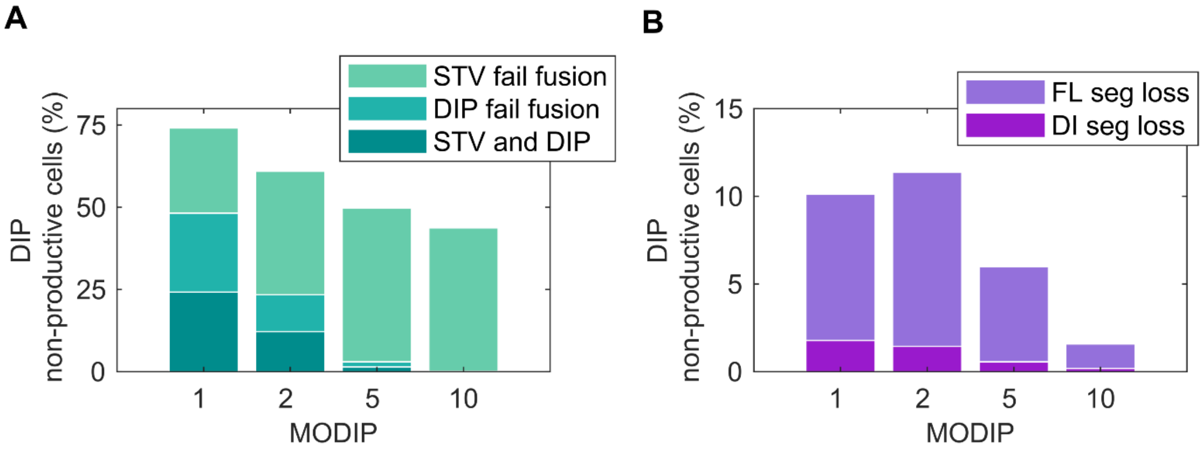


Figure 4.20 | Detailed factors which prevent DIP production at different MODIPs and a fixed MOI of 1. (A) Percentage of factors which cause failed fusions. (B) Percentage of factors which cause segment loss.

Again, we evaluated the segment loss event in individual vRNA segments. Overall our results reveal that vRNA degradation is reduced with higher MODIP with exception of FL segment 3 (Figure 4.21). It is important to highlight that, due to its lower level, segment 3 has by far the highest probability of being degraded and for all MODIP scenarios. Segment 7 has a lower segment loss probability due to our aborted simulations criterium.

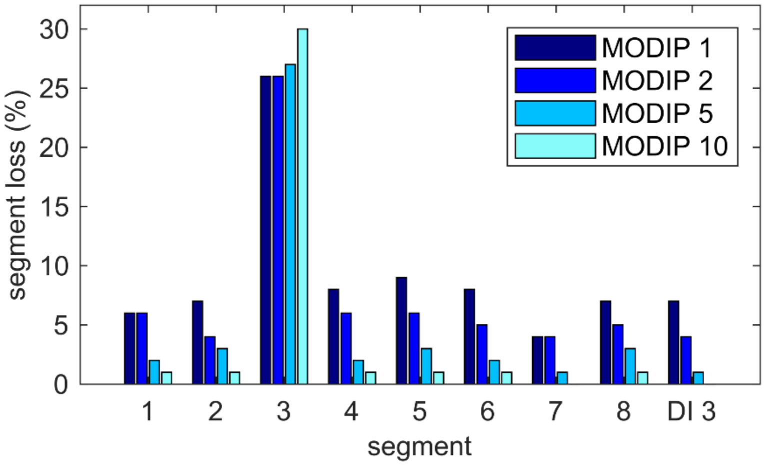


Figure 4.21 | Percentage of segment loss for each vRNA segment MOI 1 and different MODIP.

Finally, we wanted to confirm if all segment loss simulations result in non-productive cells. Surprisingly, in a total of 932 simulation that showed segment loss a few cells indeed did produce progeny particles. As expected, when the DI segment is degraded STVs can be produced (Table 4.1). When FL segment 3 was degraded, 220 simulations showed DIP release in which 57% or these runs produced more than one DIP. In addition, we found release of exactly one progeny particle in simulations with the other segment loss occurrences. We suggest that this occurred due to the vRNP being exported from the nucleus before it could replicate. In summary, the release of 1 progeny particle should not be considered production, but instead a re-packaging of the incoming vRNP.

Table 4.1 | Number of simulations which showed STV or DIP production despite losing (at least) one genome segment

	Segment lost								
	#1	#2	#3	#4	#5	#6	#7	#8	DI #3
STV Productive Cells	3	0	1	3	1	0	0	0	48
DIP Productive Cells	13	13	229	12	21	21	4	5	1
	Number of Simulations								

4.2.2 Filtering non-productive simulations

In the previous sections we analysed how different MOI and MODIP scenarios affect DIP production. We concluded that at MOI/MODIP 1/1 the differences between our implemented stochastic model and the deterministic simulation are more pronounced. Furthermore, in this scenario the stochastic effects have a higher impact which will result in increased number of non-productive simulations. Hence, first we filtered out the non-productive cells which were induced due to fusion failure and next excluded the simulations which resulted in segment loss. Then, we compared the filtered results with the deterministic model. To that end, we examined the DIP release and NP dynamics due to their importance for our analysis so far.

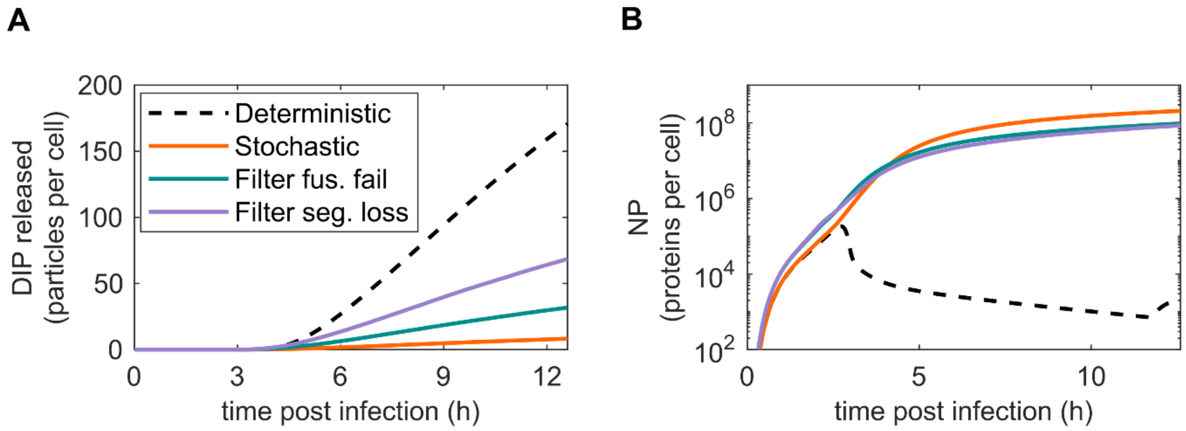


Figure 4.22 | Filtered results and comparison with the deterministic model at MOI/MODIP 1/1. Comparison of the mean stochastic result of the unfiltered and filtered simulations with the deterministic model for total DIP release (A) and NP (B).

When we exclude the non-productive simulations, we observe that the difference between our average stochastic simulations and the deterministic model are reduced, although significant differences remained. DIP release initially differed with a factor of 21 and after filtering the difference was reduced to a factor of 2.5 which reflects a good improvement (Figure 4.22A). Regarding NP, even after filtering the simulations we obtain very high deviations from the deterministic simulation (Figure 4.22B).

Since filtering of the non-productive cells did not greatly improve the observed differences to the deterministic model for DIP release and NP dynamics, we calculated the standard deviation for these molecules to understand if this parameter would be improved. We can observe that since the standard deviation is very high, when we filter the simulations we obtain a reduced standard deviation for DIP release (Figure 4.23A). However, for NP the standard deviation is lowered at the beginning but remains unchanged at 12 hpi (Figure 4.23B). This observation further supports that the wide-spread distribution of NP, which results from a mixture of different stochastic events, is an inherent characteristic of the stochastic model and is not improved when we focus on the productive simulations.

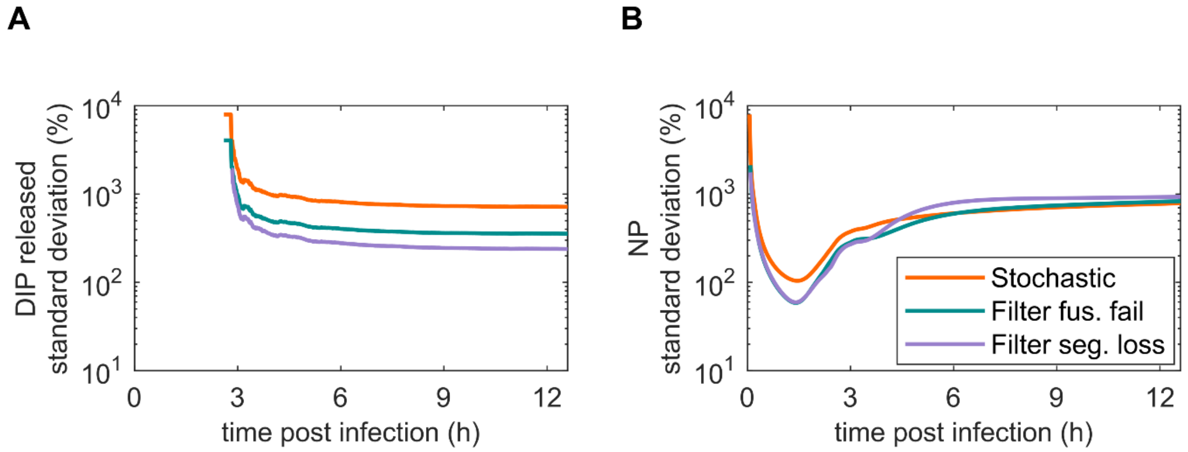


Figure 4.23 | Standard deviation and filtered results during infection at MOI/MODIP 1/1. Standard deviation of the total DIP release (A) and NP (B) divided by the mean stochastic result at the same time point during simulation.

4.3 Timing investigation

4.3.1 Randomly generated delay

In the previous section we analysed the factors that prevent particle production in our simulations and identified that both STV and DIP entry have a major influence on the infection progression. The stochastic model considers that when STV and DIP are placed in the extracellular medium at the same time, each particle can take different times to travel through the cytoplasm and reach the nucleus. The stochastic model simulates various time delays during STV and DIP import, e.g. delays for cell entry, fusion or nucleus entry. In this section we will study the time delay between STV and DIP nucleus entry, i.e. nuclear import, since this is the factor that determines the beginning of replication. We divided our simulations in the two possible scenarios regarding nucleus entry: either STV enters first so the DIP entry is delayed, or the DIP reaches first the nucleus and STV entry is delayed.

As we showed before, fusion failure is the major factor that prevents progeny production. Since we are interested in analysing the impact of the delay in production, we filtered our results: for this section we only considered simulation in which both STV and DIP fused successfully, and we obtained a total of 3313 runs for this initial infection conditions. The nucleus entry delay was divided into class sizes of 0.1 hours which resulted in a normal distribution with mean approximately zero (Figure 4.24A). However, the probability for STV and DIP entering the nucleus at exactly the same time is low, as only eight simulations showed this feature.

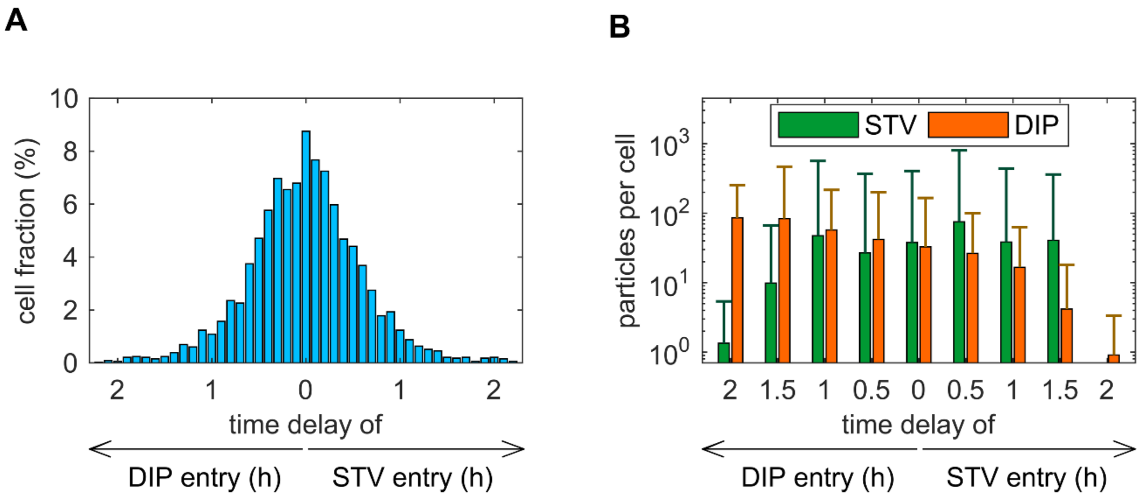


Figure 4.24 | Nucleus entry delay influence STV and DIP production. (A) Nucleus entry delay distribution. (B) Average STV and DIP release and standard deviations for different delays at 12 hpi. Only simulations in which both STV and DIP fused successfully were considered.

Our simulations reveal that on average DIP production is increased when DIP entry is delayed (Figure 4.24B). This scenario benefits DIP production, although STV release is impaired. This last result is unexpected since increasing DIP delay should benefit STV production. This indicates that the average release might not be representative due to low number of simulations obtained for 1.5 and 2 hours delay. If STV entry is delayed the production of both STVs and DIPs is impaired. When STV entry is delayed DIP release is more strongly affected than STV.

In the previous section we found that simulations which release exactly one progeny particle can be caused by re-packaging of the particle that entered the cell. Therefore, from here on we will consider these cells as non-productive. In addition, since for the scenario of a 2 h or larger entry delay we obtained a low number of simulations, we will analyse this case in the next section. The total number of simulations obtained for each time delay can be consulted in the appendix (Table A.2.2).

The distribution of STV release is wide and heterogeneous (with a maximum release of 10616 STV per cell at 0.5h post DIP entry) which impedes analysis and graphical representation. For that reason, we will focus our examination to cells that do not produce STVs. Even though we excluded all simulations with fusion failures, for each time delay most cells do not produce STVs (Figure 4.25). When DIP entry is delayed the STV non-productive cells are reduced by 8% when compared to the 0.5h and 1.5h delay scenarios (Figure 4.25A). When DIP enters first the percentage of cells that do not release STV are always higher than 95% (Figure 4.25B). This result shows that increased delays in DIP entry lead to less STV non-productive simulations.

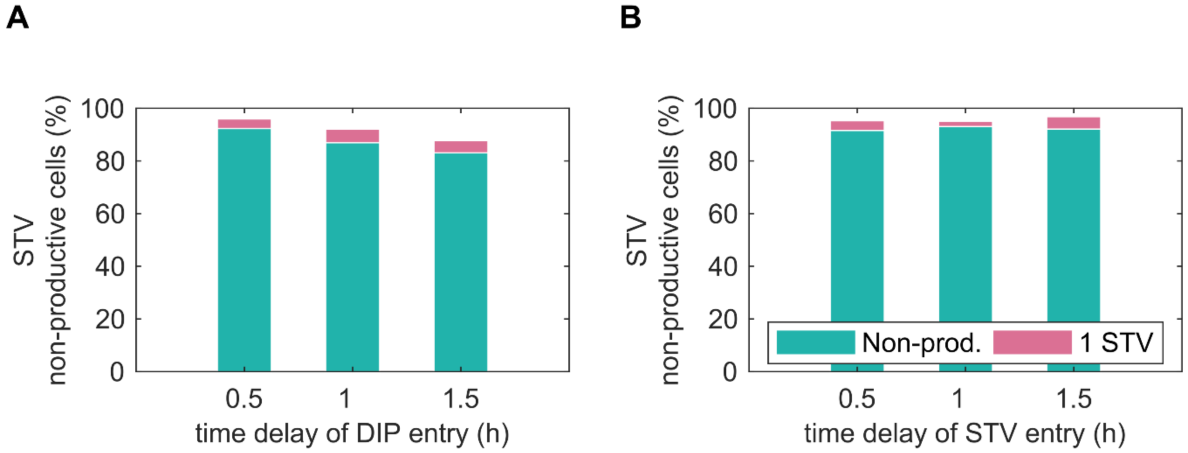


Figure 4.25 | STV non-productive and 1 STV release simulations and nucleus entry delay. Percentages of STV non-productive and 1 STV release cells when DIP entry is delayed (A) and STV entry is delayed (B) Only simulations in which both STV and DIP fused successfully were considered.

Overall both distributions of DIP production, whether STV or DIP entry is delayed, are skewed to the right and the average DIP release for all time delays is improved when DIP entry is delayed (Figure 4.26A,B). The maximum release was 3617 DIP per cell when DIP entry is delayed 0.5h. By contrast when STV entry is delayed the maximum release was only 1002 DIP per cell. This data supports that DIP production is improved when STV enters first the nucleus.

When DIP entry is delayed 1.5h the maximum value we obtained from the simulations is 3042 DIPs per cell. In this scenario have the higher average DIP production. When DIP entry is delayed 1h we obtained more simulations that release more than 200 DIPs per cell however the maximum DIP release was only 1698 DIPs per cell. This suggests that DIP production is improved when the time delay of DIP entry is increased till 1.5h, although due to low number of simulations obtained at 1.5h DIP delays we cannot confirm this hypothesis based on our simulations.

The percentage of cells that do not produce DIPs when DIP entry is delayed is much lower than when STV entry is delayed (Figure 4.26C,D). When DIP entry is delayed we observe a 5% increase in cells that do not produce DIPs although when STV entry is delayed data shows a 21% increase in the non-productive simulations.

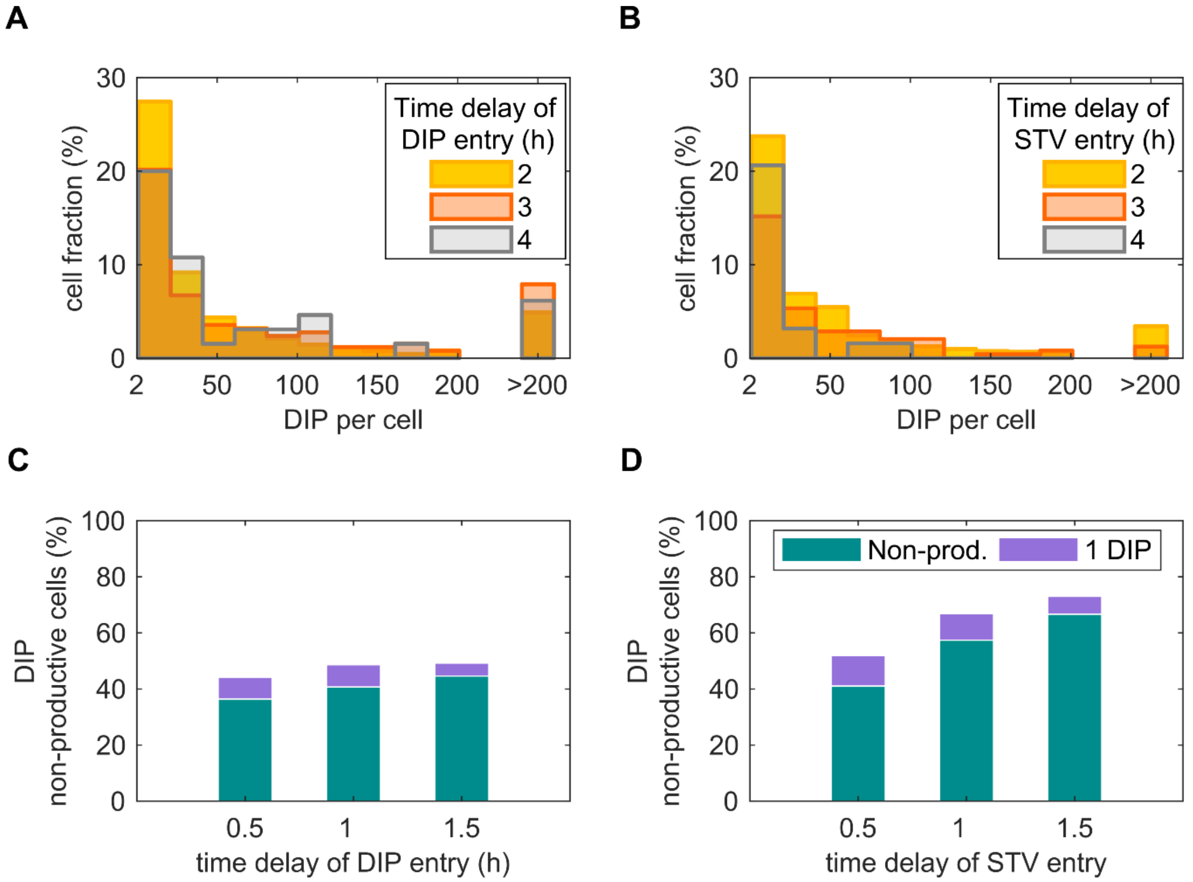


Figure 4.26 | DIP release and non-productive simulations with nucleus entry delay. (A,B) Distribution of DIP total production at 12 hpi when DIP entry is delayed (A) and STV entry is delayed (B). (C,D) Percentages of DIP non-productive and 1 DIP release cells when DIP entry is delayed (C) and STV entry is delayed (D). Only simulations in which both STV and DIP fused successfully were considered.

4.3.2 Induced delay

Previously, we investigated the effect of time delays between the nuclear import of STV and DIP by observing the results of our random simulations. However, the time delay distribution showed that delays higher than 2 hours did occur very rarely with the settings we used. Since we were also interested to analyse how higher delays would impact STV and DIP replication, we developed a strategy to induce an exact nucleus entry delay. To do that, we simulated our model with either one STV or one DIP outside the cell. The virus entry was simulated normally, and the time point of nuclear import of the respective vRNPs was determined. Then, after a specific time delay we introduced the vRNP segments of the other particle directly inside the nucleus and proceeded with the rest of the simulation. We performed simulations for time delays of 2, 3 and 4 hours in both directions and only considered simulations in which the initial particle fused successfully. The total number of simulation runs performed for each condition can be consulted in the appendix (Table A.2.3).

Increased DIP delays result in higher mean STV production (Figure 4.27A). This information can be confirmed as well by the violin plots in which dots represent the exact STV release (Figure 4.27C). We observe that delaying DIP entry by 4 hours increases STV production. This result was expected, since a delayed DIP entry provides more time for the STV to replicate without interference, which increases its production. By contrast, a delayed STV entry leads to a decrease in the mean STV release (Figure 4.27B) and more than 95% of simulations do not generate progeny virions (Figure 4.27D).

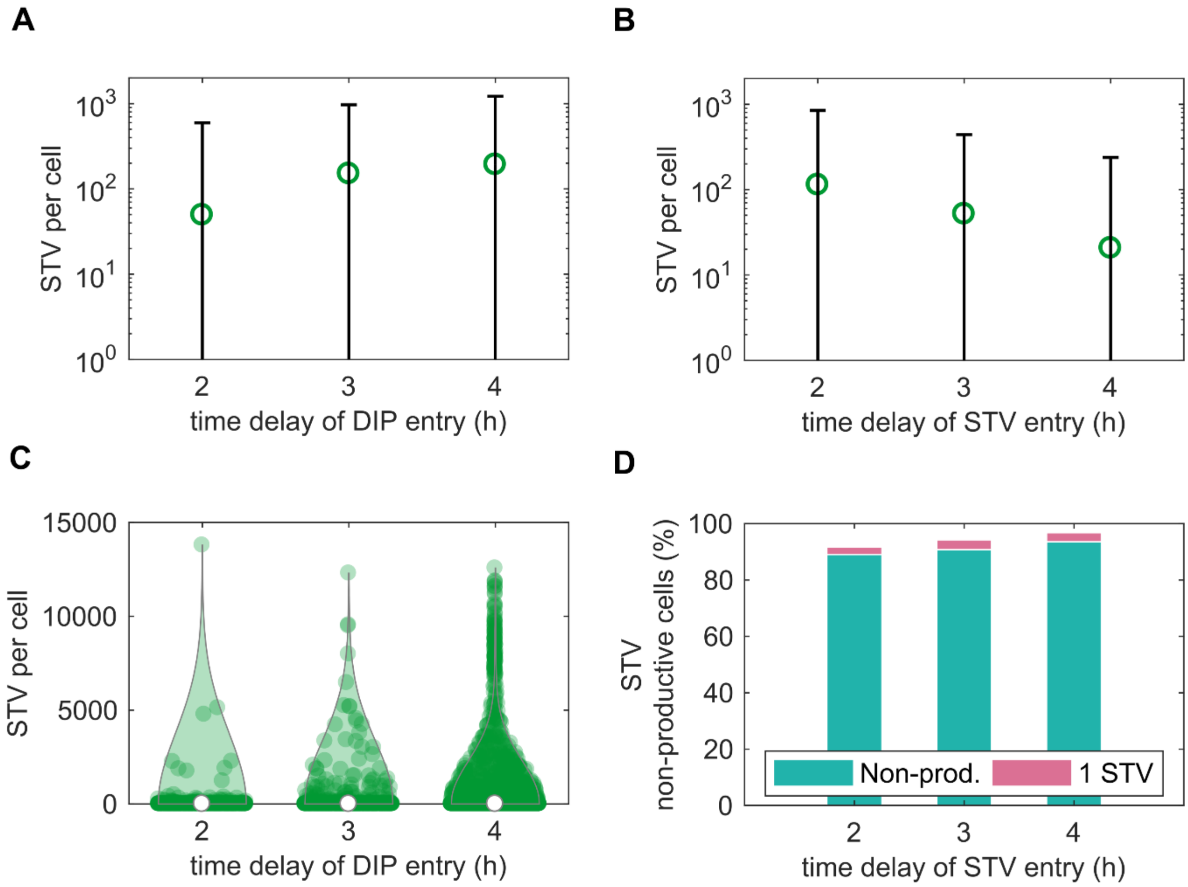


Figure 4.27 | STV release and non-productive simulations with an induced delay. (A,B) Average STV production and standard deviation at 12 hpi when DIP entry is delayed (A) and STV entry is delayed (B). (C) STV production distribution when DIP entry is delayed. Each dot represents an individual simulation result for STV release. (D) Percentage of cells that release 0 or 1 STV when STV entry is delayed. Only simulations in which both STV and DIP fused successfully were considered.

Overall our data show that the mean DIP release is increased when DIP entry is delayed (Figure 4.28A,B). However, delaying DIP entry more than 2 hours decreases its average release. This indicates that an already advanced STV replication can suppress DIP production (Figure 4.28A). When we focus on DIP distribution, we observe that a DIP entry delay between 2 and 4 hours enables some cells to produce more than 600 DIPs per cell (Figure 4.28C). In the scenario that DIP entry is delayed exactly 2 hours we observe a higher number of cells that release less than 300 DIPs per cell when comparing with higher other delays. Furthermore, the same scenario shows a higher number of simulations which produce more than 600 DIPs per cell as well. By contrast, when STV entry is delayed, it will strongly decrease DIP production and we observe that most of the productive cells release less than 150 DIPs per cell (Figure 4.28D).

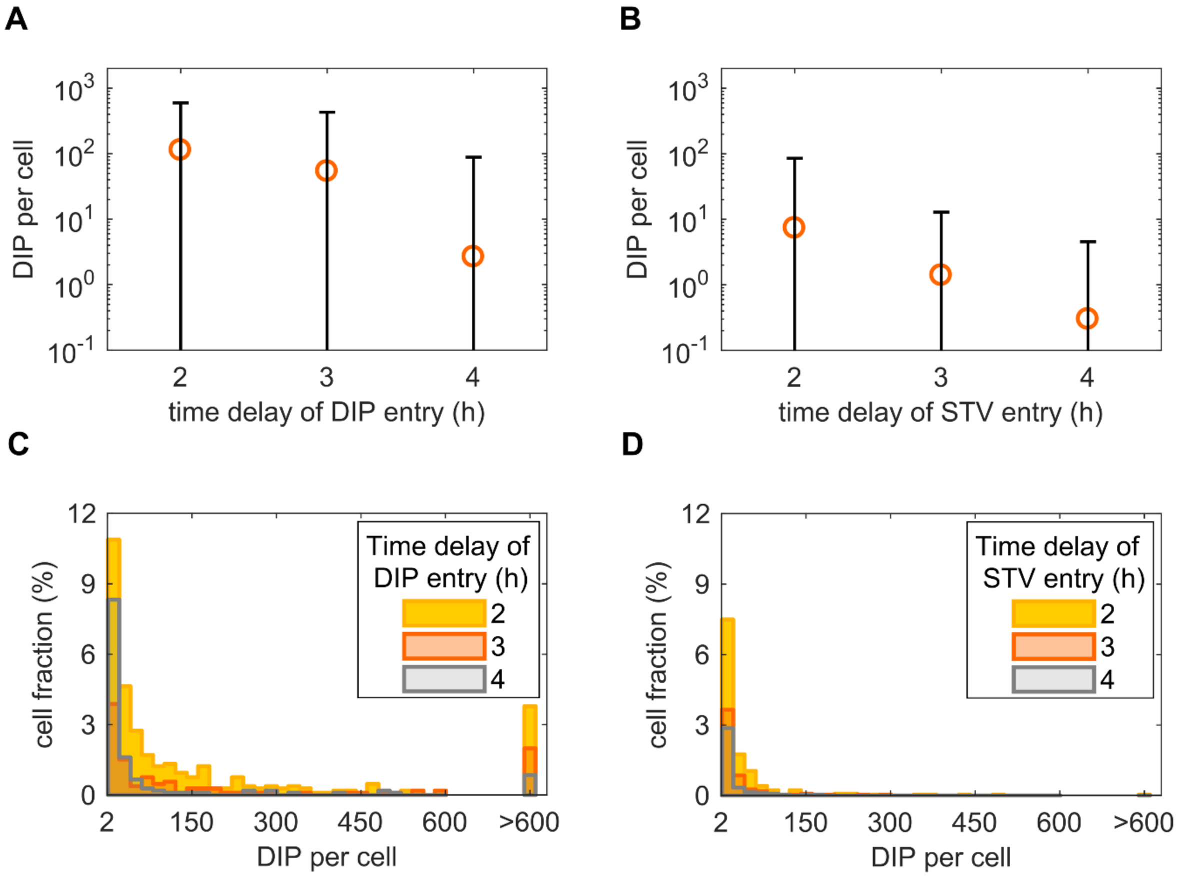


Figure 4.28 | Distribution of DIP release with induced delay. (A,B) Average DIP production and standard deviation when DIP entry (A) and STV entry is delayed (B). (C,D) Distribution of DIP production when DIP entry (C) and STV entry is delayed (D). Results were obtained at 12 hpi and only simulations in which both STV and DIP fused successfully were considered.

In the previous paragraphs we analysed the randomly generated and induced delays separately. To achieve a better overall understanding of how the time delay affects DIP production, we combined both data sets: the randomly generated delay ranges between 0.1 and 1.5 hours and the induced delay data corresponding to the range between 2 and 4 hours for a delay of either STV or DIP entry.

When DIP entry is delayed between 0 and 2 hours, the mean DIP production increases. Moreover, if the delay is increased further, DIP release decreases (Figure 4.29A). By contrast, the mean DIP production decreases with increased delays in STV nucleus entry (Figure 4.29B). However, we observe that for a delay of STV entry of exactly 2 hours, the mean DIP yield is higher than when STV entry is delayed 1.5 hours. This discrepancy in the DIP release trend can be explained by the low number of simulations obtained (63 runs) for an STV entry delay of 1.5 hours. This could result in an average DIP production that is not representative for this scenario, which is also supported by the increased standard deviation.

The maximum mean DIP production is achieved when DIP entry is delayed between 1.5 and 2 hours (Figure 4.29A). However, this range is right at the limit between the randomly generated and induced delay. In addition, similar to the observation in the previous paragraph, there is a possibility that the calculated mean DIP yield is not representative since we obtained only 65 simulations for this category. In summary, we determined that the optimum average DIP production is achieved when DIP entry is delayed between 1.5 and 3 hours.

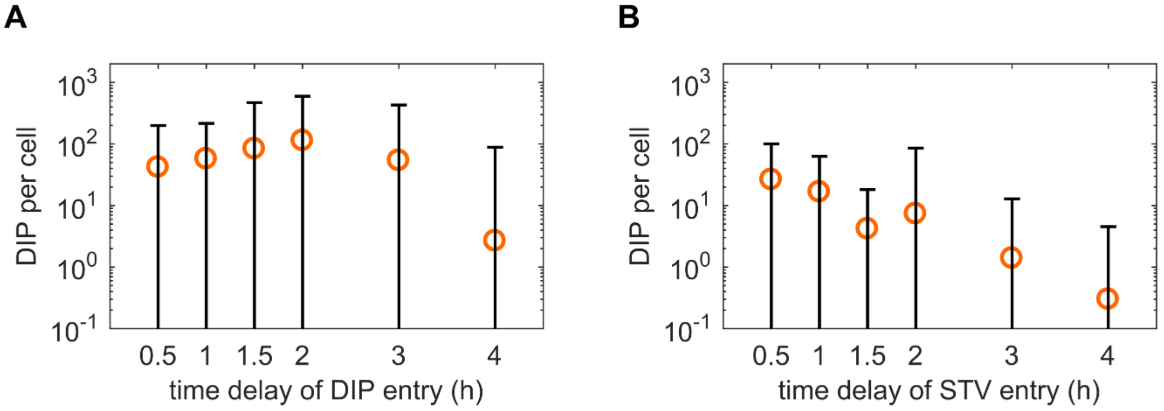


Figure 4.29 | DIP release results with compiled random and induced delays. Average DIP production and standard deviation when DIP entry (A) and STV entry is delayed (B). Results were obtained at 12 hpi and only simulations in which both STV and DIP fused successfully were considered.

In the previous sections we focused solely on the evaluation of DIPs carrying a defective segment 3, which is one of three segments that encode a protein of the polymerase complex. Since other DI segments can emerge in DIP *de novo* generation, we contemplated how DIP production would be affected by a defective segment encoding a protein that is non-essential for RNA synthesis. To that end, we analysed DIPs with a defective segment 4, which encodes for HA. We performed simulations for 2, 3 and 4 hours delay between either STV or DIP nucleus import and compared them to the results obtained for DI segment 3. The total number of simulations performed for each scenario can be consulted in the appendix (Table A.2.4).

Overall when the DIP entry is delayed, the average DIP release was very similar regarding if it was carrying a defective segment 3 or 4 (Figure 4.30A,B). However, we observe some differences in the distribution of DIP production (Figure 4.30C,D). When DIP entry is delayed 2 hours our data show that a DI segment 3 will increase the number of simulations that release DIPs (Figure 4.30C). Furthermore, delaying DIP entry by 4 hours with a DI segment 4 impairs the release since we observe that few simulations release less than 150 DIPs per cell (Figure 4.30D).

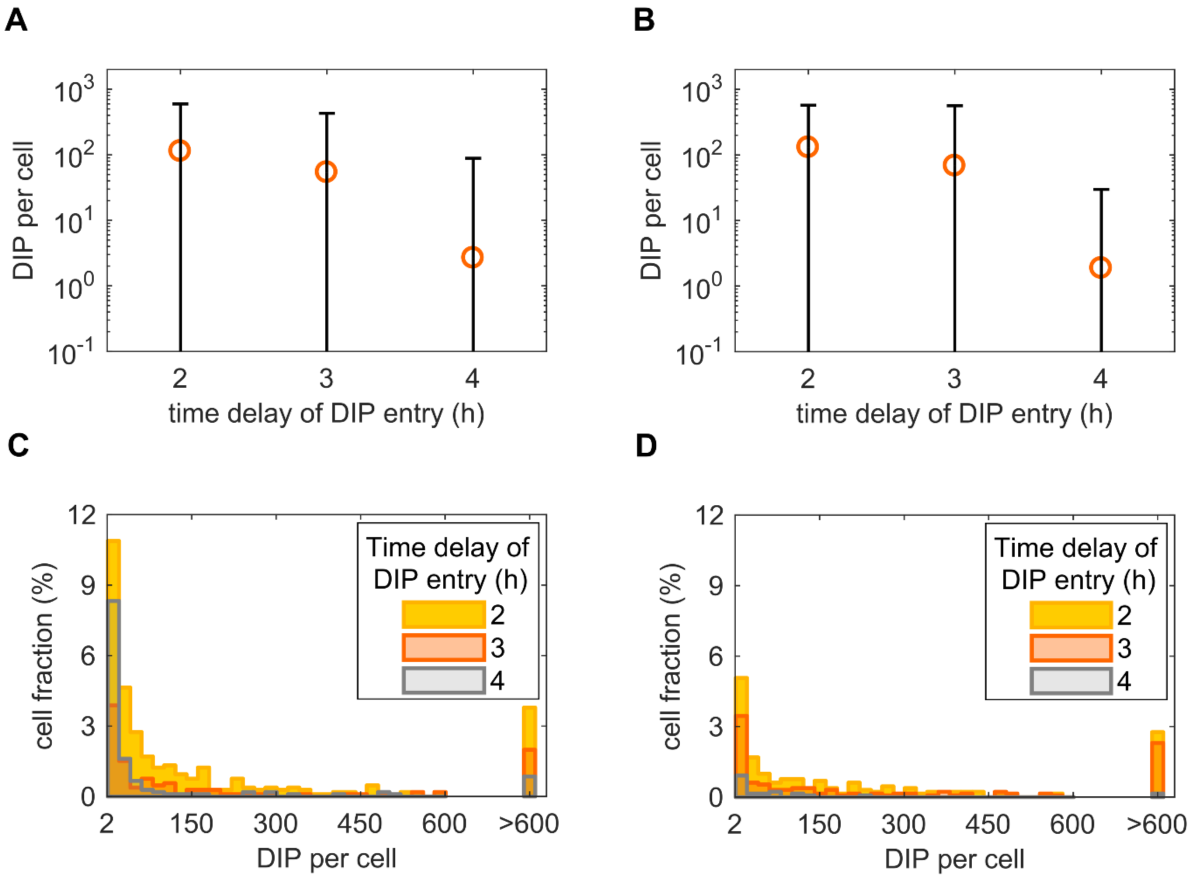


Figure 4.30 | DIP release when DIP entry is delayed for DI derived from segments 3 and 4. The DI segments originate from segment 3 (A,C) and segment 4 (B,D). (A,B) Average DIP production and standard deviation. (C,D) Distribution of DIP production with the indicated delay of DIP entry. Results were obtained at 12 hpi and only simulations in which both STV and DIP fused successfully were considered.

In the scenario that DIP carries a defective segment 3, when STV entry delay is increased, the mean DIP production decreases strongly (Figure 4.31A). However, if DIP is originated from a DI segment 4, our result shows a less pronounced decrease in the average DIP yield (Figure 4.31B).

When we focus in the DIP production distribution when STV entry is delayed, we can observe some differences regarding if DIP carries a defective segment 3 or 4 (Figure 4.31C,D). Less than 6% of the cells produce less than 10 DIPs per cell with DI segment 3 when STV entry is delayed 2 hours (Figure 4.31C). However, in the same scenario when DIP carries a defective segment 4 more than 12% of the cells release less than 10 DIPs per cell (Figure 4.31D).

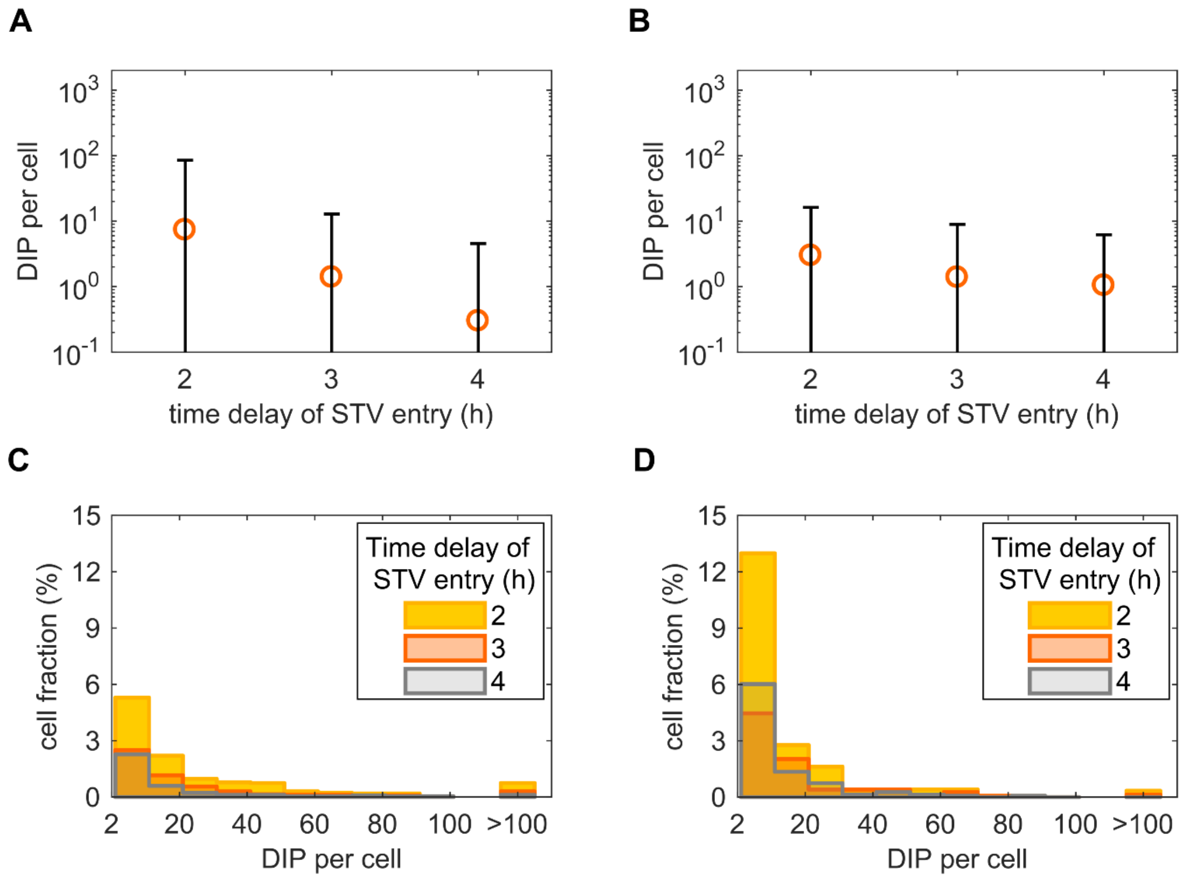


Figure 4.31 | DIP release when STV entry is delayed for DI derived from segments 3 and 4. The DI segments originate from segment 3 (A,C) and segment 4 (B,D). (A,B) Average DIP production and standard deviation. (C,D) Distribution of DIP production with the indicated delay of STV entry. Results were obtained at 12 hpi and only simulations in which both STV and DIP fused successfully were considered.

4.4 Model variations

4.4.1 Minimum packaging model

As explained in the previous chapters of this thesis, we know that DIP has an advantage over STV replication. However, is it still unknown in what part of the replication process this advantage emerges and how to quantify it. In this section we implemented variations in our model regarding the packaging advantage of DIPs and in the next section we will study different values for the DIP replication advantage. Our goal was to analyse how different approaches of implementing a DIP advantage affect both STV and DIP production.

In our model the packaging of STVs and DIPs includes two key steps: first all eight vRNP genome segments congregate to form a complex and in the next step STV and DIP complexes compete for proteins to incorporate into progeny particles. In the original model the complex formation rate is proportional to the product of each vRNP abundance in the cytoplasm (Equations (3.47) and (3.48)).

This leads to an almost instantaneous formation of the complexes, because normally the most abundant vRNPs in the cytoplasm reach high numbers. We modified this complex formation rate and in the new model version this rate is proportional to the least abundant vRNP in the cytoplasm (Equations (3.49) and (3.50)). To distinguish both models during our analysis the original model was named *Complex Model* and the modified one *Minimum Packaging Model.*

Since we want to analyse our result in a more precise manner, we decided to exclude possible sources of variability in our data. To that end, we moved away from analysing the data by the MOI and focused on the amount of STVs or DIPs that were imported into the nucleus. These particles which successfully entered the nucleus will be referred to as fused STVs and fused DIPs. Thus, we excluded the effects of fusion failures that strongly increase the heterogeneity in the simulations. Furthermore, we can observe that in an MOI/MODIP 10/10 scenario the fused STVs and DIPs show a Poisson distribution and there are various different combinations possible for the initial number of particles that entered the nucleus successfully (Figure 4.32A,B).

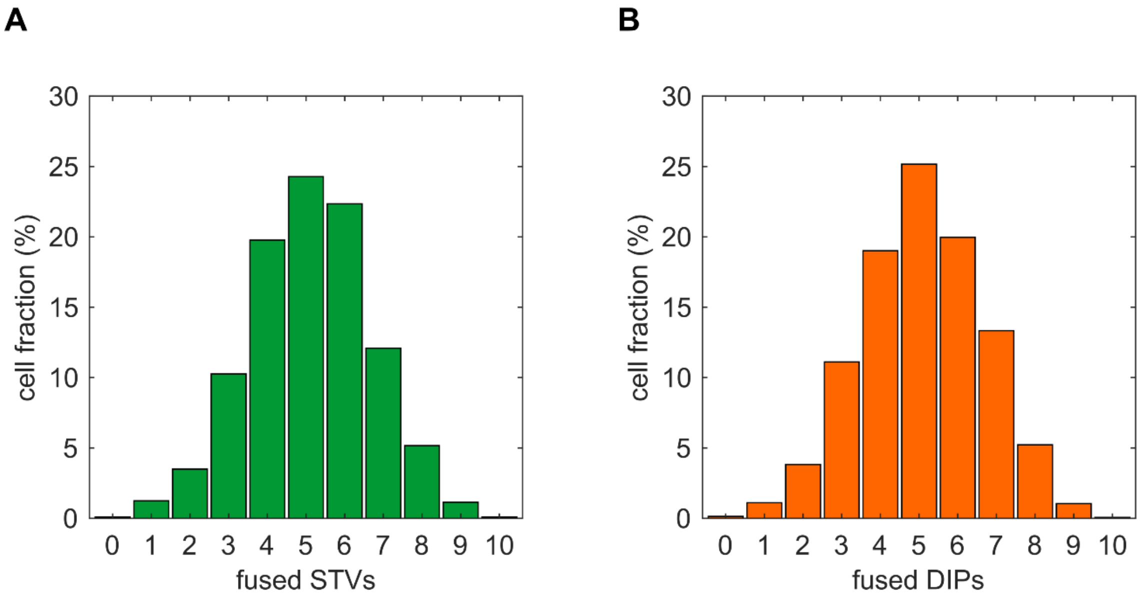


Figure 4.32 | Distribution of fused particles at MOI/MODIP 10/10. Distribution of fused STVs (A) and DIPs (B).

To elucidate how the number of fused STVs and DIPs impacts the release we performed several simulations with different MOI/MODIP combinations and grouped our results by the amount of fused particles. For each group we calculated the mean STV and DIP production. The results are illustrated in heat maps where the grey colour represents areas for which we did not obtain any simulations and the dark blue regions correspond to an average production close to zero.

As it would be expected, STV production is maximized when DIP fails to fuse and increases with the number of STVs fused (Figure 4.33A). Furthermore, the mean STV release is nearly zero when at least two DIPs fused. Interestingly, our data show that increasing the number of fused DIPs does not benefits its release. However, DIP yield is increased in a scenario that few numbers of DIPs successfully enter a cell and higher amount of STV fused (Figure 4.33B). More precisely, the maximum average DIP production obtained was when 9 STV and 1 DIP entered the nucleus.

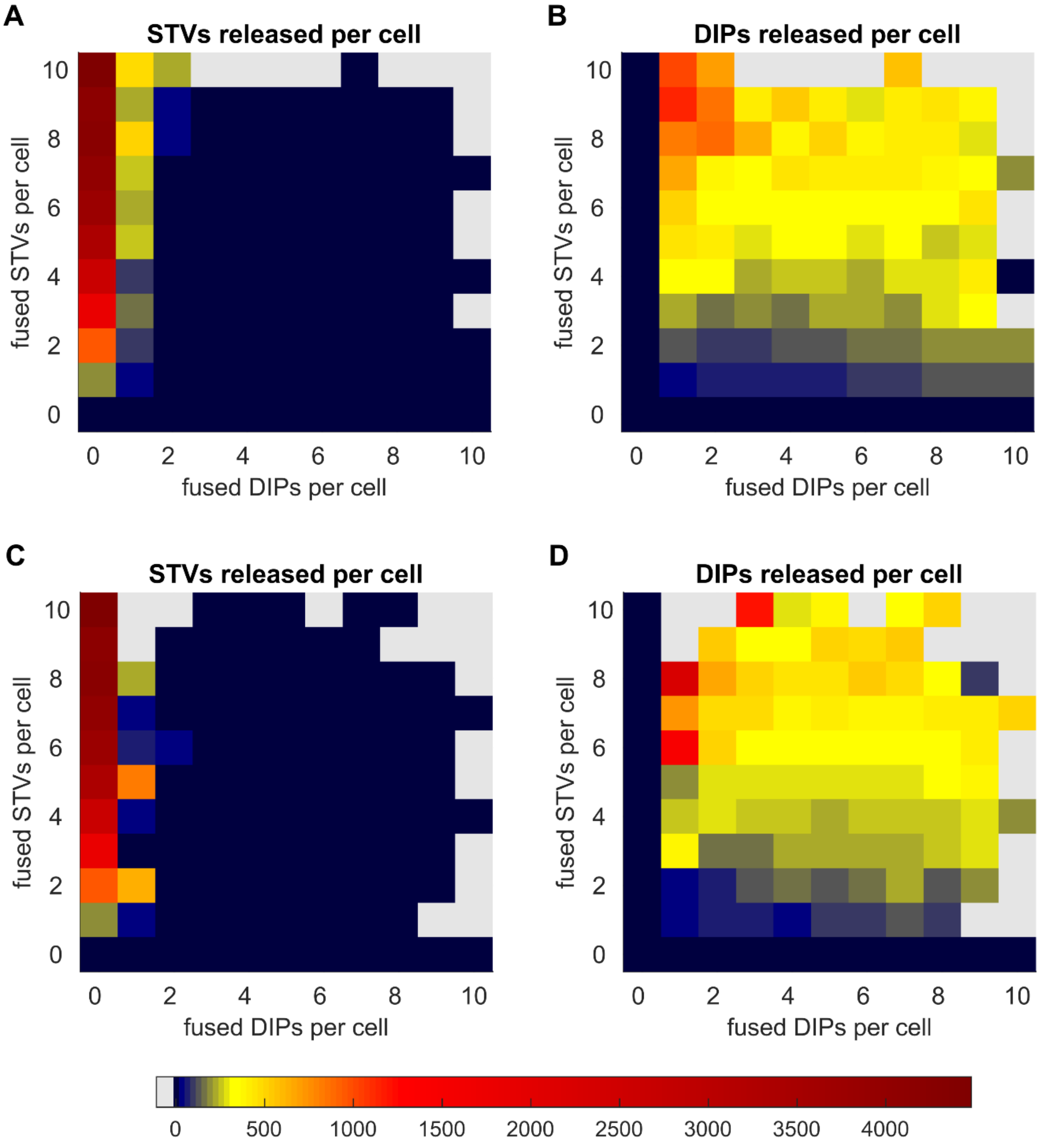


Figure 4.33 | STV and DIP release with different implementations of particle packaging. (A,B) Average STV (A) and DIP (B) production considering a complex formation rate proportional to the product of each vRNP abundance in the cytoplasm – *Complex Model*. (C,D) STV (C) and DIP (D) average production considering a complex formation rate proportional to the least abundant vRNP segment – *Minimum Packaging Model*. Results were obtained at 12 hpi and only simulations in which both STV and DIP fused successfully were considered.

When we compare both the *Complex Model* and the *Minimum Packaging Model* we do not observe significant differences. We can highlight that in the scenarios where 1 fusing DIP is combined with 2 and 5 STV entering the cell, the *Minimum Packaging Model* reveal a mean STV release higher than the mean DIP production (Figure 4.33C,D). Furthermore, in agreement with the *Complex Model,* the maximum average DIP yield in the *Minimum Packaging Model* is achieved when 1 DIP and 8 STV successfully fused (Figure 4.33D).

Since overall the average production was not affected by the new implemented complex formation rate, we wondered if we can observe some differences in the release dynamics. Furthermore, we want to investigate if this variation in the model has a bigger impact either on STV or DIP release. During this analysis we focused in the scenario that exactly one STV and one DIP fused successfully. We performed 1927 simulations with the *Minimum Packaging Model* at these initial infection conditions.

Our results reveal that the STV release dynamics is slightly lower for the *Minimum Packaging Model* (Figure 4.34A). Since the new complex formation rate is proportional to the least abundant vRNP this will slow down the packaging step which affects STV production. By contrast, the DIP release dynamics is very similar for the both models (Figure 4.34B). In summary, the complex formation rate implemented in the *Minimum Packaging Model* mainly impacts STV release, which indicates that in most simulations the FL segment 3 should be the least abundant segment in the cytoplasm.

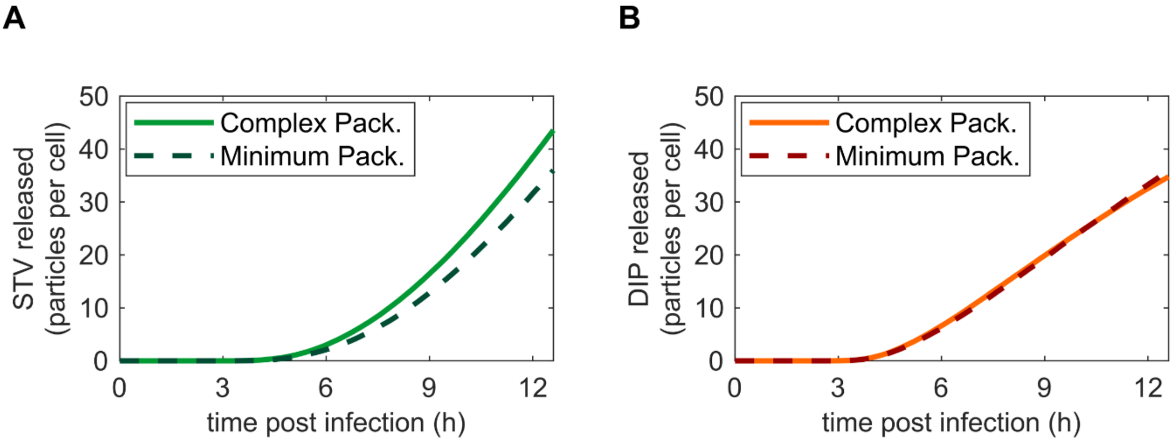


Figure 4.34 | Release dynamics using different packaging approaches. STV (A) and DIP (B) release dynamics over the course of infection. Only simulations in which both STV and DIP fused successfully were considered.

Furthermore, we calculated the STV and DIP complex formation and release rates based on the level of complexes and progeny produced between two consecutive time points. As it would be expected from what we observed previously, the STV complex formation and release rates are higher for the original *Complex Model* (Figure 4.35A,C). Surprisingly, we observed some differences in the two models regarding the complex formation and release rates for DIPs (Figure 4.35B,D). When the cells start to release the first particles after 3 hpi we can observe that both rates are lower for the *Minimum Packaging Model*. However, at approximately 6 hpi, the rates shift and during later infection stages the *Complex Model* has lower complex formation and release rates.

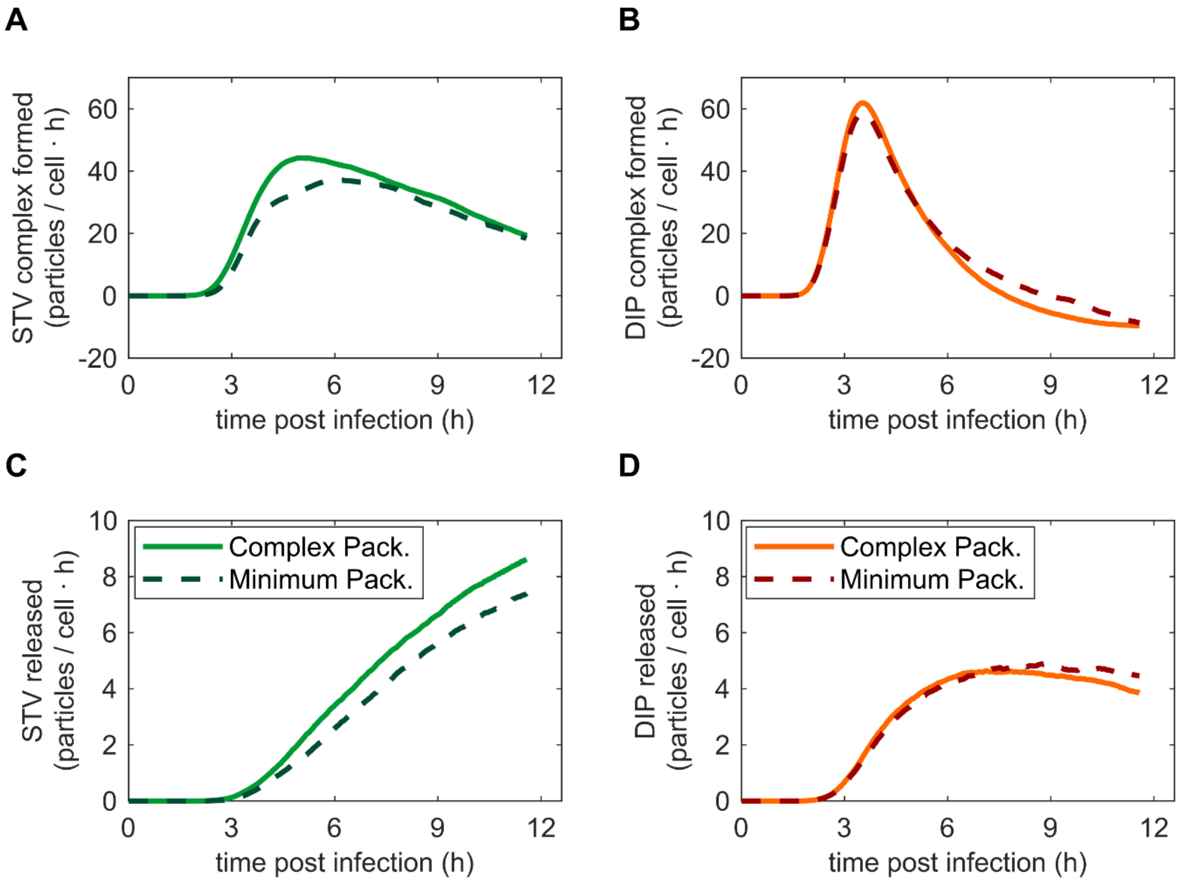


Figure 4.35 | Complex formation and release rates using different packaging approaches. (A,B) Complex formation rate of STV (A) and DIP (B). (C,D) Release rate of STV (C) and DIP (D). Only simulations in which both STV and DIP fused successfully were considered.

4.4.2 Replication model

In the previous section we evaluated how a different approach of implementing the packaging of vRNPs affects both STV and DIP release dynamics. In this section we focused on the DIP replication advantage as we were interested to compare the replication and packaging advantage. Therefore, we analysed if an increased advantage factor will benefit DIP propagation and how this impacts the level of different genome segment.

We modified the original model by excluding the complex formation step (Equations (3.35) and (3.36)). In the new model we assumed that the packaging of particles occurs in a single reaction step which assembles the eight vRNPs and all proteins necessary for particle release (Equations (3.51) and (3.52)). Although the packaging advantage was excluded in the modified model, DIP replication advantage is maintained due to the implemented advantage factor. We will compare the original model which includes both replication and packaging advantage (*Complex Model*) with the new model that comprises only the replication advantage (*Replication Model*).

Similarly, as in the previous section we used heat maps to compare the general release for each combination of STV and DIP fusion. The results show a clear difference between STV and DIP production when we compare the *Complex Model* (Figure 4.33A,B) and the *Replication Model* (Figure 4.36A,B). However, we can highlight some similarities with the original model, e.g. in the scenario that 1 DIP successfully fuses. Here an increasing number of fused STVs leads to an increased average DIP production for both models (Figure 4.36B). Although in this scenario, we observe that the STV production overcomes the DIP release in the *Replication Model* (Figure 4.36A).

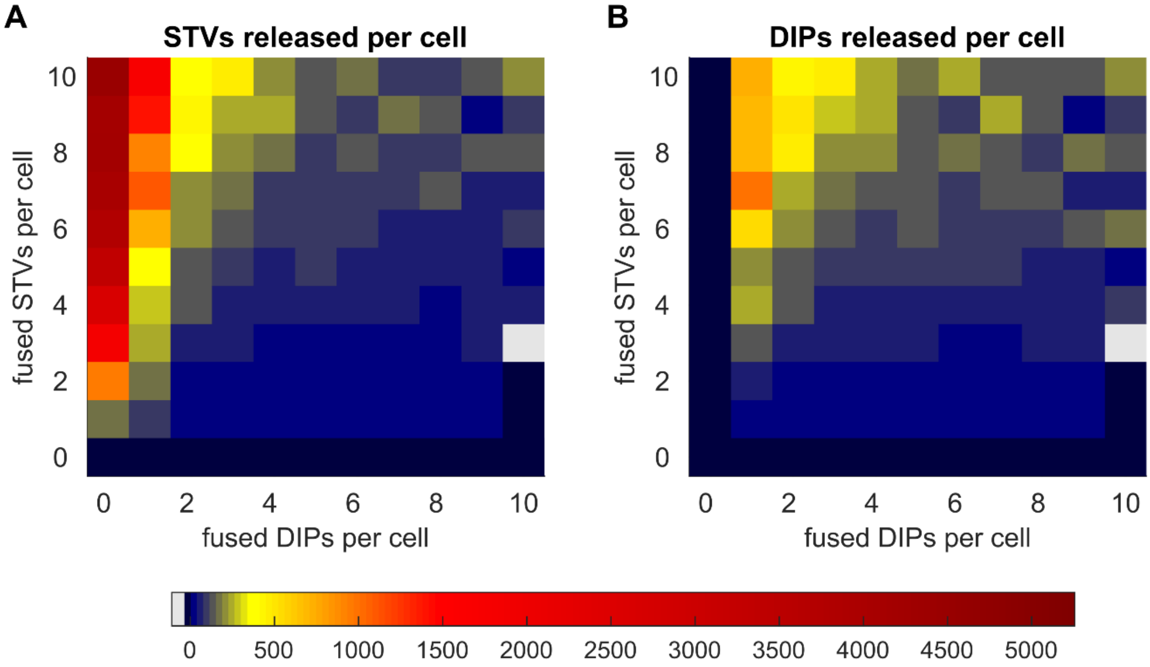


Figure 4.36 | STV and DIP release in the *Replication Model*. Average STV (A) and DIP (B) production considering an advantage factor of 3.65 and packaging occurring in a single step reaction which assembles the eight vRNPs and all proteins necessary for particle release. Results were obtained at 12 hpi and only simulations in which both STV and DIP fused successfully were considered.

Overall, we observed that the average STV and DIP production are very similar in the *Replication Model.* These data indicate that when the DIP advantage is only implemented at the replication level, this will not improve DIP release. Consequently, the DIP advantage should be implemented simultaneously at the replication and packaging level (with a two-step packaging process) to achieve a maximum DIP production and higher interference with the STV replication.

As described before, it has been suggested that DIP has a replication advantage over STV due to the shorter length of its DI segment that allows a faster synthesis and consequent accumulation of the defective segment. We used the *Replication Model* to evaluate how the length dependent advantage would impact STV and DIP replication. We performed simulations for different replication advantage factors ranging from 0 to 5 and calculated the average particle production and segment levels at 12 hpi. For this analysis we only considered simulations in which exactly one STV and one DIP fused successfully.

The results obtained for the average STV and DIP production confirm what we observed previously in the heat maps: when only the replication advantage is implemented, and the complex formation step is disregarded, the average production of STVs and DIPs is very similar (Figure 4.37A). Surprisingly, when there is no advantage factor $(F_{Adv} = 0)$, the average release of both particles is different. Furthermore, we observe a higher production when the advantage factor is lower than 1. By contrast, increased values of advantage impair production. It is important to notice that when the advantage factor is higher than 3 the average DIP release is close to zero, although STV still produces low amounts of progeny.

The factor advantage impacts directly the genome segment levels since it affects the cRNA synthesis rate. To understand how this advantage would impact the abundance of different segments abundance we calculated the average levels of different vRNPs inside the cytoplasm for each factor value. Our data show the FL segments 3 and 5 are higher than the DI segment when the advantage is lower than 1 (Figure 4.37B). However, in this regime the DI segment has an increasing trend and the FL segments are decreasing. By contrast, the DI segment level decreases with higher factor advantage values and the FL segments suffer a more accentuated reduction.

Furthermore, we used a 3.65 factor advantage in the simulations performed for the original *Complex Model*, which represents a disadvantageous scenario for DIP replication as indicated by the results obtained in this section. Overall, increased advantage factors up to 1 increase the level of DI segment which consequently induce a higher average DIP production. However, then the advantage factor is higher than 1, the results for the DI segments strongly fluctuate. This result suggests that probably the number of simulations performed was not sufficient to obtain a consistent result.

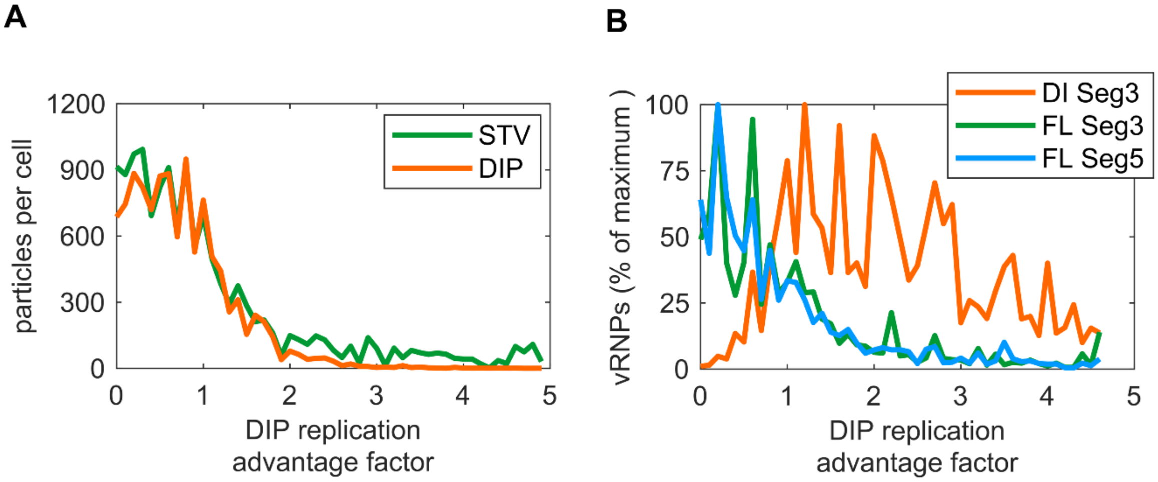


Figure 4.37 | Effect of the defective cRNA replication advantage on particle release and genome segment levels. (A) Average STV and DIP production for different factor advantages. (B) Average vRNP levels inside the cytoplasm depending on the applied advantage factor. Results were obtained at 12 hpi and only simulations in which both STV and DIP fused successfully were considered.

5 Discussion

In recent years, DIPs have become a topic of high interest for researchers due to their potential use for antiviral therapy that could support the conventional vaccination methods [13,14]. Due to the increased necessity of understanding the mechanism inherent to DIP interference, we implemented the replication of DIPs in a general stochastic model of IAV infection, which was developed by Heldt & Dorl [3,67]. Our model uses the model adaptation approach previously applied in the deterministic model of DIP replication, which was implemented by Laske *et al*. [9]. We will compare and discuss our results in comparison to the deterministic DIP model in this section. To our knowledge, this is the first time that the intracellular IAV replication in the presence of DIPs is described with a stochastic model.

At MOI/MODIP 10/10 due to high number of molecules the stochastic effects should be negligible, as a result the average stochastic intracellular dynamics is very similar to the deterministic model. We would expect this small deviations at high initial infection condition since differences were detected as well when the stochastic model of IAV was compared with the deterministic approach [1]. However, significant differences emerge regarding particle release. STV production distribution is skewed to the right towards low productivity and 60% of simulations do not produce progeny virions which confirms DIP interference in STV replication. Experimental data [7,23] support our model results since most of our simulations showed high DIP production (99% of simulations showed between 1 and 1500 DIPs released per cell) and impaired STV release (almost 40% of the cells produced between 1 and 15 STVs per cell). The average DIP release dynamics is lower than the deterministic result, although the average DI vRNA dynamics is slightly higher when compared to the previous deterministic model of DIP replication (Figure 4.2B,C). In theory, a higher level of available defective segment should result in an increased DIP release. However, this does not occur in our simulations since we defined that all genome segments are independent from one another. This assumption was implemented because IAV has a segmented genome and consequently each vRNA replicates independently. Therefore, vRNA levels inside a cell can vary significantly [67]. By contrast, in the deterministic model vRNA segments replicate simultaneously maintaining similar levels. In our simulations the DI segment spans up to 4 orders of magnitude and the FL segments range over 6 orders of magnitude (Figure 4.4A,B). Consequently, even though DI segment is in high abundance, the absence or low level of one FL segment necessary for DIP replication can impair its release. Hence, the segmented genome of IAV can be the major source of the variability observed in STV and DIP release at high initial infection conditions.

The simulations performed at MOI/MODIP 1/1 showed an even higher deviation from the deterministic result. Specifically, in 94% and 85% of simulations cells did not produce STV or DIP, respectively. This supports the assumption that stochastic effects, which prevents cells from producing progeny, are more pronounced at low initial infection conditions [67]. The major factor which increases the number of non-productive cells at low MOI/MODIP in our model is the occurrence of failed particle fusions (Figure 4.16 and Figure 4.17). The deterministic model does not account for such events since it considers that all variables are continuous: when one STV or DIP enter the cell, a fraction of the particle successfully fuses and induces progeny production while another fraction fails to fuse, consequently complete failed fusion is not possible. By contrast, the stochastic model approaches the system as discrete, so each particle has a chance of successfully entering the cell (roughly ~50% in our model). Since in nature one complete viral particle is necessary to infect a cell, the stochastic model provides a more realistic representation and describes four possible fusion scenarios: STV fuses alone, DIP fuses alone, both fuse or both fail fusing, each event occurring with a ~25% probability in a MOI/MODIP 1/1 scenario. Stochastic simulations result in a mixture of these events which contribute to the observed cell-to-cell variability. These events determine the outcome of the system and explain the increase in the percentage of simulations that do not produce DIPs that was not observed previously at MOI/MODIP 10/10. At low initial infection conditions, a DIP will not replicate in the three outlined fusion failure scenarios since even if it fuses alone, it does not produce progeny due to the lack of resources which can only be provided by the STV [64]. The increased percentage of cells that do not produce STV at low initial infection conditions when compared with the MOI/MODIP 10/10 scenario is less pronounced (from 60% to 98% of non-productive cells) since it has the possibility to replicate in one of the three scenarios: when DIPs fail to fuse, the STV can propagate without interference.

Another essential event that prevents cells from producing progeny particles and is not considered in the deterministic model is segment loss. vRNP segments inside the nucleus can be degraded by nucleases before starting replication [65]. In the deterministic model a fraction of the vRNP segment is degraded, however part of this genome segment can replicate. By contrast, in the stochastic model one degradation event can determine the end of the replication process. Depending in which segment is lost, different scenarios can occur in at MOI/MODIP 1/1: either the DI segment is degraded and only STV replicates or its corresponding FL segment is lost and just DIP produces progeny. In theory, if one of the other FL segments is degraded, this would prevent both STV and DIP propagation since both require the incorporation of the other functional segments in their released particles. However, in multiple simulations we observed the release production of exactly one progeny particle despite the loss of a genome segment necessary for release. We suggest that these events occurred due to the respective vRNP being exported from the nucleus before it could replicate, i.e. the RNP segments which arrive at the nucleus could still be used for the formation of a single progeny particle. Obviously, this event can only occur if M1 and NEP was already synthesised and available to perform the nuclear export. Hence, the release of exactly one progeny particle can be considered as a re-packaging of the incoming vRNPs when segment loss occurs or release of exactly one particle when no segment was lost.

Fusion failure and segment loss are two events that affect the success of particle replication. These factors, combined with the segment independent replication, can induce fluctuations in the vRNA levels which at MOI/MODIP 1/1 span over 8 orders of magnitude (Figure 4.9). The differences in vRNA levels also impact protein synthesis. When vRNP enters the nucleus, both mRNA and cRNA can be synthesized. Since the mRNA synthesis rate is considerably higher than the cRNA degradation rate, primary transcription can occur before the loss of a vRNA segment. Consequently, small amounts of mRNA are produced that can synthesise low protein levels resulting in a wide-spread distribution of viral proteins [67]. At MOI/MODIP 1/1 the average stochastic RdRp and NP dynamics are considerably higher than in the deterministic model (Figure 4.10B). Specifically, in the deterministic model the viral polymerase and NP suffer an abrupt reduction at 3 hpi which is caused by a protein limitation due to DIP interference. This limitation was not observed in our average stochastic dynamics due to existence of different protein subpopulations (Figure 4.11). Specifically, the RdRp cluster observed at low protein levels could be caused by the loss of one of the genome segments that encodes for the subunits of the polymerase complex. Furthermore, since FL segment 3, which encodes for RdRp, is in low levels inside the cell, this might also result in the reduction of polymerase synthesis and establish subpopulations with low protein levels [62]. However, the occurrence of low NP levels could be caused either by the loss of segment 5 or substantially increased replication of DI vRNA which would require more NP for stabilization. The NP and DIP production distributions at MOI/MODIP 10/10 support this assumption since we observed that simulations with low levels of NP at 12 hpi showed an increased DIP production (Figure 4.6B). Furthermore, simulations in which more than 700 DIPs per cell were released exhibited a similar trend in NP dynamics as observed in the deterministic result. This result was not shown, because at MOI/MODIP 1/1 we obtained only four simulations that showed such high DIP yield and their average would be not representative for a comparison with the deterministic result. Overall, our model simulations suggest that, besides the regular DIP and STV production described in the deterministic model, different scenarios regarding DIP and STV replication emerge. These include the fusion failure of the STV or DIP particle and the loss of any of the genome segments. These events combined with the assumption of an independent segment replication cause large fluctuations in the vRNA levels which will strongly affect protein levels and particle release. NP is especially affected by these events which results in a wide distribution even at high initial infection conditions.

As explained previously, in a MOI/MODIP 1/1 scenario stochastic effects are inducing a high percentage of non-productive cells which increases the differences observed between the stochastic and deterministic model results. To investigate how different combinations of MOI/MODIP would impact DIP production, we tested different combinations. Therefore, we fixed the MODIP at 1 and increased the MOI. Our simulation results reveal that a higher number of initial STV particles infecting a single cell increases the average DIP release and the differences between the stochastic average result and the deterministic model are reduced with increasing MOI (Figure 4.13). Furthermore, the standard deviation of DIP release is higher in the MOI/MODIP 10/1 scenario. This supports the concept that higher MOI values not only increase the DIP average production, but also its potential for releasing more progeny particles. We suggest that a higher number of initial STV benefits DIP production due to increased resources availability, i.e. proteins and FL vRNAs, which are essential for DIP propagation. Additionally, we evaluated how DIP production is affected when we fix the MOI to 1 and increase the MODIP. The stochastic model shows that when MODIP is increased, the average DIP release increases although the standard deviation does not vary significantly (Figure 4.14). However, experimental studies of IAV infection showed that high amounts of initially provided DIPs result not only in a reduction in virus titer, but also the total particle concentration decreases [76]. Furthermore, experiments performed in VSV showed that increasing the MODIP decreases the overall DIP release [77]. Hence, we suggest that increased MODIP values do not improve DIP production but rather that the increased DIP release we observe is caused by the reduction of non-productive simulations which increases the average DIP release. To further examine this effect, we calculated the ratio of DI segment 3 to its complementary FL segment ratio. We observed that lower ratios, i.e. higher levels of FL vRNP segment 3 (and others FL segments) compared to the DI segment 3, benefit DIP production since this indicates the presence of more FL segments that would increase the availability of resources for DIP release. Furthermore, higher levels of DI vRNP seem to result in a disadvantageous self-inference. This might occur due to exhaustion of essential proteins, since more DI segments would require stabilization by NP. Finally, to evaluate how the levels of FL segment impact STV production, we calculated the ratio of the average of the individual vRNP segment levels in a cell over the FL vRNP segment 3. Our results reveal that STV production occurs optimally when the ratio is around 1. This indicates that a balanced replication of all functional segments is required for high STV production. This requirement is challenged by the independent segment replication which can lead to unbalanced replication and impair both STV and DIP production.

The stochastic model considers STV and DIP as two independent particles. At the start of the simulation, each particle is transported independently in the cytoplasm. As transport is a random process, each particle can reach the nucleus at different time points. Different delays of nuclear import can also be simulated with the deterministic model by introducing each particle at a different time point. However, when both are placed outside the cell at the same time, STV and DIP are transported together and reach the nucleus simultaneously. The time delay events observed in stochastic simulation are one factor that increases the cell-to-cell variability observed in our results. To analyse the time delay we performed simulations at MOI/MODIP 1/1 but only considered for our analysis simulations in which both particles fused successfully. Delays below 1.5 hours were analysed based on our regular simulation results in which the differences in time of nuclear import occurred due to the random nature of particle transport. Delays from 2 to 4 hours were implemented by inducing the delay. However, the average results for delays below 1.5 hours, which occurred in our regular simulation, might not be fully representative for each scenario. This is caused by the reduced number of simulations performed for an individual scenario since the time delays showed a normal distribution with mean approximately zero. Therefore, time delays around zero hours occurred more frequently and their average results are more reliable than for delay around 1.5 hours. Altogether, our data show that when the DIP entry is delayed more than 3 hours, STV can start its replication without interference which constitutes an advantage for STV production. By contrast, when STV entry is delayed its replication begins later which increases the number of cells not producing STV. The average DIP production is increased when DIP entry is delayed up to 2 hours. This data supports the hypothesis that DIP production is improved when

STV enters first since STV starts the synthesis of the proteins earlier, which will increase resource availability for DIP production. However, if the STV replication is advanced by a too large margin, DIP propagation will be impaired. The maximum DIP release is achieved when its entry is delayed between 1.5 and 2 hours. However, this range is right at the limit between the randomly generated and induced delay and the average calculated at 1.5 hours might not be representative due to the low number of simulations performed. Therefore, we suggest that the optimum average DIP production is obtained when DIP entry is delayed between 1.5 and 3 hours. This is in agreement with experiments which show that DIP interference is only observed if DIP entry is delayed no longer than 3 hours [78]. In addition, we also analysed how the delay of nuclear import would affect DIPs carrying a defective segment 4. A defective segment 4 allows DIP to replicate and increase its genome segments levels on its own. The simulation of this scenario with the deterministic model results in a prevention of STV replication if STV entry is delayed more than 3 hours and, consequently, the DIP would lack necessary proteins and not release any progeny. However, this complete shutdown of DIP release was not observed in our simulations. We assume that, such a complete shutdown of DIP production could also be observed in the stochastic model or delays higher than 4 hours, which should be investigated in the future.

Finally, we analysed different approaches of DIP advantage implementation at the replication and packaging steps. In the original model the packaging process comprises two steps: first all eight vRNP genome segments get together forming a complex [57] and then both STV and DIP complexes compete for proteins to incorporate in their release. The complex formation rate is proportional to the product of each vRNP levels in the cytoplasm which leads to an almost instantaneous formation of the complexes. We consider that this considerably high rate (in the order of magnitude of 10^{20} molecules$\cdot$cell$^{-1}\cdot$h^{-1}) is unrealistic due to transport factors that can affect the complex establishment. We modified this assumption by implementing a complex formation rate that depends on the least abundant vRNA which reduces the rate of formation. Overall, our results show that this modification reduced the average STV release dynamics. We suggest that this occur dur to low levels of FL segment 3 in the cell, which constitutes a limiting factor for STV production. By contrast, the model modification did not impact DIP production since the DI segment is usually present in high levels due to the DIP replication advantage. However, we observed that the DIP complex formation and release rates in the modified model is higher at 6 hpi than the original model (Figure 4.35B,D). This interesting dynamic may occur since at the beginning there are few vRNPs in the cytoplasm, so the complex formation rate is very low, and few particles are released from the cell. Consequently, vRNPs will accumulate inside the cell and later in infection the complex formation and release rate will be higher due to increased number of available genome segments.

The last model variation was implemented by discarding the complex formation step assuming that the assembly of the eight vRNPs and all proteins necessary for packaging occurs in a single step reaction. Our results revealed that by implementing only the replication advantage, the DIP and STV production are, on average, very similar. These results highlight the importance of the implemented complex formation step in the model to achieve a substantial DIP advantage and inhibition of STV production. Furthermore, we performed simulations with increased replication advantage factor values.

Our results showed that increasing DIP advantage results in a reduction of the FL segments levels. By contrast, the DI segment abundance increases with advantage factors up to 1 which improves DIP production. For advantage factors above one the DI segment levels decrease which impairs DIP release. The previously implemented deterministic model of DIP coinfection also showed a decrease in DIP production with an increasing replication advantage factor. In the deterministic simulation this effect was related to the depletion of unbound NP. The decrease of FL segment 5 (and others FL segments) results in a reduction of NP synthesis and consequently the lack of NP to stabilize vRNA impairs the overall production [9]. In summary, the optimal DIP production is achieved when the replication advantage factor is below than 1. Furthermore, DIP release is impaired when this factor is too high since this will cause an unbalanced replication of FL segments which will eventually decrease protein availability.

In conclusion, we implemented a stochastic intracellular model of influenza A virus and DIP replication which highlights the importance of random events during the infection and their effect on virus release. In contrast to the deterministic model, the stochastic implementation accounts for the independent replication of viral genome segments, the failure of particle fusion, the loss of genome segments and the time delay of nucleus entry. Our model shows that these events have a major impact during coinfections of STVs and DIPs. However, other sources of cell-to-cell variability highly observed in nature, e.g. host factors, immune response or mutations were not considered in our model. Is important to note that our model still requires experimental data for validation, however, this is a challenge as the experimental methods for quantification are still in development. Furthermore, the stochastic simulations performed with our model are highly computationally demanding due to the high number of interacting molecules. In addition, a high number of simulations should be performed to achieve representative averages of a cell populations. We established to perform more than 1000 simulations for each condition since we observed that beyond this number of runs we obtained a stable average. However, when filtering the results to analyse specific aspects and subpopulations among the obtained data, the amount of suitable simulations can be quite low. This constitutes a limitation for our modelling results.

6 Conclusion

We developed an intracellular stochastic model of IAV and DIP replication with the aim to study how different stochastic effects impact DIP production. We could demonstrate that the average stochastic dynamics significantly differs from the previous implemented deterministic model. Specifically, at MOI/MODIP 1/1 most of the simulations are non-productive as a result of random events such as fusion failure and the loss of genome segments. Simulations performed at high MOI/MODIP conditions still show wide-spread distributions as a consequence of the implemented independent replication of IAV genome segments. Moreover, our simulations are susceptible to different time delays between the nuclear import of DIPs and STVs which further increases the heterogeneity of progeny production. In addition, these stochastic effects especially impact levels of NP which resulting in the emergence of three sub-populations in its distribution, which also strongly affect DIP production.

The stochastic model we implemented can be used to conduct a comprehensive study of DIPs originating from different genome segments in the future. Furthermore, it would be interesting to study the competition of different DIPs to identify if a specific DI segment will emerge as a dominant species implying that it may be a potent antiviral agent. However, to simulate this scenario a model expansion from the intracellular to the population level is necessary. In addition, a more detailed analysis of the NP sub-populations could further support the importance of its role in DIP replication. The stochastic model could also be extended by the implementation of DIP *de novo* generation to investigate how this process could impact vaccine production and DIP interference.

Bibliography

---

[1] Heldt, F.S. (2014) Mathematical models of influenza A virus infection: from intracellular replication to virus growth in cell populations. *PhD Dissertation, Faculty of Process and Systems Engineering, Otto von Guericke University, Magdeburg*.

[2] Bouvier, N.M. and Palese, P. (2008) The biology of influenza viruses. *Vaccine*. 26 (Suppl 4), D49–D53.

[3] Dorl, S. (2013) Stochastic modelling if the intracellular replication of influenza virus during vaccine production. *Bachelor Thesis, Faculty of Process and Systems Engineering, Otto von Guericke University, Magdeburg*.

[4] Aggarwal, K., Jing, F., Maranga, L., and Liu, J. (2011) Bioprocess optimization for cell culture based influenza vaccine production. *Vaccine*. 29 (17), 3320–3328.

[5] Horimoto, T. and Kawaoka, Y. (2005) Influenza: lessons from past pandemics, warnings from current incidents. *Nature Reviews Microbiology*. 3 (8), 591–600.

[6] Simonsen, L., Spreeuwenberg, P., Lustig, R., Taylor, R.J., Fleming, D.M., Kroneman, M., et al. (2013) Global mortality estimates for the 2009 influenza pandemic from the GLaMOR project: a modeling study. *PLoS Medicine*. 10 (11),.

[7] Duhaut, S.D. and McCauley, J.W. (1996) Defective RNAs inhibit the assembly of influenza virus genome segments in a segment-specific manner. *Virology*. 216 (2), 326–337.

[8] Frensing, T., Heldt, F.S., Pflugmacher, A., Behrendt, I., Jordan, I., Flockerzi, D., et al. (2013) Continuous influenza virus production in cell culture shows a periodic accumulation of defective interfering particles. *PLoS One*. 8 (9),.

[9] Laske, T., Heldt, F.S., Hoffmann, H., Frensing, T., and Reichl, U. (2016) Modeling the intracellular replication of influenza A virus in the presence of defective interfering RNAs. *Virus Research*. 213 90–99.

[10] Odagiri, T. and Tashiro, M. (1997) Segment-specific noncoding sequences of the influenza virus genome RNA are involved in the specific competition between defective interfering RNA

and its progenitor RNA segment at the virion assembly step. *Journal of Virology*. 71 (3), 2138–2145.

[11] Frensing, T. (2015) Defective interfering viruses and their impact on vaccines and viral vectors. *Biotechnology Journal*. 10 (5), 681–689.

[12] Nayak, D.P., Chambers, T.M., and Akkina, R.K. (1985) Defective-interfering (DI) RNAs of influenza viruses: origin, structure, expression, and interference. *Current Topics in Microbiology and Immunology*. 114 103–151.

[13] Marriot, A.C. and Dimmock, N.J. (2010) Defective interfering viruses and their potential as antiviral agents. *Reviews in Medical Virology*. 20 51–62.

[14] Dimmock, N.J. and Easton, A.J. (2014) Defective Interfering Influenza Virus RNAs: Time To Reevaluate Their Clinical Potential as Broad-Spectrum Antivirals? *Journal of Virology*. 88 (10), 5217–5227.

[15] Dimmock, N.J., Rainsford, E.W., Scott, P.D., and Marriott, A.C. (2008) Influenza Virus Protecting RNA: an Effective Prophylactic and Therapeutic Antiviral. *Journal of Virology*. 82 (17), 8570–8578.

[16] Dimmock, N.J., Dove, B.K., Scott, P.D., Meng, B., Taylor, I., Cheung, L., et al. (2012) Cloned Defective Interfering Influenza Virus Protects Ferrets from Pandemic 2009 Influenza A Virus and Allows Protective Immunity to Be Established. *PLoS ONE*. 7 (12),.

[17] Beauchemin, C. a a and Handel, A. (2011) A review of mathematical models of influenza A infections within a host or cell culture: lessons learned and challenges ahead. *BMC Public Health*. 11 Suppl 1 (Suppl 1), S7.

[18] Ullah, M. and Wolkenhauer, O. (2010) Stochastic approaches in systems biology. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*. 2 (4), 385–397.

[19] Raj, A. and van Oudenaarden, A. (2008) Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*. 135 (2), 216–226.

[20] Gillespie, D.T. (2007) Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*. 58 (1), 35–55.

[21] von Magnus, P. (1954) Incomplete forms of influenza virus. *Adv Virus Res*. 2 (11), 59–79.

[22] Huang, A.S. and Baltimore, D. (1970) Defective viral particles and viral disease processes. *Nature*. 226 325–327.

[23] Duhaut, S.D. and Dimmock, N.J. (2002) Defective segment 1 RNAs that interfere with production of infectious influenza A virus require at least 150 nucleotides of 5′ sequence: Evidence from a plasmid-driven system. *Journal of General Virology*. 83 (2), 403–411.

[24] Calder, L.J., Wasilewski, S., Berriman, J.A., and Rosenthal, P.B. (2010) Structural

organization of a filamentous influenza A virus. *Proceedings of the National Academy of Sciences*. 107 (23), 10685–10690.

[25] Fodor, E. (2013) The RNA polymerase of influenza A virus: mechanism of viral transcription and replication. *Acta Virologica*. 57 (4), 113–122.

[26] Zheng, W. and Tao, Y.J. (2013) Structure and assembly of the influenza A virus ribonucleoprotein complex. *FEBS Letters*. 587 (8), 1206–1214.

[27] Noda, T., Sagara, H., Yen, A., Takada, A., Kida, H., Cheng, R.H., et al. (2006) Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature*. 439 (7075), 490–492.

[28] Huang, T., Palese, P., and Krystal, M. (1990) Determination of influenza virus proteins required for genome replication. *Journal of Virology*. 64 (11), 5669–5673.

[29] Vasin, A. V., Temkina, O.A., Egorov, V. V., Klotchenko, S.A., Plotnikova, M.A., and Kiselev, O.I. (2014) Molecular mechanisms enhancing the proteome of influenza A viruses: an overview of recently discovered proteins. *Virus Research*. 185 53–63.

[30] Jennings, P.A., Finch, J.T., Winter, G., and Robertson, J.S. (1983) Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? *Cell*. 34 (2), 619–627.

[31] Davis, A.R. and Nayak, D.P. (1979) Sequence relationships among defective interfering influenza viral RNAs. *Proceedings of the National Academy of Sciences of the United States of America*. 76 (7), 3092–3096.

[32] Hutchinson, E.C., von Kirchbach, J.C., Gog, J.R., and Digard, P. (2010) Genome packaging in influenza A virus. *Journal of General Virology*. 91 (2), 313–328.

[33] Heldt, F.S., Frensing, T., and Reichl, U. (2012) Modeling the intracellular dynamics of influenza virus replication to understand the control of viral RNA synthesis. *Journal of Virology*. 86 (15), 7806–7817.

[34] Skehel, J.J. and D.C., W. (2000) Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annual Reviews in Biochemistry*. 69 531–569.

[35] Nunes-Correia, I., Ramalho-Santos, J., Nir, S., and Pedroso De Lima, M.C. (1999) Interactions of influenza virus with cultured cells: detailed kinetic modeling of binding and endocytosis. *Biochemistry*. 38 (3), 1095–1101.

[36] Chou, Y., Heaton, N.S., Gao, Q., Palese, P., Singer, R., and Lionnet, T. (2013) Colocalization of different influenza viral RNA segments in the cytoplasm before viral budding as shown by single-molecule sensitivity FISH analysis. *PLoS Pathogens*. 9 (5),.

[37] Boulo, S., Akarsu, H., Ruigrok, R.W.H., and Baudin, F. (2007) Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes. *Virus Research*. 124 12–21.

[38] Hutchinson, E.C. and Fodor, E. (2012) Nuclear import of the influenza A virus transcriptional machinery. *Vaccine*. 30 (51), 7353–7358.

[39] Portela, A. and Digard, P. (2002) The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. *Journal of General Virology*. 83 (4), 723–734.

[40] Honda, A., Uéda, K., Nagata, K., and Ishihama, A. (1988) RNA polymerase of influenza virus: role of NP in RNA chain elongation. *Journal of Biochemistry*. 104 (6), 1021–1026.

[41] Medcalf, L., Poole, E., Elton, D., and Digard, P. (1999) Temperature-sensitive lesions in two influenza A viruses defective for replicative transcription disrupt RNA binding by the nucleoprotein. *Journal of Virology*. 73 (9), 7349–7356.

[42] Mullin, A.E., Dalton, R.M., Amorim, M.J., Elton, D., and Digard, P. (2004) Increased amounts of the influenza virus nucleoprotein do not promote higher levels of viral genome replication. *Journal of General Virology*. 85 (12), 3689–3698.

[43] Vreede, F.T., Jung, T.E., and Brownlee, G.G. (2004) Model suggesting that replication of influenza virus is regulated by stabilization of replicative intermediates. *Journal of Virology*. 78 (17), 9568–9572.

[44] Vreede, F.T. and Brownlee, G.G. (2007) Influenza virion-derived viral ribonucleoproteins synthesize both mRNA and cRNA in vitro. *Journal of Virology*. 81 (5), 2196–2204.

[45] Jorba, N., Coloma, R., and Ortín, J. (2009) Genetic trans-complementation establishes a new model for influenza virus RNA transcription and replication. *PLoS Pathogens*. 5 (5),.

[46] Hatada, E. and Hasegawa, M. (1989) Control of influenza virus gene expression: quantitative analysis of each viral RNA species in infected cells. *Journal of Biochemistry*. 105 (4), 537–546.

[47] Yángüez, E. and Nieto, A. (2011) So similar, yet so different: selective translation of capped and polyadenylated viral mRNAs in the influenza virus infected cell. *Virus Research*. 156 1–12.

[48] Nayak, D.P. and Hui, E.K.W. (2002) Assembly and morphogenesis of influenza viruses. *Recent Research Developments in Virology*. 4 35–54.

[49] Baudin, F., Petit, I., Weissenhorn, W., and Ruigrok, R.W.H. (2001) In vitro dissection of the membrane and RNP binding activities of influenza virus M1 protein. *Virology*. 281 (1), 102–108.

[50] Perez, D.R. and Donis, R.O. (1998) The matrix 1 protein of influenza A virus inhibits the transcriptase activity of a model influenza reporter genome in Vivo. *Virology*. 249 (1), 52–61.

[51] Wakefield, L. and Brownlee, G.G. (1989) RNA-binding properties of influenza A virus matrix protein M1. *Nucleic Acids Research*. 17 8569–8580.

[52] Martin, K. and Heleniust, A. (1991) Nuclear transport of influenza virus ribonucleoproteins: The viral matrix protein (M1) promotes export and inhibits import. *Cell*. 67 (1), 117–130.

[53] Bui, M., Wills, E.G., Helenius, A., and Whittaker, G.R. (2000) Role of the influenza virus M1 protein in nuclear export of viral ribonucleoproteins. *Journal of Virology*. 74 (4), 1781–1786.

[54] O'Neill, R.E., Talon, J., and Palese, P. (1998) The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. *EMBO Journal*. 17 (1), 288–296.

[55] Huang, X., Liu, T., Muller, J., Levandowski, R.A., and Ye, Z. (2001) Effect of influenza virus matrix protein and viral RNA on ribonucleoprotein formation and nuclear export. *Virology*. 287 (2), 405–416.

[56] Noda, T. and Kawaoka, Y. (2010) Structure of influenza virus ribonucleoprotein complexes and their packaging into virions. *Reviews in Medical Virology*. 20 (6), 380–391.

[57] Fournier, E., Moules, V., Essere, B., Paillart, J.C., Sirbat, J.D., Isel, C., et al. (2012) A supramolecular assembly formed by influenza A virus genomic RNA segments. *Nucleic Acids Research*. 40 (5), 2197–2209.

[58] Fournier, E., Moules, V., Essere, B., Paillart, J.C., Sirbat, J.D., Cavalier, A., et al. (2012) Interaction network linking the human H3N2 influenza A virus genomic RNA segments. *Vaccine*. 30 (51), 7359–7367.

[59] Gavazzi, C., Yver, M., Isel, C., Smyth, R.P., Rosa-Calatrava, M., Lina, B., et al. (2013) A functional sequence-specific interaction between influenza A virus genomic RNA segments. *Proceedings of the National Academy of Sciences*. 110 (41), 16604–16609.

[60] Ali, A., Avalos, R.T., Ponimaskin, E., and Nayak, D.P. (2000) Influenza virus assembly: effect of influenza virus glycoproteins on the membrane association of M1 protein. *Journal of Virology*. 74 (18), 8709–8719.

[61] Rossman, J., Jing, X., Leser, G., and Lamb, R. (2010) Influenza virus M2 protein mediates ESCRT-independant membrane scission. *Cell*. 142 (6), 902–913.

[62] Akkina, R.K., Chambers, T.M., and Nayak, D.P. (1984) Mechanism of interference by defective-interfering particles of influenza virus: Differential reduction of intracellular synthesis of specific polymerase proteins. *Virus Research*. 1 (8), 687–702.

[63] Odagiri, T., Tominaga, K., Tobita, K., and Ohta, S. (1994) An amino acid change in the non-structural NS2 protein of an influenza A virus mutant is responsible for the generation of defective interfering (DI) particles by amplifying DI RNAs and suppressing complementary RNA synthesis. *Journal of General Virology*. 75 (1), 43–53.

[64] Widjaja, I., de Vries, E., Rottier, P.J.M., and de Haan, C.A.M. (2012) Competition between Influenza A virus genome segments. *PLoS ONE*. 7 (10),.

[65] Srivastava, R., You, L., Summers, J., and Yin, J. (2002) Stochastic vs. determninistic modeling of intracellular viral kinetics. *Journal of Theoretical Biology*. 218 309–321.

[66] Frazier, J.M., Chushak, Y., and Foy, B. (2009) Stochastic simulation and analysis of biomolecular reaction networks. *BMC Systems Biology*. 3 1–21.

[67] Heldt, F.S., Kupke, S.Y., Dorl, S., Reichl, U., and Frensing, T. (2015) Single-cell analysis and stochastic modelling unveil large cell-to-cell variability in influenza A virus infection. *Nature Communications*. 6 (January), 1–12.

[68] McQuarri, D.A. (1967) Stochastic approach to chemical kinetics. *Journal of Applied Probability*. 4 (3), 413–478.

[69] Gillespie, D.T. (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*. 22 (4), 403–434.

[70] Gillespie, D.T. (1977) Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*. 81 (25), 2340–2361.

[71] Banks, H.T., Hu, S., Joyner, M., Broido, A., Canter, B., Gayvert, K., et al. (2012) A comparison of computational efficiencies of stochastic algorithms in terms of two infection models. *Mathematical Biosciences and Engineering*. 9 (3), 487–526.

[72] Gillespie, D.T. (2001) Approximate accelerated stochastic simulation of chemically reacting systems. *Journal of Chemical Physics*. 115 (4), 1716–1733.

[73] Cao, Y., Gillespie, D.T., and Petzold, L.R. (2006) Efficient step size selection for the tau-leaping simulation method. *Journal of Chemical Physics*. 124 (4),.

[74] Stegmann, T., Schoen, P., Bron, R., Wey, J., Bartoldus, I., Ortiz, A., et al. (1993) Evaluation of viral membrane fusion assays. Comparison of the octadecylrhodamine dequenching assay with the pyrene excimer assay. *Biochemistry*. 32 (42), 11330–11337.

[75] Enami, M., Fukuda, R., and Ishihama, A. (1985) Transcription and replication of eight RNA segments of influenza virus. *Virology*. 142 68–77.

[76] Frensing, T., Pflugmacher, A., Bachmann, M., Peschel, B., and Reichl, U. (2014) Impact of defective interfering particles on virus replication and antiviral host response in cell culture-based influenza vaccine production. *Applied Microbiology and Biotechnology*. 98 (21), 8999–9008.

[77] Sekellick, M.J. and Marcus, P.I. (1980) Viral interference by defective particles of vesicular stomatitis virus measured in individual cells. *Virology*. 104 247–252.

[78] Nayak, D.P., Tobita, K., Janda, J.M., Davis, a R., and De, B.K. (1978) Homologous interference mediated by defective interfering influenza virus derived from a temperature-sensitive mutant of influenza virus. *Journal of Virology*. 28 (1), 375–386.

[79] Arava, Y., Wang, Y., Storey, J.D., Liu, C.L., Brown, P.O., and Herschlag, D. (2003) Genome-wide analysis of mRNA translation profiles in Saccharomyces cerevisiae. *Proceedings of the*

National Academy of Sciences of the United States of America. 100 (7), 3889–3894.

[80] Robb, N.C., Jackson, D., Vreede, F.T., and Fodor, E. (2010) Splicing of influenza A virus NS1 mRNA is independent of the viral NS1 protein. *Journal of General Virology*. 91 (9), 2331–2340.

[81] Amorim, M.J., Bruce, E.A., Read, E.K.C., Foeglein, A., Mahen, R., Stuart, A.D., et al. (2011) A Rab11- and microtubule-dependent mechanism for cytoplasmic transport of influenza A virus viral RNA. *Journal of Virology*. 85 (9), 4143–4156.

[82] Babcock, H.P., Chen, C., and Zhuang, X. (2004) Using single-particle tracking to study nuclear trafficking of viral genes. *Biophysical Journal*. 87 (4), 2749–2758.

[83] Spirin, A.S. (1986) Ribossome structure and protein biosynthesis. *Benjamin/Cummings, Menlo Park, Calif.*

[84] Lamb, R.A. and Krung, R.M. (2001) Orthomyxoviridae: the viruses and their replication. *In: Knipe, D.M., Howley, P.M. (Ed.), Fields Virology 4th Edition. Lippincott Williams & Wilkins: Philadelphia, Pa*. 1487–1531.

Appendix

This chapter comprises additional data regarding the developed intracellular stochastic model of IAV and DIP replication. Specifically, includes the implemented parameters values and the number of simulations performed for each initial infection condition.

A.1 Implemented parameters

Table A.1.1 | List of implemented parameters of the model.

Parameter	Value	Unit	Description	Source
B_{Hi}^{Tot}	150	sites	total number of high-affinity binding sites	[35]
B_{Lo}^{Tot}	1000	sites	total number of low-affinity binding sites	[35]
D_{Rib}	160	nucleotides	distance between two adjacent ribosomes on an mRNA	[79]
F_{Adv}	3.65[a]	-	replication advantage of DI RNA	[9]
F_{Fus}	0.51	-	fraction of fusion-competent virions	[33]
F_{Spl7}	0.02	-	fraction of M2-encoding mRNAs	based on ratio of M1 to M1 in a virion
F_{Spl8}	0.125	-	fraction of NEP-encoding mRNAs	[80]
$K_{V^{Rel}}$	10	virions	influence of protein concentration on virus release	[9]
k_{Hi}^{Att}	$8.09 \text{x} 10^{-2}$	$\text{site}^{-1} \cdot \text{h}^{-1}$	attachment rate to high-affinity binding sites	adjusted to data in [35]
k_{Lo}^{Att}	$4.55 \text{x} 10^{-4}$	$\text{site}^{-1} \cdot \text{h}^{-1}$	attachment rate to low-affinity binding sites	adjusted to data in [35]
k_{M1}^{Bind}	$1.39 \text{x} 10^{-6}$	$\text{molecule}^{-1} \cdot \text{h}^{-1}$	binding rate of M1 to nuclear vRNPs	[33]
k_{NP}^{Bind}	$3.01 \text{x} 10^{-4}$	$\text{molecule}^{-1} \cdot \text{h}^{-1}$	binding rate of NP to RdRp-RNA complexes	[33]
k_{RdRp}^{Bind}	1	$\text{molecule}^{-1} \cdot \text{h}^{-1}$	binding rate of RdRp-complexes to vRNA/cRNA	[33]

k^{Cplx}	1	molecule^{-7}·h^{-1}	formation rate of complexes containing eight vRNPs	fast formation is assumed
k_{M}^{Deg}	0.33	h^{-1}	degradation rate of mRNAs	[33]
k_{R}^{Deg}	36.36	h^{-1}	degradation rate of naked cRNA/vRNA	[33]
k_{Rnp}^{Deg}	0.09	h^{-1}	degradation rate of RNPs	[33]
k_{RRdRp}^{Deg}	4.25	h^{-1}	degradation rate of RdRp-RNA complexes	[33]
k^{En}	4.8	h^{-1}	endocytosis rate	[33]
k_{Hi}^{Eq}	1.13x10^{-2}	site^{-1}	equilibrium constant of high-affinity binding sites	[35]
k_{Lo}^{Eq}	8.33x10^{-5}	site^{-1}	equilibrium constant of low-affinity binding sites	[35]
k^{Exp}	1x10^{-6}	molecule^{-1}·h^{-1}	rate of NEP binding and nuclear export	adjusted to [81]
k^{Fus}	3.21	h^{-1}	fusion with endosomes rate	[33]
k^{Imp}	6	h^{-1}	nuclear import rate	[82]
k^{Rel}	3.7x10^{-3}	virions· molecule^{-1}·h^{-1}	virus release rate	[33]
k_{C}^{Syn}	1.38	h^{-1}	cRNA synthesis rate	[33]
k_{M}^{Syn}	2.5x10^{5}	nucleotides·h^{-1}	mRNA synthesis rate	[33]
k_{P}^{Syn}	64800	nucleotides·h^{-1}	protein synthesis rate	[83]
k_{V}^{Syn}	13.86	h^{-1}	vRNA synthesis rate	[33]
L_{1}^{M}	2320	nucleotides	length of the mRNA of segment 1	[84]
L_{2}^{M}	2320	nucleotides	length of the mRNA of segment 2	[84]
L_{3}^{M}	2211	nucleotides	length of the mRNA of segment 3	[84]
L_{4}^{M}	1757	nucleotides	length of the mRNA of segment 4	[84]
L_{5}^{M}	1540	nucleotides	length of the mRNA of segment 5	[84]
L_{6}^{M}	1392	nucleotides	length of the mRNA of segment 6	[84]
L_{7}^{M}	1005	nucleotides	length of the unspliced mRNA of segment 7	[84]
L_{8}^{M}	868	nucleotides	length of the unspliced mRNA of segment 8	[84]
L_{1}^{V}	2341	nucleotides	length of the vRNA and cRNA of segment 1	[84]
L_{2}^{V}	2341	nucleotides	length of the vRNA and cRNA of segment 2	[84]

L_3^V	2233	nucleotides	length of the vRNA and cRNA of segment 3	[84]
L_4^V	1778	nucleotides	length of the vRNA and cRNA of segment 4	[84]
L_5^V	1565	nucleotides	length of the vRNA and cRNA of segment 5	[84]
L_6^V	1413	nucleotides	length of the vRNA and cRNA of segment 6	[84]
L_7^V	1027	nucleotides	length of the vRNA and cRNA of segment 7	[84]
L_8^V	890	nucleotides	length of the vRNA and cRNA of segment 8	[84]
L_{DI}^V	480[a]	nucleotides	length of the vRNA and cRNA of DI segment	[9]
$N_{P_{HA}}$	500	molecules·virion^{-1}	number of HA molecules in a virus particle	[84]
$N_{P_{M1}}$	3000	molecules·virion^{-1}	number of M1 molecules in a virus particle	[84]
$N_{P_{M2}}$	40	molecules·virion^{-1}	number of M2 molecules in a virus particle	[84]
$N_{P_{NA}}$	100	molecules·virion^{-1}	number of NA molecules in a virus particle	[84]
N_{M1}^{Nuc}	200	nucleotides	number of nucleotides bound by one M1 molecule	[51]
N_{NP}^{Nuc}	24	nucleotides	number of nucleotides bound by one NP molecule	[39]

[a] Advantage and length for DI RNA of segment 3. Simulations performed for a DI RNA of segment 4 use $F_{Adv} = 4.39$ and $L_{DI}^V = 330$ nt.

Table A.1.2 | List of additional implemented parameters in the *Replication Model*.[b]

Parameter	Value	Units	Description	Source
$N_{P_{NP}}$	1000	molecules·virion^{-1}	number of NP molecules in a virus particle	[84]
$N_{P_{NEP}}$	165	molecules·virion^{-1}	number of NEP molecules in a virus particle	[84]
$N_{P_{RdRp}}$	45	molecules·virion^{-1}	number of RdRp molecules in a virus particle	[84]

[b] In the modified model was implemented the parameters shown and the values from Table A.1.1.

A.2 Number of simulations performed

Table A.2.1 | Number of simulations performed for different initial infection conditions.

MOI	MODIP	Number of simulations
1	1	6352
2	1	2201
5	1	854
10	1	1177
1	2	1872
1	5	2857
1	10	3867

Table A.2.2 | Number of simulations obtained for each random generated delay.

	Average delay (h)	Delay range (h)	Number of simulations
Time delay of DIP entry	2	[2.25 - 1.75]	23
	1.5	[1.75 - 1.25]	65
	1	[1.25 - 0.75]	253
	0.5	[0.75 - 0.25]	853
	0	[0.25 - 0.25]	990
Time delay of STV entry	0.5	[0.75 - 0.25]	791
	1	[1.25 - 0.75]	244
	1.5	[1.75 - 1.25]	63
	2	[2.25 - 1.75]	22

Table A.2.3 | Number of simulations performed for each induced delay and a DIP carrying a DI segment 3.

	Exact delay (h)	Number of simulations
Time delay of DIP entry	2	877
	3	1057
	4	7528
Time delay of STV entry	2	1750
	3	2683
	4	6407

Table A.2.4 | Number of simulations performed for each induced delay and a DIP carrying a DI segment 4.

	Exact delay (h)	Number of simulations
Time delay of DIP entry	2	570
	3	1302
	4	1304
Time delay of STV entry	2	1412
	3	1279
	4	1471