

Modeling and Simulation of HIV-1 Intracellular Replication

MSc Thesis

Author

Narges Zarrabi

Supervisor:

Prof. Dr. Peter M.A. Sloot

Submitted to Faculty of Science in partial fulfillment of the requirements for
the degree of Master of Science in Grid Computing, Computational Science

Computational Science
Faculty of Science
University of Amsterdam



UNIVERSITEIT VAN AMSTERDAM

2009

Modeling and Simulation of HIV-1 Intracellular Replication

MSc Thesis

Author

Narges Zarrabi

Supervisor:

Prof. Dr. Peter M.A. Slood

Committee Members:

Prof. Peter Slood

Prof. Jaap Kaandorp

Dr. Inge Bethke

Mr. Emiliano Mancini (MSc)

Abstract

Modeling HIV infection has a long history since 1980s when HIV was discovered and the AIDS symptoms were first observed.

In recent decades, treating Human Immunodeficiency Virus infection and AIDS has been the main research topic of many experiments in laboratories around the world. In spite of the broad effort in designing antiretroviral drugs for treating HIV patients, there is still no specific vaccine or drug available to completely block the virus replication. Many mathematical and computational models have been developed to investigate the complexity of HIV dynamics, immune response and drug therapy.

The objective of this thesis is to model the intracellular replication cycle of HIV-1. The model is basically designed to study HIV replication inside a single cell before initiation of drug therapy. Two modeling approaches have been used to implement the virus replication model: *Rate-limited approach* and *diffusion-limited approach*. Results of the model are discussed based on the number of cDNA created and integrated into the host cell DNA, the number of viral mRNA transcribed, and the amount of translated viral proteins after a single round of virus replication. Model validation is demonstrated graphically and quantitatively based on comparison between simulation results and available experimental data. A sensitivity analysis is also performed on some of the parameters used in the model. This analysis helps us to determine which steps, in the viral replication process, are more sensitive and uncertain. These results also indicate which parameters are critical

in the replication process and which parameters are less sensitive. Results obtained from the simulation give insights about the detail of HIV replication dynamics inside the cell at protein level. Therefore the model can be used for future studies of HIV intracellular replication in vivo and drug treatment.

Acknowledgments

First and foremost, I would like to take this opportunity to express my most sincere gratitude and appreciation to my supervisor, Professor Peter Sloot, for his countless efforts in guiding and encouraging me throughout my research. He has supported me and given me one of the best opportunities in my academic life to accomplish this research thesis. His knowledgeable suggestions and heuristic advice always inspired me on my way.

This work was made possible through the ViroLab project (www.virolab.org) and I am grateful to all ViroLab group members, especially Dr. Victor Muller and Dr. David van de Vijver. I am also grateful to Dr. Tay Joc Cing and Assoc. Prof. Ong Yew Soon from Nanyang Technological University.

I would like to thank Emiliano for his great helps throughout this thesis and the model and also thank to Gokhan and Rick for their feedbacks on the model.

Again thanks to all my friends in Amsterdam including: Gokhan, Emiliano, Rick, Chinh, David, Toktam, Erik and all the people in the Section of Computational Science who made it a wonderful experience for me at SCS and Universiteit van Amsterdam.

Last and not least, I would like to thank my parents for their endless love, a special thanks to beloved Shayan for his countless support, and thank to all my family and friends back in Iran who encouraged and helped me during my whole life.

Contents

| | |
|--|------------|
| Abstract | i |
| Acknowledgments | iii |
| 1 Introduction | 1 |
| 1.1 Background and Motivation | 1 |
| 1.2 Scope and Objectives | 3 |
| 1.3 Organization of the thesis | 3 |
| 2 Literature Review | 5 |
| 2.1 HIV Structure and Classification | 5 |
| 2.1.1 HIV-1 Structure | 6 |
| HIV-1 Enzymes | 7 |
| HIV-1 Genes | 7 |
| 2.1.2 HIV-1 Subtypes | 8 |
| 2.2 HIV-1 Replication and Intracellular Growth | 9 |
| 2.2.1 HIV Replication cycle | 9 |
| 2.2.2 Replication Inhibitors | 13 |
| 2.2.3 Replication Dynamics in Lymph Node | 14 |

| | | |
|----------|---|-----------|
| 2.3 | Existing Models of HIV Replication | 15 |
| 2.4 | Summary | 17 |
| 3 | Model Design and Implementation | 19 |
| 3.1 | General Model Design | 20 |
| 3.1.1 | States and Transitions | 21 |
| 3.1.2 | Cell Infection | 22 |
| 3.2 | Rate-limited Approach | 24 |
| 3.2.1 | Model Entities | 24 |
| 3.2.2 | Parameters and Assumptions | 25 |
| 3.3 | Diffusion-limited Approach | 27 |
| 3.3.1 | Model Entities | 29 |
| 3.3.2 | Parameters and Assumptions | 30 |
| 3.4 | Model Implementation | 34 |
| 3.5 | Summary | 34 |
| 4 | Simulation Results, Analysis, and Validation | 35 |
| 4.1 | Primary Results and Validation | 36 |
| 4.1.1 | Integration | 36 |
| | Nuclear Transfer | 38 |
| | Transcription | 38 |
| | Translation | 40 |
| 4.2 | Sensitivity Analysis | 41 |
| | Nuclear Transfer Rate | 43 |
| | Integration Rate | 46 |
| | Transcription Rate | 48 |
| 4.3 | Comparison of the Two Approaches | 50 |

| | | |
|----------|---|-----------|
| 4.3.1 | Simulation Performance | 52 |
| 4.4 | Summary | 53 |
| 5 | Conclusions and Future Work | 55 |
| 5.1 | Summary of Contributions | 55 |
| 5.2 | Recommendations for Future Work | 56 |
| | Bibliography | 59 |

List of Figures

| | | |
|-----|---|----|
| 2.1 | HIV-1 structure | 6 |
| 2.2 | HIV genome | 7 |
| 2.3 | Different subtypes of HIV-1 | 8 |
| 2.4 | HIV intracellular replication cycle | 10 |
| 2.5 | Three steps in HIV reverse transcription | 11 |
| 2.6 | HIV infected CD4+ T cells in lymph node | 15 |
| 3.1 | 7 main steps of HIV replicayion process | 20 |
| 3.2 | Cell states and Transitions | 22 |
| 3.3 | Proportion of cells getting infected under different MOIs | 23 |
| 3.4 | Cell internal variables and intracellular rates | 25 |
| 3.5 | Flowchart of the rate-limited model | 28 |
| 3.6 | Simulation visualization of early stages of simulation: viral RNAs (red) are assigned to the cell cytoplasm environment. The purple circles are the cellular tRNA available in the cell | 31 |
| 3.7 | Simulation visualization of late stages of simulation: viral mRNA (small green) mainly in the cell nucleus, and translated viral proteins (big green) in the cell cytoplasm. The purple circles are the cellular tRNA available in the cell | 31 |

| | | |
|------|---|----|
| 3.8 | Flowchart of the diffusion-limited approach showing the main loop of the simulation | 33 |
| 4.1 | Copies of integrated cDNA per cell at MOI 0.4 | 37 |
| 4.2 | Integrated cDNA for low and high values of MOI | 38 |
| 4.3 | Integrated cDNA with two different nuclear transport rate | 39 |
| 4.4 | Expression of viral mRNA in experimental results and simulation results . | 40 |
| 4.5 | mRNA transcripts at different MOI | 41 |
| 4.6 | Viral protein translated at different MOI | 42 |
| 4.7 | Sensitivity analysis for nuclear transport rate on the output of integrated cDNA | 44 |
| 4.8 | Sensitivity of integrated cDNA to nuclear transport rate at different time points | 44 |
| 4.9 | Sensitivity analysis for nuclear transport parameter on the mRNA transcripts | 45 |
| 4.10 | Sensitivity analysis for integration parameter on the integrated cDNA . . . | 47 |
| 4.11 | Sensitivity of integrated cDNA to integration rate at different time points . | 48 |
| 4.12 | Sensitivity analysis for transcription parameter on the translated viral proteins | 49 |
| 4.13 | Sensitivity of viral proteins to transcription rate at different time points . . | 50 |
| 4.14 | Simulation results of integrated cDNA for the two approaches: Diffusion-limited approach and Rate-limited | 51 |

List of Tables

| | | |
|-----|--|----|
| 3.1 | Model Parameters (Rate-limited approach) | 26 |
| 3.2 | Model Parameters (Diffusion-limited approach) | 32 |
| 3.3 | Agent Collision Rules | 33 |
| 4.1 | Comparison of Approaches: Rate-limited vs. Diffusion-limited | 52 |

Chapter 1

Introduction

1.1 Background and Motivation

Human immunodeficiency virus (HIV), a member of retrovirus family, causes acquired immunodeficiency syndrome (AIDS). HIV declines the number of immune cells in an infected body and causes the immune systems to lack the ability to fight off other infectious diseases. AIDS has so far killed more than 25 million infected people around the world [1]. Hence, HIV infection and AIDS have turned out to one of the most miserable epidemics since 1980s when the first evidences of the disease were observed.

In the field of virology and infectious disease, mathematical and computational models are developed to help a better understanding of the underlying biological phenomena and medical processes. Some virological or medical experiments may be expensive or time-consuming to conduct, which makes them infeasible to study. For example progression of HIV infection is slow and development of AIDS is a long term process (approximately 10-12 years under anti-retroviral therapies) [2]. In this case, conducting medical experiments to study the dynamics of HIV is time-consuming and infeasible. Therefore, computer models of HIV dynamics and immune system response could be a time-efficient and cost-

effective approach to this problem. These models assist researchers and pharmacologists to study the viral dynamics and investigate the effect of different drugs on disease progression (ViroLab, the virtual laboratory for decision support in viral disease treatment [3]).

In recent years, many models that vary from mathematical and equation-based to cellular automata and agent-based [2, 4, 5, 6, 7] have been developed to investigate the complexity of HIV dynamics, immune system response, and drug therapy. Simulation of HIV infection can vary from modeling the entire immune response (under virus infection and drug therapy) down to simulating the intracellular processes in single cell infected by a virus. However, there are not many models which consider the viral intracellular replication at single cell level.

To study viral dynamics at a single-cell level analysis, some techniques have been developed such as microarrays, expressed tag sequencing (EST), serial analysis of gene expression (SAGE) and polymerase chain reaction (PCR). However, these techniques have some limitations which still remain a challenging issue. These limitations basically include: sensitivity level of the technique, artifact and the accuracy level of the results [8]. Modeling intracellular processes at a single-cell level is a possible approach to address these limitations. In addition, in the case of HIV infection, most antiretroviral drugs are designed to inhibit HIV replication inside infected cells. For example blocking reverse transcription and integration through the use of reverse transcription inhibitors, protease inhibitors, or a combination of both [9]. Therefore, quantitative measurements of HIV intracellular replication at single-cell level is critical for drug development, and pathogenesis studies [10].

In this thesis we modeled the intracellular replication of HIV using a combination of mathematical and agent-based modeling techniques. Cells are individual entities in the model and the state of each cell changes over time based on its internal variables and quantities.

1.2 Scope and Objectives

To make the content of this thesis clear, we will focus on the word *replication*. Replication is applied in different contexts in microbiology which refers to different functionalities. For example cell replication causes cell division and generation of daughter cells from a source cell. Virus replication refers to the 'growth' of a virus inside a living cell which leads to production and release of new virions from the infected cell. For simplicity replication is referred to virus replication throughout this thesis.

The objective of the thesis is to model HIV intracellular replication from the early phases of infection including reverse transcription and integration to late phases including mRNA transcription and production of new viral proteins inside the cell.

1.3 Organization of the thesis

The remainder of the thesis is organized as follow: In chapter 2 we first introduce HIV structure and subtypes. Then we explain HIV replication and intracellular growth and afterward introduce current retroviral drugs for inhibiting HIV replication. HIV replication dynamics in lymph nodes, as reservoirs sites for viral replication, is explained. A literature review on existing models of viral replication is also presented at the end of the chapter. Chapter 3 describes the model design and implementation. We first introduce the general model design and then describe 2 approaches for implementing the model: Rate-limited approach and diffusion-limited approach. Section 4 is divided into 2 sections. First, in Section 4.1 we will present the primarily results of the simulation results and demonstrate the validation of the results. Second, in section 4.2 we perform a sensitivity analysis on the key parameters in the model and made a discussion on the analysis results. Chapter 5 concludes the thesis with a summary and recommendations for future work.

Chapter 2

Literature Review

2.1 HIV Structure and Classification

There are two major types of human immunodeficiency virus (HIV): HIV type 1 and type 2 (HIV-1 and HIV-2). Both HIV-1 and HIV-2 originate from simian immunodeficiency viruses (SIVs), spread directly from human to human (i. e. sexual contact, through blood, mother to child) and appear to cause AIDS. However, it seems that HIV-2 has a lower rate of disease transmission and progression and there has been direct evidence of lower viral replication rates in vivo in HIV-2 infection than in HIV-1 [11]. Similarly, plasma viral loads are lower in HIV-2 infection, suggesting that HIV-2 replication is limited in vivo in comparison to that of HIV-1 [11].

Why did we choose HIV type 1? As mentioned above, studies have shown that HIV-1 is more pathogenic than HIV-2 and has a higher rate of disease progression [11]. Also a huge amount of data and knowledge on the kinetics of HIV-1 and dynamics of its replication, which helps to validate the results of the simulation, is available in the literature. Hence, we chose HIV-1 data for simulation and validating our model. In the rest of the thesis the word HIV explicitly refers to HIV-1, unless specified.

2.1.1 HIV-1 Structure

HIV has a simple structure. Figure 2.1 shows the structure of HIV. It contains two identical copies of positive single-stranded RNA. Each RNA is about 9500 nucleotides long and encodes for 9 genes. About 2000 viral p24 proteins form the capsid. The viral capsid contains the viral RNA-strands, viral enzymes and the nucleocapsid proteins p7. The capsid is surrounded by a matrix composed of the viral protein p17. The whole matrix and viral proteins are packed in the viral envelope. On the surface of the virus there are embedded glycoproteins, gp120 and gp41, to the viral envelope which enables the virus to attach and fuse with the target cells [12].

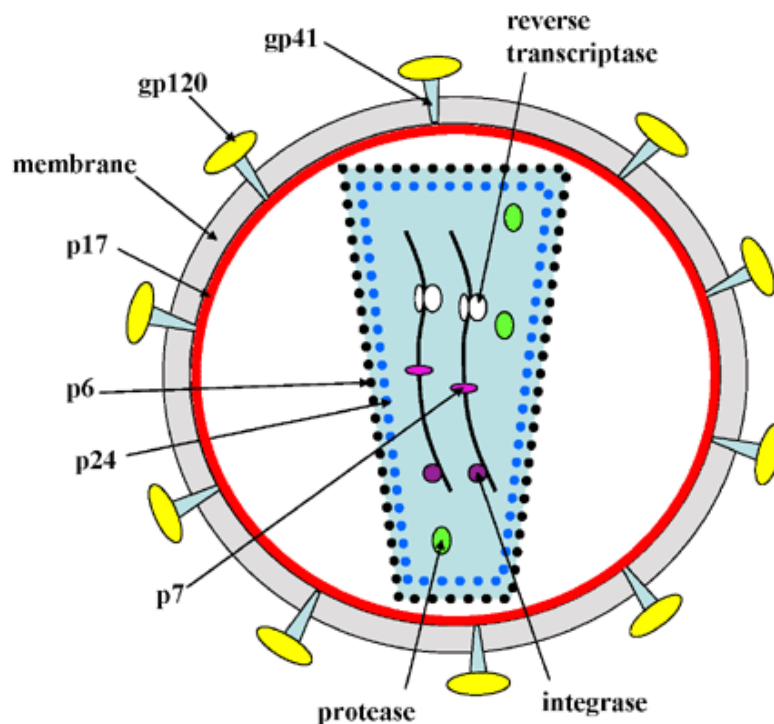


Figure 2.1: HIV-1 structure: Viral envelope protein forms the surface of the virus. Embedded on the viral surface there are glycoproteins (gp120 and gp41) which help the binding and fusion of the virus to the target cell. The viral capsid (p24, p6) inside the cell contains the two viral RNA-strands, viral reverse transcriptase, integrase, and protease enzymes and the nucleocapsid proteins p7

HIV-1 Enzymes

HIV contains three major enzymes: Reverse transcriptase (RT), integrase (IN) and protease (PR). These three enzymes are encoded by the viral pol gene.

HIV-1 Genes

HIV RNA consists of 9 genes. Some genes are common in all retroviruses like Gag, Pol and Env, while some others like Vpu are specific to HIV-1 retrovirus. The first three genes (Gag, Pol and Env) are called the main genes and the rest are called auxiliary genes.

Figure 2.2 shows the structure and genes on the HIV-1 RNA genome.

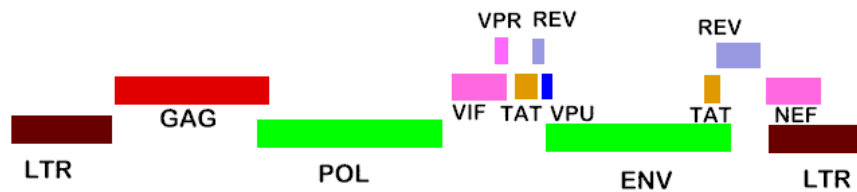


Figure 2.2: HIV genome consists of 9 genes: Gag, Pol, Env, Tat, Rev, Nef, Vif, Vpr and Vpu. LTR is the Long Terminal Repeat at each end of the genome

In what follows, we will have a closer look to the function of each gene.

- **Gag:** Gag is a group of antigen genes which codes for various HIV structural proteins (p24 capsid protein, p6, p7, p17 matrix protein)
- **Pol:** The polymerase gene encodes the necessary enzymes for virus replication (RT, IN, PR).
- **Env:** The envelope gene codes for viral envelope proteins (gp120 and gp41) which are required for viral entry process.
- **Tat:** Tat is the transactivator gene which is expressed early in the life cycle of HIV. It is essential for viral gene expression and controls the transactivation of all HIV proteins.

- **Rev**: Rev is a regulator of viral protein expression.
- **Nef**: Nef is a negative regulator factor which retards HIV replication.
- **Vif**: Vif is a virus infectivity factor gene that disrupts antiviral activity.
- **Vpr**: The virus protein R gene encodes for viral protein R and plays as a regulator in Macrophages.
- **Vpu**: The virus protein U gene is required for efficient viral replication and release. It is found only in HIV-1.

2.1.2 HIV-1 Subtypes

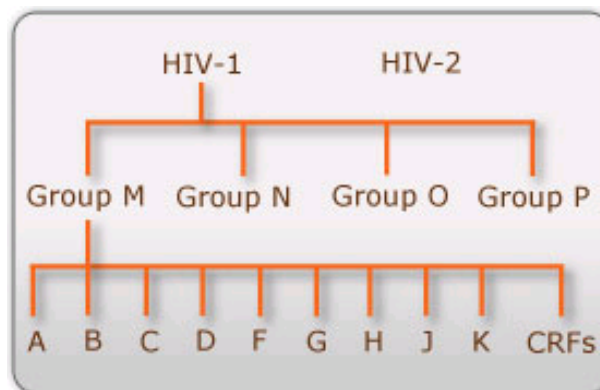


Figure 2.3: Different subtypes of HIV-1

The strains of HIV-1 are classified into four groups: the major group M, and the groups N, O and P. These four groups may represent separate introductions of simian immunodeficiency virus into humans. Groups N, O and P are not spread worldwide and are mainly confined to West and Central Africa, while more than 90 percent of HIV-1 infections belong to HIV-1 group M [13]. Due to the high error rate in reverse transcription and RNA replication [14], group M of HIV-1, has a complex classification. As shown in Figure 2.3 subtypes of group M are subclassified as: A, B, C, D, F, G, H, J, K and CRFs (circulating recombinant forms). Among them, subtype B is the most common subtype in Europe,

America, Japan, Australia, and some Asian countries, while subtype A, C and D are more common in Africa. The less common subtypes F, G, H, J, K are mainly spread through Africa, South America, and some parts of Europe. If two viruses of different subtypes happen to enter a naive cell, the genetic material of both strains may mix together during the replication process. This leads to a new viral strain called circulating recombinant forms or CRFs. Having virus mutation and replication, newly HIV genetic subtypes and CFRs are likely to appear [13].

2.2 HIV-1 Replication and Intracellular Growth

Despite a simple structure, HIV has a complex dynamics inside the cell.

2.2.1 HIV Replication cycle

HIV replication first requires the virus to enter an uninfected host cell such as CD4+ cells or macrophages. After the virus enters the cell, it can utilize the cell machinery for replication and release of new virions. Figure 2.4 illustrates the intra-cellular replication process of HIV-1 from entering the cell to releasing new virus particles [15]. HIV replication consists of several sub processes: entry, reverse transcription, nuclear transfer, integration, transcription, export to cytoplasm, translation, assembly, budding and maturation [9]. In what follows, each sub process in the viral replication cycle is briefly described.

- **Entry:** HIV operates by first entering into the target cell. Entry itself is a complex multi-step process which is done through conformational changes and receptor bindings [16]. Cell entry consists of the three following steps:
 - **Binding to CD4 receptors:** Entry process starts with HIV surface glycoprotein (gp120) binding to the CD4 receptors located on the surface of T cells.
 - **Co-receptor binding:** After attaching to CD4 receptors, HIV is pulled close to

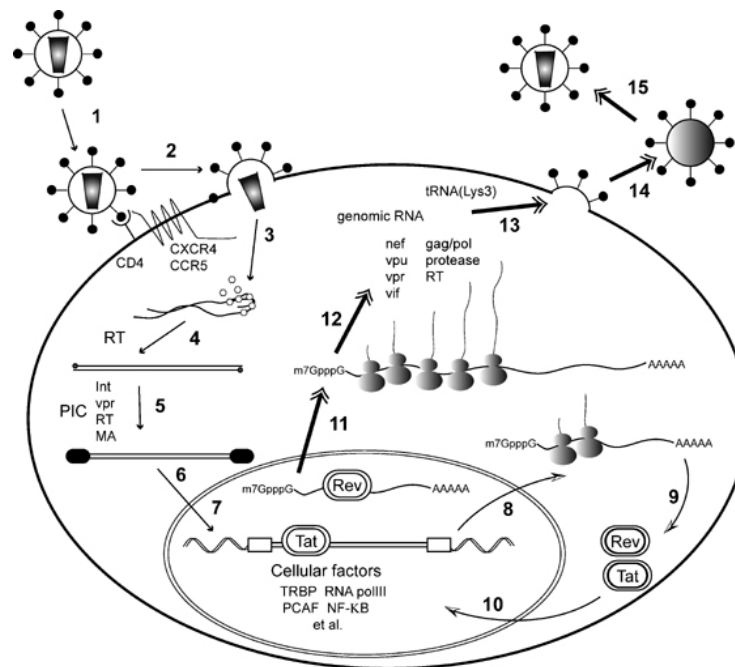


Figure 2.4: HIV intracellular replication cycle: Entry(Binding to CD4 receptor(1), Coreceptor binding(2), Fusion(3)), Reverse transcription(4,5), Transfer to the nucleus(6), Integration (7), Transcription and regulation(8,9,10), Export to the cytoplasm (11), Translation (12), Assembly, budding and maturation(13,14,15)

cell surface. Then HIV gp120 receptor mediates an interaction with chemokine co-receptors (i.e. CCR5 or CXCR4) embedded in the surface of immune cells.

- **Fusion:**entry process is completed by the fusion of cell membrane with viral membrane which is subsequent to conformational change in gp41. Gp41 is a glycoprotein on the surface of HIV virus. Once fusion takes place, viral RNA and HIV enzymes such as reverse transcriptase, integrase, and protease are injected into the cell.
- **Reverse Transcription:** As mentioned earlier, HIV genetic information is in the form of RNA. This RNA genome needs to be converted into double stranded DNA in order to be integrated into the host cell DNA. Without reverse transcription the viral genome is can not be incorporated into the host cell and thus cannot be replicated to form new copies of virions. Reverse transcription is composed of several steps and

usually occurs after the release of the viral core into the cytoplasm of the target cell [17]. Viral RNA serves as the template for reverse transcription. HIV uses cellular tRNA as primer and RT enzyme to complete the reverse transcription process. As HIV-1 has positive RNA strands, the minus-strand HIV DNA (minus-strand strong-stop DNA) is the first transcript. The efficiency of this reaction is about 45% [18].

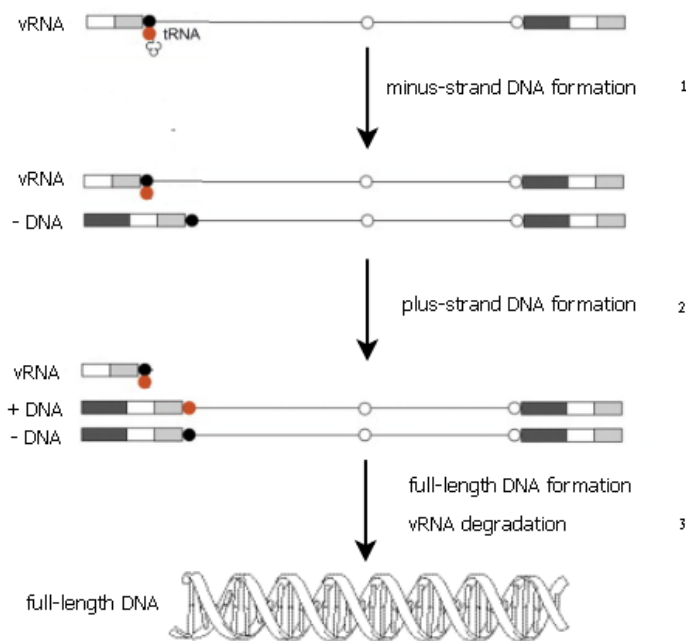


Figure 2.5: Three main steps of HIV reverse transcription: First the minus-strand DNA is synthesized from viral RNA (vRNA). Second step the plus-strand DNA is synthesized. At third step these is the full-length double-strand DNA formation and vRNA degradation

The minus-strand DNA species will elongate and the plus-strand HIV DNA species will be synthesized from it later in the reverse transcription process. The reaction efficiency of the plus-strands HIV DNA synthesis is 40% [18]. Viral RNA synthesis to DNA is usually completed 12 to 16 hours post infection and the template RNA at each synthesis is subsequently degraded [19]. Figure 2.5 is inspired from [20] and shows the three main steps in HIV reverse transcription.

- **Nuclear Transport:** The synthesized double-strand viral DNA is moved to the cell nucleus. The nuclear transport of viral DNA is an active and rapid process, and all the viral DNA molecules synthesized in the cytoplasm are transported to the nucleus within 24 hours post infection [21].
- **Integration:** Integration is the process of viral DNA insertion into the host cell DNA. The integrated form of viral DNA is necessary for viral replication [11]. HIV integrase enzyme is responsible for integration process and is a key factor in determining the site of integration. The integrated DNA is also called provirus. Integration can occur at many locations in the human genome. About 40% of integrations occur in transcription units of the host genome [22]. Once integration occurs the host cell is known as an infected cell carrying provirus in its DNA. Further, the infected cell can remain silent and latently infected or become activated and turn to a productively infected cell.
- **Transcription:** After integration, the cell may turn to an activated state and become transcriptionally active. As long as the cell is not activated the transcription of integrated virus does not occur. Activation of the cell induces transcription of integrated proviral DNA into messenger RNA (mRNA). Acutely infected cells synthesize high levels of virus, whereas latently infected cells transcribe little or no viral RNAs. Transcription of the integrated viral genome produces full-length mRNA transcripts that either remain full-length or become spliced. The half-life of the mRNA is 0.2 per hour in both the nucleus and the cytoplasm [23].
- **Export to cytoplasm:** After transcription process in the nucleus the viral mRNAs are migrated to the cell cytoplasm where building blocks for a new virus are synthesized.
- **Translation:** In the cytoplasm viral mRNA undergoes translation resulting in the production of viral proteins. These proteins are produced as long polypeptide chains

that must be cleaved into smaller proteins in order to become functional. This step is performed by the viral protease enzyme. The structural proteins required for new viral particles are produced during this step.

- **Assembly and Budding:** Functional viral proteins and viral RNA assemble at the cell membrane. The viral RNA together with related proteins are packed and released from the cell surface by the budding process. The cellular plasma membrane is the primary site of HIV-1 budding [24].

2.2.2 Replication Inhibitors

Many drugs have been developed with the objective of blocking HIV replication and infection. The most important ones are Reverse Transcriptase Inhibitors (RTIs), Protease Inhibitors (PIs), and Entry Inhibitors (EIs). Inhibitors of HIV replication act at specific stages of the viral replication cycle [10]. Only the entry inhibitor is an extracellular inhibitor, while the others are acting intracellularly and inhibit the virus replication inside the cell.

- **Entry Inhibitor:** A new class of drugs which prevents the entry process by inhibiting either the binding process or the fusion of the virus and cell membrane.
- **Reverse Transcription Inhibitor:** A class of drugs which prevents reverse transcription process by inhibiting the reverse transcription enzyme. Blocking reverse transcription process prevents the single strand viral RNA to be converted into double strand viral DNA.
- **Integrase Inhibitor:** A class of drugs which prevents integration process by inhibiting the integrase enzyme. Blocking integration process prevents the viral DNA to be integrated into the host cell DNA.
- **Protease Inhibitor:** A class of drugs which act by inhibiting the protease enzyme. Protease enzyme cleaves viral proteins into smaller functional proteins.

2.2.3 Replication Dynamics in Lymph Node

The amount of HIV-1 RNA and viral load in the blood has been a measure of disease progression and response to antiretroviral therapy in HIV infected patients. However, a large proportion of viruses reside in lymphoid tissue. It is observed that only a small fraction (2%-4%) of the cells are circulating through the blood, while the majority of cells are in the lymphatic system [25,26]. 98% of lymphocytes are known to be in compartments of the lymphatic system such as thymus and lymph nodes, supporting the hypothesis that lymph nodes are major reservoirs of HIV infection in vivo [27]. Cell cycle analysis and labeling experiments in lymphoid histocultures (Eckstein et al. 2001 [27]) showed that within lymphoid microenvironment, HIV-1 replication can occur in all CD4+ T cell types apart from their phenotypic classification, proliferation status, or cell cycle progression [27]. Thus, resting and naive T cells in lymph nodes can also be permissive hosts for infection and may be infected by HIV-1, although quantitative measurements show that there is difference between viral burst size of infected memory and naive T cells (infected naive T cells have a reduced burst size compared to productively infected memory T cells due to co-receptor tropism [27,16]).

After entering the lymph node, T cells reside in the paracortex area (T zones). Within the lymph node, it is shown that the fibroblastic reticular cell (FRC) network regulates naive T-cells to access the paracortex area. In the paracortex area of the lymph node the cell movement is limited due to the dense population of cells within this domain [27]. Figure 2.6 illustrates the lymph node from an HIV infected patient. The infected cells are stained bright red. At the time of lymph node biopsy, the HIV infected patient had a CD4+ lymphocyte count of over 300 per mm^3 which is yet not an AIDS stage. The picture is taken from Wisconsin Viral Research Group [28].

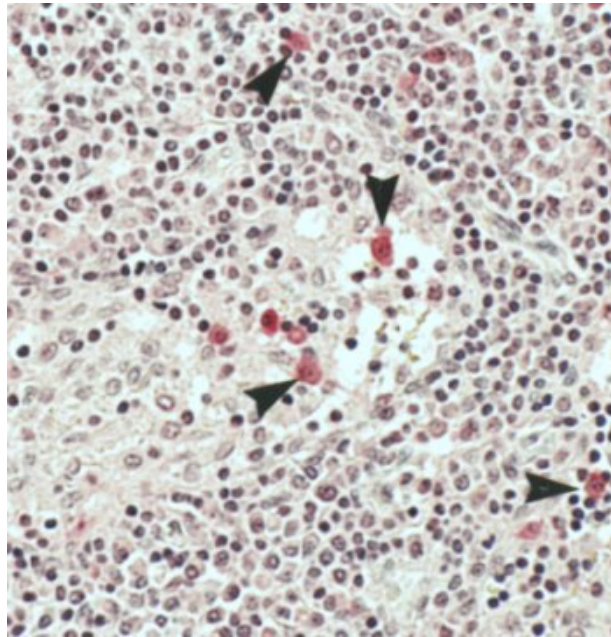


Figure 2.6: HIV infected CD4+ T cells in lymph node (marked in red)

2.3 Existing Models of HIV Replication

As HIV was discovered in 1980s, models of HIV infection and replication have been built in order to better understand the dynamics of the virus infection and progression to disease (AIDS) [2, 4, 5, 6, 7]. These models were built at different levels, from cellular up to population level. For a concrete example of this wide-range multiscale modeling see *HIV Decision Support: From Molecule to Man* by Sloat et al. 2009 [29,30]. Despite the wide range of modeling for HIV dynamics, more and more attention is being paid toward single cell analysis to understand intracellular processes.

At the cellular and molecular level different modeling techniques, from deterministic to stochastic [31] and from mathematical to agent-based models [23,32,33], have been used to simulate the intracellular process of viral replication.

The first model used to study intracellular virus replication was made by Reddy et al. 1999 [23] and represented the intracellular kinetics of HIV-1 by a mathematical model.

Their model is based on a system of coupled ordinary differential equations (ODEs) that is solved numerically. The simulation accounts for the kinetics of reverse transcription, integration of proviral DNA into the host genome, transcription, mRNA splicing and transport from the nucleus, translation, feedback of regulatory proteins to the nucleus, transport of viral proteins to the cell membrane, particle assembly, budding, and maturation. Each subprocess of the viral replication cycle is represented by one or more equations and the model provides concentration-based insights on how the overall replication cycle depends on its constituent reactions.

Although mathematical models are good for modeling the overall behavior of the system, a disadvantage is that they are less useful in capturing the individual interactions on system behavior. In such models only population level factors are considered, while ignoring spatial and topological dependencies on individual interactions.

On the other hand Agent-Based Simulation (ABS), an emerging field in modeling microbiological phenomena, models the system by treating cells and molecules as independent agents. The system behavior in ABS emerges through modeling the interaction between agents. Thus, in modeling complex behavior at molecular and cellular level such as receptor-binding kinetics, intra-cellular signaling, and virus-cell interactions ABS would be an efficient technique for modeling.

In an agent-based model entities are specified as agents at individual level and the complexity of the system emerges from simple rules defined between them. In modeling microorganisms at cellular level agent-based modeling has been widely used [34, 32, 35, 36, 5]. A previous model using ABS principles was *CellAK* by Webb et al. 2004 [32], the model included cellular details such as membrane with lipid bilayers, multiple compartments such as a variable number of mitochondria, substrate molecules, enzymes with reaction rules and metabolic pathways. Another example of ABS is used by Guo et. al. 2005 [36] for hypothesis testing at a cellular level by modeling cells as individual agents which interact.

Multi-agent simulation of HIV pathogenesis of different micro-biological hypotheses in progression of HIV and developing AIDS has been studied using CAFISS, a complex adaptive framework for immune system simulation. More detailed information about CAFISS is available in [37]. Although, agent-based modeling has advantages in modeling biological and micro-biological systems, such models are computationally more expensive than mathematical models. This problem is becoming less important nowadays with fast supercomputers and distributed computing.

In the modeling section 3 we have described in detail how we modeled the HIV-1 replication process and the intracellular kinetics of the virus.

2.4 Summary

In this section we described the classification and structure of HIV. We described the major types of HIV (HIV-1 and HIV-2) and their differences. We explained why we chose HIV type 1 for our simulation and introduced the structure of HIV-1 and its main viral compartments including viral genome and enzymes. We also introduced different subtypes of HIV-1 and their diversity. Then we introduced the HIV-1 replication and intracellular growth, the steps in replication cycle and current existing inhibitors of HIV replication. Since lymphocytes mostly reside in the lymphatic system we had a closer look into the dynamics of replication in the lymph node. In the end, we introduced the existing models and modeling techniques that have been used for studying virus intracellular growth and HIV replication dynamics.

Chapter 3

Model Design and Implementation

In this model we simulate the HIV-1 replication and intracellular growth. We consider two different approaches for implementing the model: *Rate-limited approach* and *diffusion-limited approach*. The rate-limited approach defines the internal state of a cell as a set of variables. The internal variables of the cell keep track of the molecular quantities inside the cell. These variables change based on rates inferred from literature. On the other hand, the diffusion-limited approach defines the internal state of the cell by simulating the molecular entities inside the cell as agents. Every entity's movement is described by a random walk at each time step and events occur based on the rules defined upon collision between entities.

The main idea of the model is explained in the Section General Model Design 3.1. Each of the approaches used to implement the model are explained separately in sections Rate-Limited Approach 3.2 and Diffusion-Limited 3.3

3.1 General Model Design

As mentioned, HIV has a complex dynamics inside the cell and the knowledge about the inner viral replication is still inadequate. So, many assumptions and simplifications are being made in modeling. For simplicity, we have only simulated the steps that play a significant role in the replication process. The following 7 main steps are considered in the model: Reverse transcription, nuclear transport, integration, cell activation, mRNA transcription, transport to the cytoplasm and translation.

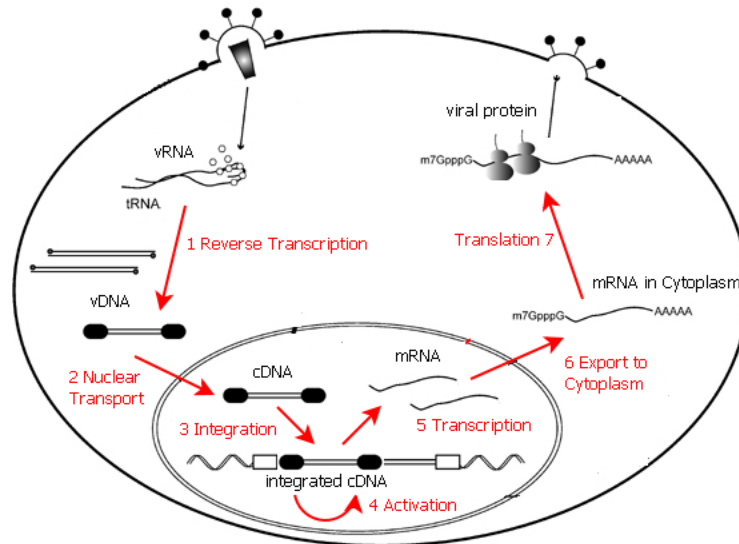


Figure 3.1: Seven main steps of HIV replication considered in the model: Reverse transcription, nuclear transport, integration, cell activation, transcription, transport to the cytoplasm and translation (The red arrows are the steps simulated in the model)

During the course of virus replication we keep track of the following quantities inside the cell by saving them in history files:

- viral RNA
- viral DNA in both cytoplasm and nucleus
- integrated cDNA
- transcribed viral mRNA

- translated viral protein

This data is later used for results analysis.

3.1.1 States and Transitions

The life cycle of HIV inside the cell is arbitrarily divided into two distinct phases: The early phase which refers to the steps of replication from cell entry to integration of viral cDNA into the cell genome, whereas the late phase refers to the post-integration steps from cell activation and mRNA transcription to budding and release of new viruses from the cell. In our simulation we have defined 6 states for the cell, 3 states in the early phase of viral replication and 3 in the late phase. The 6 cell states are described as follows:

- **Healthy:** All cells are healthy when they are created. Later they may become infected.
- **Early Infected:** Healthy cells randomly become infected with certain number of viruses based on a stochastic process.
- **Latently Infected:** The cell is latently infected if the viral DNA integrates into the cell DNA and the cell is still silent.
- **Actively Infected:** The cell is actively infected by viral gene expression and activation, which results in mRNA transcription process.
- **Productively Infected:** The cell is productively infected when it starts to produce new viral particles through the budding process.
- **Death and Lysis:** The cell is dead at the end of the simulation. Viral particles present inside the cell are released at its death by lysis.

Hence, each cell in the model has one of the above states at each time. The state of a cell may change during its life span. Figure 3.2 shows different state transition of a cell during the viral replication process.

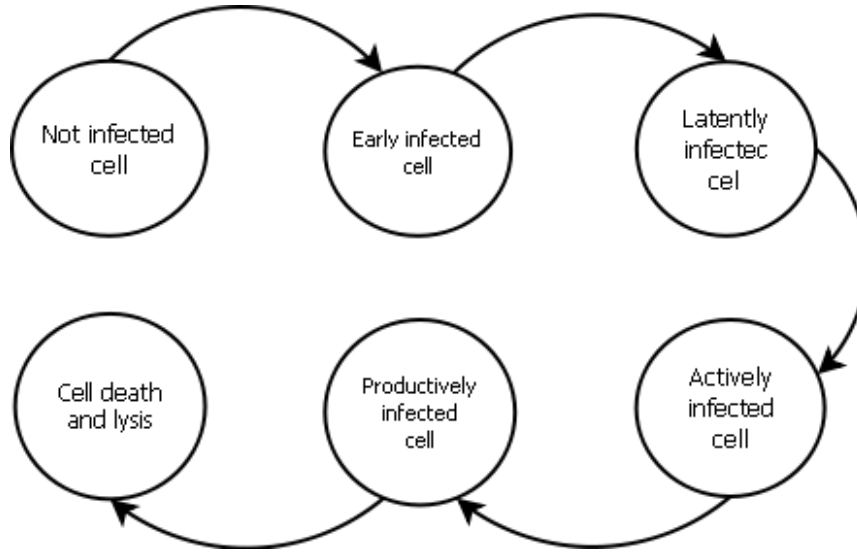


Figure 3.2: Cell six possible states. Arrows indicate the possible transitions between the cell states.

3.1.2 Cell Infection

Cell Infection is simulated as a stochastic process in the model. Each cell may get infected (with 1 or more viruses) or remain uninfected based on the multiplicity of infection (MOI), which is specified as an input value in the model. The MOI is obtained by the ratio of virus particles to cells or the average number of virus particles per cell. The MOI in real experiments is determined by dividing the number of viruses per ml by the number of cells per ml.

$$MOI = \frac{\text{Viral load (per ml)}}{\text{number of cells (per ml)}} \quad (3.1)$$

Although the MOI represents the average number of viruses per cell, the actual number of viruses that infect any given cell follows a statistical distribution. The proportion of cells in a population infected by a specific number of viruses can be calculated from the Poisson distribution [38]. Hence, for a certain MOI, the number of viruses infecting a cell is following a Poisson distribution,

$$P(n) = \frac{m^n \cdot e^{-m}}{n!} \quad (3.2)$$

where m is the average number of viral particles per cell (MOI), n is the specific number of viruses infected the cell, and $P(n)$ is probability that the cell will get infected at least with n viruses. For example if $m = 1$, a higher proportion of cells will get infected with one virus or will not get infected (infected with 0 viruses), while some cells may get infected with two or more viruses. At a low MOI most cells will be infected with fewer viruses and at a high MOI most cells will be infected with more viruses. These results agree with what we would expect from Poisson distribution as shown in Figure 3.3.

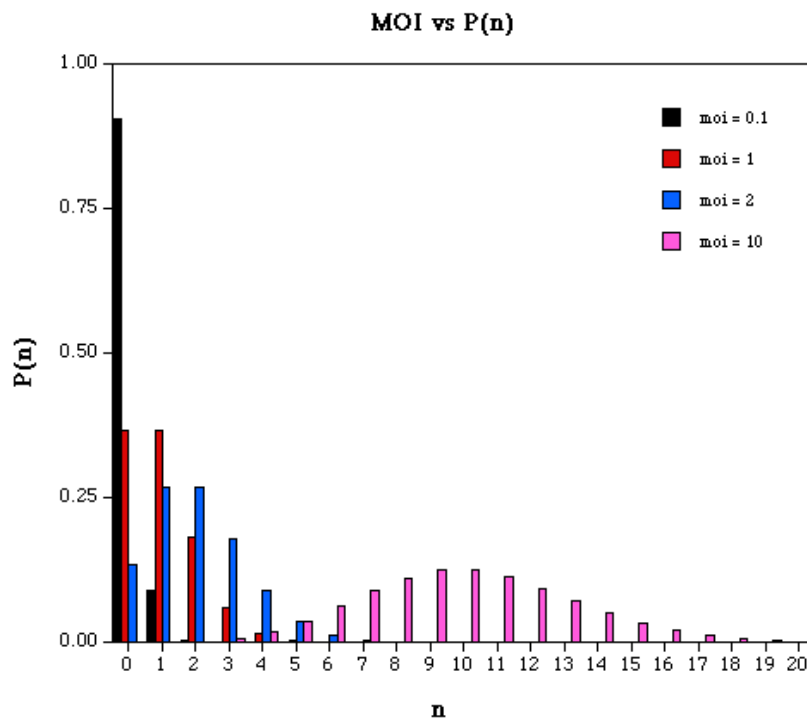


Figure 3.3: Proportion of cells getting infected under different MOIs

Hence, the cell infection is modeled as a stochastic process in which the number of viruses infecting the cell follows a Poisson distribution with mean of m ($m = \text{MOI}$). The MOI is an input variable in the model which is specified at the start of the simulation.

3.2 Rate-limited Approach

In the rate-limited approach we developed the model of HIV intracellular replication based on the rates inferred from literature. These rates include: Reverse transcription, nuclear transport, integration, activation, mRNA transcription, exporting to cytoplasm and translation rates. The list of parameters and rates is summarized in Section 3.2.2. In this approach the cell infection and viral replication is modeled as a stochastic process. The model is time-driven and time advances uniformly by a constant time-step during the simulation.

3.2.1 Model Entities

In this approach cells are individual entities in the model. Each cell has associated a specific state and a set of variables. The internal variables of the cell are a measure of the intracellular quantities and the internal state of the cell (See section 3.1.1) changes based on these variables. The cell internal variables are: VRNA (viral RNA), VDNA (viral DNA in both cytoplasm and nucleus), PROVIRUS (integrated cDNA), MRNA (mRNA transcribed), VP(translated viral protein). At the start of the simulations all variables are set to 0 and the number of viral RNAs infecting the cell is assigned to VRNA. The number of viral RNA is taken from a Poisson distribution as specified in Section 3.1.2. Figure 3.4 shows the internal variables and rates changing these variables inside a single cell. Squares are the cell internal variables that we keep track of in the simulation and arrows are their rates of change. An incoming arrow shows an increase in the variable value, while an outgoing arrow shows a decrease. The change in each quantity (or internal variable) is based on both the rate of its production and the amount of quantity that is producing it.

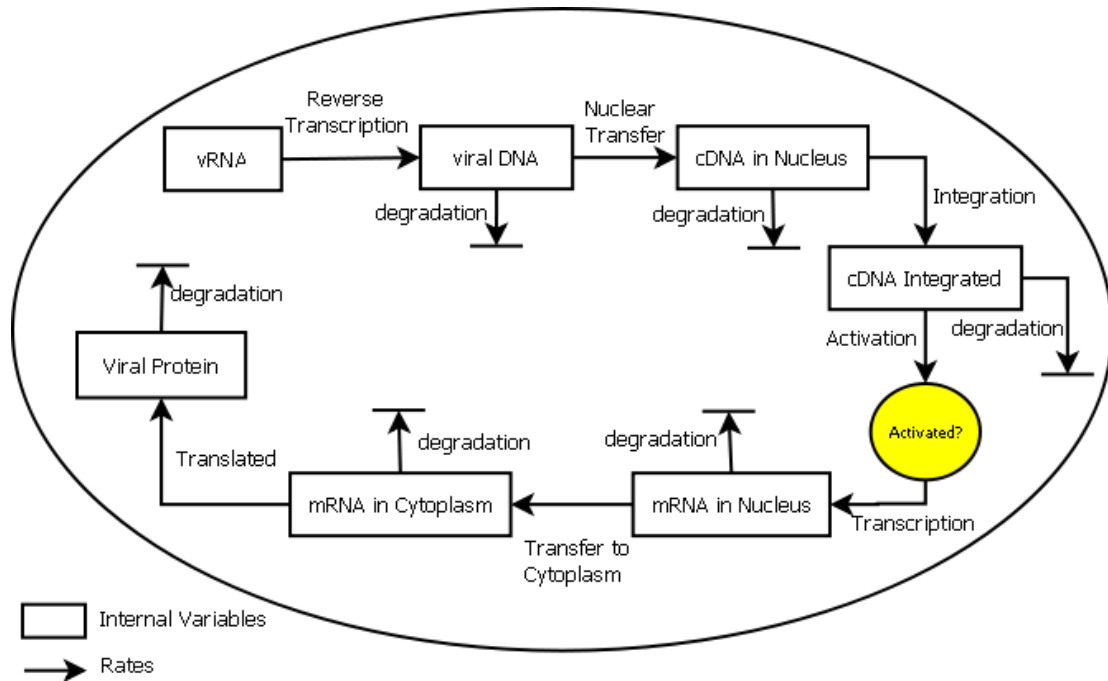


Figure 3.4: Cell Internals variables and intracellular rates. Squares are the cell internal variables that we keep track of in the simulation and arrows are their rates of change. An incoming arrow shows an increase in the variable value, while an outgoing arrow shows a decrease.

3.2.2 Parameters and Assumptions

The initial population of cells simulated and the value of MOI are input data in the simulation. Input data are specified by the user at the start of the simulation. We chose a time step of 0.2 minute (12 seconds) in the simulation. This time is small enough to capture the level of detail we are looking at and long enough for completion of chemical reactions occurring that we are trying to model such as DNA synthesis and transcription. The total simulation time is 2.5 days which is approximately the average life-span of an infected cell. A full list of rates and parameters used in this approach is presented in Table 3.1.

Each rate is rescaled to the time step of the simulation and is used as a probability of occurrence for the event. In what follows we will briefly explain the use of each parameter

Table 3.1: Model Parameters (Rate-limited approach)

| Parameters | Value | Reference |
|---|--|-----------|
| Reverse transcription synthesis rate (k_{rt}) | 120 nucleotide sec^{-1} | [19] |
| Reaction efficiency for RT first strand DNA | 0.45 | [18] |
| Reaction efficiency for RT second strand DNA | 0.40 | [18] |
| Activation rate ($k_{activation}$) | $3 \times 10^{-3} day^{-1}$ | [39] |
| Integration rate ($k_{integration}$) | 4.6 copies h^{-1} | [23] |
| Transcription rate ($k_{transcription}$) | 1000 transcripts h^{-1} | [23] |
| Translation rate ($k_{translation}$) | 262 proteins h^{-1} | [23] |
| Export rate from cytoplasm | 2.6 copies h^{-1} | [23] |
| Transfer rate to nucleus | 0.012 copies h^{-1} | [21] |
| Half life of mRNA | 0.2 h^{-1} | [23] |
| Elimination rate of protein | 1.4×10^{-5} proteins sec^{-1} | [40] |

in the model.

Reverse Transcription: The simulation starts with reverse transcription if the cell is already infected. We assume that reverse transcription process is fast enough and starts once the cell gets infected by the virus. Reverse transcription occurs with the rate of $k_{rt} = 120$ nucleotide per second ¹. As mentioned, the reverse transcription process mainly consists of 3 steps: first-strand DNA synthesis with reaction efficiency of 45%, second-strand DNA synthesis with reaction efficiency of 40% and the full-length double-strand DNA formation and vRNA degradation. We have considered these 3 steps in modeling reverse transcription.

Activation: In the previous model of HIV intracellular replication by Reddy et al. [23] all cells were considered to be in an activated state at the time of infection. In our model we distinguish between activated (actively infected) and inactivated cells (latently infected). After the provirus is integrated into the host cell DNA, the cell is in a latently infected stage. A latently infected cell may become activated with an activation rate of $k_{activation} = 3 \times 10^{-3}$

¹Nucleotide is a common length unit for single strand RNA

per day. At every time step we check for the cell activation. As soon as the cell is activated, transcription process is initiated in the model.

Transcription: Initiation of mRNA transcription is obtained by cellular factors and cell activation [41]. Activation of the cell is important for a latently infected cell to produce infected mRNA molecules by transcription. There are three major classes of transcripts: unspliced, singly spliced, and multi-spliced. In our simulation, we are not modeling the splicing event of mRNA transcripts and only consider the full-length mRNAs. Splicing plays regulatory roles and is required for efficient mRNA transcription. It basically adds a delay to the process of transcription which is balanced in our simulation by the time delay added for the cell activation. After cell activation, the cell produces mRNA transcripts at every time step with an average transcription rate of $k_{transcription} = 1000$ transcripts per hour.

The modeling flowchart of the virus intracellular replication cycle (in the rate-limited approach) is presented in Figure 3.5. The function *Replication* is executed at each iteration in the simulation. Details of the replication function are shown in the flowchart with a remark that the function checks the occurrence of an event at each step by taking a random number between 0 and 1. The event occurs if the random number is within the range of the event rate, otherwise it goes to the next step.

3.3 Diffusion-limited Approach

In the diffusion-limited approach we are modeling the HIV intracellular replication as in the rate-limited approach. This diffusion-based approach models the movement and spatial effects of quantities inside the cell. Entities in this approach are intracellular compartments which have a significant role in viral replication and they are modeled as agents. We have defined a 2-dimensional grid of 100×100 for placing the cell. The position of the cell is

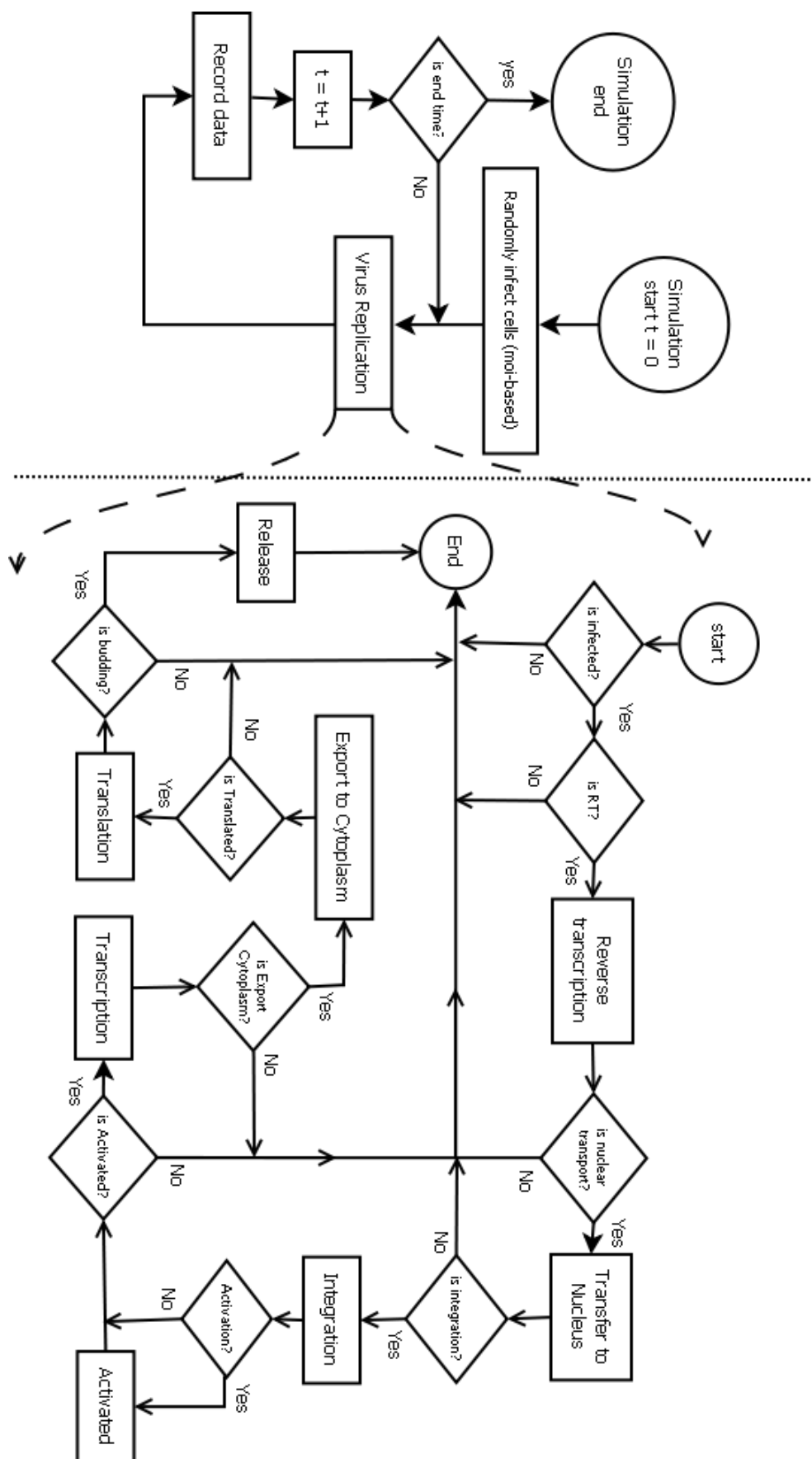


Figure 3.5: Flowchart of the rate-limited model showing the simulation of HIV intracellular replication.

assumed to be constant during the simulation run. The simulation is time-driven and at each time step the moving particles move one step, either left, right, up or down, within the cell area (Finite size effects are not studied). Events occur upon collision between agents and collisions will be processed by the agent-collision rules listed in Table 3.3. The modeling agents are explained in Section 3.3.1. At each iteration all agents move one step and check for collisions.

The advantage of this approach is that by modeling the intracellular entities as agents we take into account spatial effects that cannot be represented in the rate-limited model.

3.3.1 Model Entities

The modeled cellular entities are the cell DNA, nucleus, cytoplasm (cell area) and tRNA (transfer RNA), while the modeled viral entities are viral RNA, DNA, mRNA and viral proteins. The modeling entities are listed as following:

- **Cell:** The cell has a position and a size. It is placed on the center of the simulation grid and its position remains unchanged during the simulation. Other intracellular particles move within the cell area.
- **Cellular nucleus:** The cell nucleus has a size and a fixed position in the center of the cell.
- **Cellular DNA:** The cell DNA has a size and a fixed position in the center of the cell nucleus.
- **Cellular tRNA:** transfer RNA inside the cell cytoplasm binds to primary binding site (PBS) of viral genome. tRNA acts as a primer and is essential for the initiation of reverse transcription process.
- **Viral RNA:** HIV particle consists of 2 viral RNAs which are released inside the cell after infection. A random walk is defined for vRNA inside the cytoplasm area of the cell.

- **Viral DNA:** Viral DNA is the product of reverse transcription. Viral DNA randomly move inside the cell area(both the cytoplasm and the nucleus).
- **Integrated DNA:** Viral cDNA is the viral DNA which is transferred to the Nucleus and can be integrated into the host cell DNA.
- **Viral mRNA:** Viral mRNA is produced in the nucleus and can be exported to the cytoplasm. mRNAs are randomly moving in both the nucleus and the cytoplasm.
- **Viral protein:** Viral proteins are produced in the cytoplasm by translation of viral mRNAs.

Figure 3.6 shows a snapshot of the model visualization at the early stages of the simulation. The modeling entities that are visualized are: the cell, the nucleus and the cellular DNA which are all positioned on the center of the grid, while viral RNAs (red) and cellular tRNAs (purple) are randomly positioned in the cell cytoplasmic area. Figure 3.7 shows a snapshot of the model at the late stages of the simulation where viral mRNA (small green) is transcribed from the cell DNA inside the nucleus and viral proteins (big green) are translated in the cell cytoplasm.

3.3.2 Parameters and Assumptions

In this section we will briefly describe the parameter choices and initial values. The time step used in this approach is exactly the same as in the rate-limited approach (each time step = 0.2 minute). Also the MOI is an input data to the simulation and is specified at the start of the simulation. We chose the cell radius equal to 40 grid points (gp). The average T cell radius is between 4 to 4.5 μm [42]. Considering the cell size 4 μm each grid point would be 0.1 μm . In T lymphocytes the diameter of the cell is approximately twice of its nucleus. Hence, we used the value of 2 for the cell to nucleus size ratio [42, 43], so the nucleus radius would be 20 (gp). If the T cell DNA is 0.3 μm , with each grid point 0.1 μm , the DNA diameter would be 6 gp (radius = 3 gp). The amount of tRNA in the cell is a

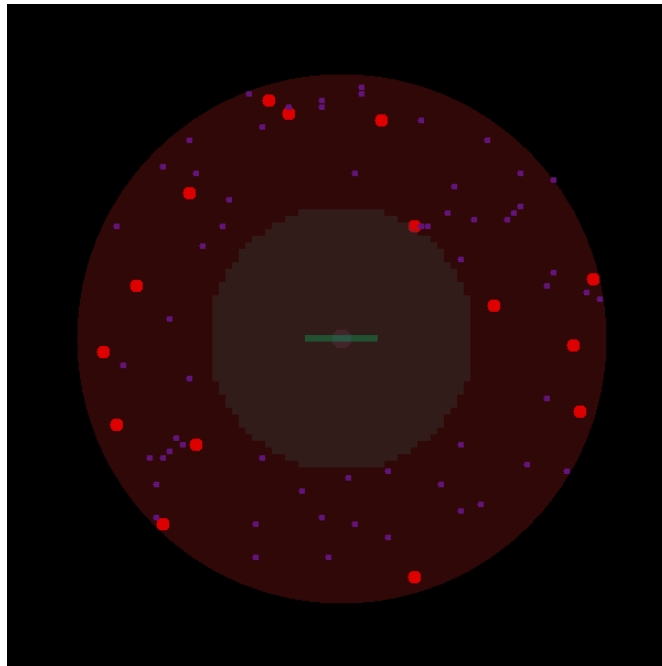


Figure 3.6: Simulation visualization of early stages of simulation: viral RNAs (red) are assigned to the cell cytoplasm environment. The purple circles are the cellular tRNA available in the cell

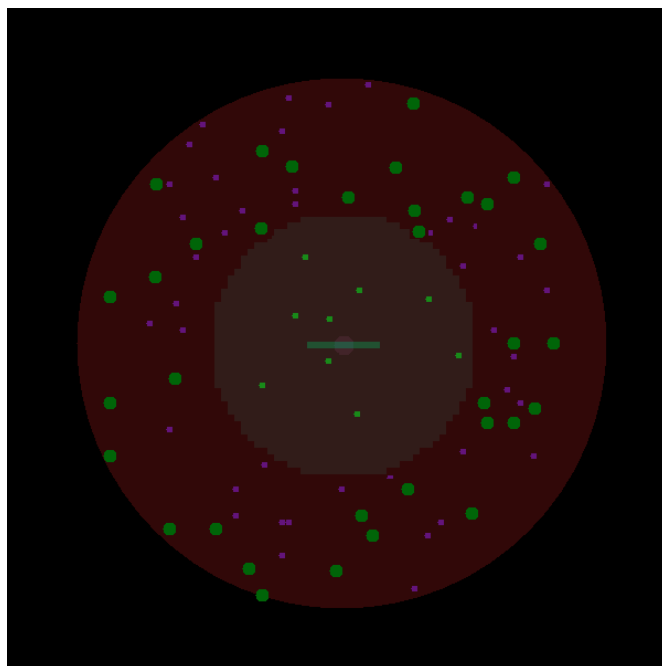


Figure 3.7: Simulation visualization of late stages of simulation: viral mRNA (small green) mainly in the cell nucleus, and translated viral proteins (big green) in the cell cytoplasm. The purple circles are the cellular tRNA available in the cell

high value (198000 per cell [42]). We chose a relatively high value per surface area of the cell. The cell activation, transcription, and translation rate is the same as the rate-limited approach. After integration the cell may get activated with rate $k_{activation}$. If the cell is activated, mRNA transcription starts with rate $k_{transcription}$ and transcribed mRNAs are transported to the cell cytoplasm. In the cytoplasm mRNAs are translated into the viral proteins with rate $k_{translasion}$. Table 3.2 shows the list of parameters, constants and initial values.

The movement of viral particles inside the cell is known to be based on both diffusion and moving along micro-tubes [44]. Movement along micro-tubes leads to a faster transfer of particle toward the cell nucleus. In this simulation we are modeling the movement of particles only based on the diffusion. Diffusion coefficient of particles is tuned to compensate this assumption in the model.

Table 3.2: Model Parameters (Diffusion-limited approach)

| Parameter | Value | Description |
|--|--|-------------|
| Activation rate ($k_{activation}$) | $3 \times 10^{-3} \text{ day}^{-1}$ | [39] |
| Transcription rate ($k_{transcription}$) | 1000 transcripts h^{-1} | [23] |
| Translation rate ($k_{translasion}$) | 262 proteins h^{-1} | [23] |
| Half life of mRNA | 0.2 h^{-1} | [23] |
| Elimination rate of protein | $1.4 \times 10^{-5} \text{ proteins sec}^{-1}$ | [40] |
| Diffusion coefficient of particles inside the cell | $2.08 \times 10^{-3} \mu\text{m}^2 \text{ s}^{-1}$ | tuned |
| Initial values and constants | | |
| Grid size | 100×100 | assumed |
| Cell radius | 40 gp | assumed |
| Nucleus radius | 20 gp | derived |
| Cellular diameter | 6 gp | derived |
| Initial number of tRNA | 70 per cell surface area | derived |

The modeling flowchart of the virus intracellular replication cycle (in the diffusion-limited approach) is presented in Figure 3.8. At start of simulation the cell is getting infected with a certain number of viruses, randomly taken from a Poisson distribution, based on MOI (For more on initial infection based on Poisson distribution see 3.1.2).

Then the agents are assigned to the simulation grid. At each iteration every moving agent performs a one step random walk. After that, each agent checks for the collision with other entities and process it based on agent-collision rules (See Table 3.3). At the end of each iteration if there are no more collisions to be processed, it goes to the next step.

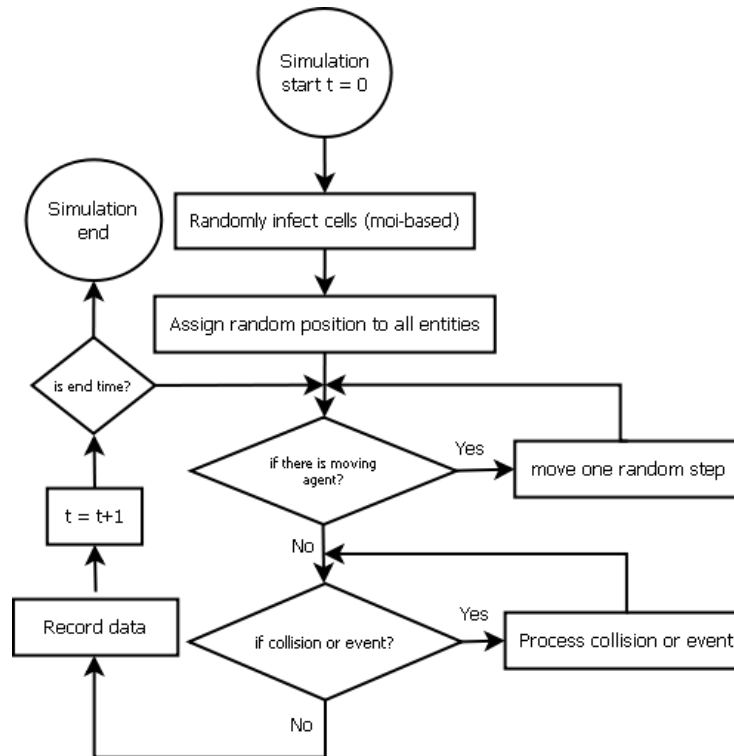


Figure 3.8: Flowchart of the diffusion-limited approach showing the main loop of the simulation

Table 3.3: Agent Collision Rules

| IF COLLISION(agent1 - agent2) | THEN (event) | Rules |
|-------------------------------|-------------------------|---|
| vRNA - tRNA | Reverse Transcription | Produce vDNA and vRNA degradation |
| vDNA - cell nucleus membrane | Nuclear Transfer | vDNA transfer from the cytoplasm to the nucleus(cDNA) |
| cDNA - cell DNA | Integration | produce integrated cDNA |
| mRNA - cell nucleus membrane | Export to the cytoplasm | cDNA export from the nucleus to the cytoplasm |

3.4 Model Implementation

The model is implemented in Java programming language for both approaches. We have used MASON (A Multi-Agent Simulation library core in Java [45]) libraries and 2D visualization tools in our simulation. The simulation is time-driven and the update scheme is by time advance. The simulation time advances uniformly by a constant time-step at each iteration. The result data of each cell at each run are recorded in a Hash Table at certain time points. These times can be changed at start of the simulation according to the interest of the experiment. The data saved in the history files are used later for analyzing the results in MATLAB.

3.5 Summary

In this chapter first the general model design 3.1, cell states and transitions, and the cell infection are described. The two different approaches used to implement the model are explained in detail: Rate-limit approach 3.2 and diffusion-limit approach 3.3. At the end we explained the model implementation in both approaches.

Chapter 4

Simulation Results, Analysis, and Validation

In this simulation we have modeled a single cycle of virus replication inside an infected cell. The cell is getting infected by a certain number of viruses and undergoes viral replication. Cell infection is a stochastic function whose parameter is the initial MOI (See Section 3.1.2). Although we are modeling intracellular processes at a single cell level, we are analyzing the results over a population of cells. This is closer to what happens in in-vivo and in-vitro experiments such as cell in a ml volume of the blood or a petri dish. Hence, in a population of cells infected with a certain MOI, some cells are getting infected and some not. The final results were obtained by running the simulation for a population of 1000 cells and averaging over several populations.

At each run we compute the first replication cycle up to 60 hours post infection, corresponding to approximately 2 and a half days, which is the average life-span of an infected cell. At each run a set of cells is getting infected with different virus numbers and cells are not interacting with each other. Using this model we analyzed each step of the viral replication cycle by tracking the following entities: cDNA integration into host cell DNA,

mRNA transcription after cell activation, and number of translated viral proteins produced by each infected cell after the first round of replication. Results are mainly validated based on the number of cDNA integrated into the host DNA and the amount of mRNA transcripts produced per cell. Validation of results is based on the number of cDNA molecules integrated into the host cell DNA, and the number of mRNA transcribed after cell activation. In what follows we will discuss the simulation results of the rate-limited approach (Sections primary results 4.1 and sensitivity analysis 4.2), and a comparison between the two approaches is being made in the section 4.3.

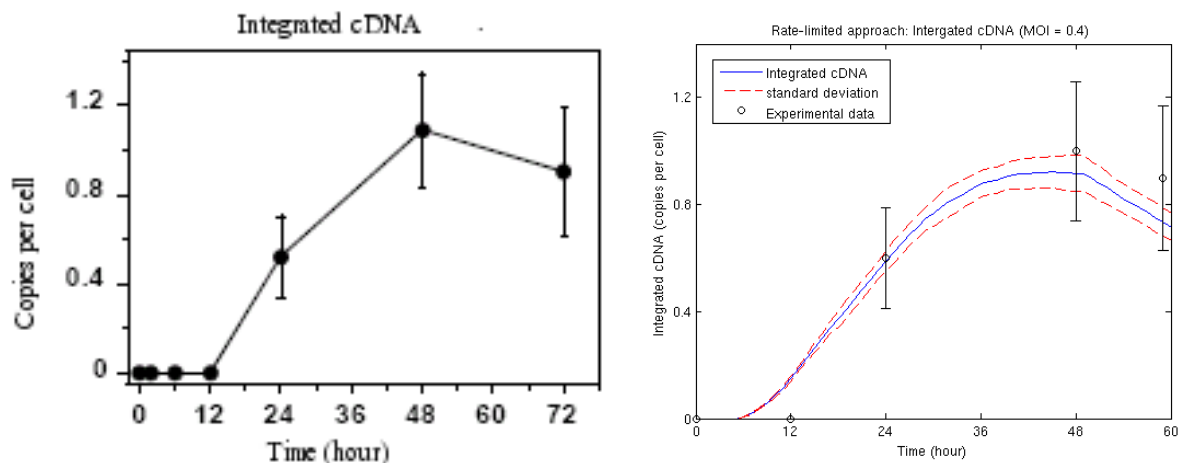
4.1 Primary Results and Validation

At the start of the simulation the initial number of cells is set to 10^3 and cells are randomly getting infected based on MOI. As mentioned earlier, we didn't model the entry process and cell infection is a stochastic process. For each virus particle that infects the cell, 2 viral RNAs are added to the cell. If a cell is infected, reverse transcription occurs with the rate specified in Table 3.1. The viral RNA is reverse transcribed and viral DNA is synthesized. After production of viral DNA in the cytoplasm, it is transferred to the nucleus and gets integrated into the cell DNA with the rate specified in Table 3.1.

4.1.1 Integration

Integration of viral DNA into the host cell DNA is a key step in the virus replication cycle. If integration doesn't occur, the virus is incapable to fully replicate and produce new copies of viruses through the cell. Reverse Transcription generates DNA in the cytoplasm, the concentration of this DNA in the cytoplasm decreases due to the transfer of viral DNA to the nucleus. The DNA transferred to the nucleus is rapidly integrated into the host genome, circulated or degraded. We have taken into account the integration and

degradation of viral DNAs in the nucleus. DNA circularization is neglected in our model because it is very limited [22]. In experiments conducted by Scott et al. [22] detection of integrated proviruses was accomplished using fluorescence-monitored PCR. They used the $\text{MOI} = 0.4$ in their experiments and limited their measurements to a single cycle of HIV replication. The concentration of the integrated cDNA rises and approaches its maximum value (1.02 copies per cell at $\text{MOI} = 0.4$) about 24 hour post infection. The simulation results of integrated cDNA is consistent with the experimental results and exactly follows the same trend as experimental data.



(a) Experimental results of copies per cell of the integrated cDNA in DNA metabolism after infection of Average T cells with HIV-based vectors
(b) simulation results of integrated cDNA and the standard deviation over 10 runs of 1000 cells

Figure 4.1: Copies of integrated cDNA per cell at $\text{MOI} 0.4$

In PCR-based experiments a low MOI is used because of the high sensitivity of the PCR technique [19]. We also have used a low multiplicity of infection ($\text{MOI} = 0.4$) and fitted the simulation results with experimental data presented in Figure 4.1. Then we run more simulations for lower and higher MOIs such as 0.1, 0.4, 1 and 10 to determine the maximum value and the trend of integrated cDNA into the host cell DNA (Figure 4.2). From these simulations we observe that only the availability of DNA molecules limits the actual integration. This observation is consistent with experimental results in [24].

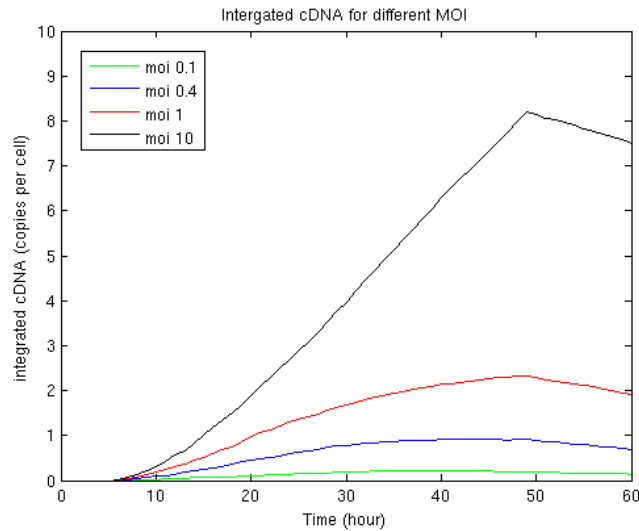


Figure 4.2: Simulation results for lower and higher MOIs (0.1, 0.4, 1 and 10) show the trend of integrated cDNA into the host cell DNA per cell

Nuclear Transfer

The value we used for the nuclear transfer parameter in our model is 0.012 (copies per hour) [3.1]. This value is 10 times less than the value used in the mathematical model by Reddy et al. 1999 [23]. Figure 4.3 compares the amount of integrated cDNA for two different nuclear transfer rates. The blue line is the simulation results using the experimental value from [21] for nuclear transfer parameter (0.012 copies per hour), while the red line is the same result using the parameter value (0.12 copies per hour) from the previous model by Reddy et al. 1999 [23]. Our model shows that using the experimental value from [24] has better results than using the value reported in Reddy et al. The sensitivity analysis for the nuclear transfer parameter is presented in the Sensitivity Analysis section 4.2.

Transcription

Viral mRNA is the product of transcription process in an infected cell. After integration cells are getting activated with the cell activation rate specified in Table 3.1, subsequently

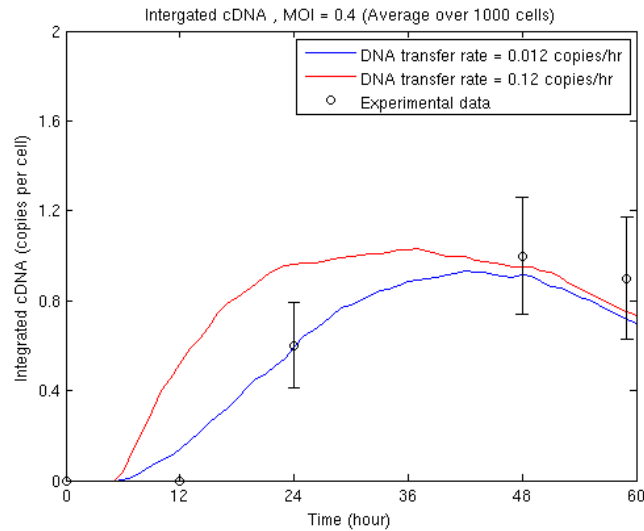


Figure 4.3: Integrated cDNA with two different values reported for nuclear transport rate. The blue line is the simulation results using the experimental value (0.012 copies per hour) from [21], while the red line is the same result using the reported value (0.12 copies per hour) from the previous model by Reddy et al. 1999 [23]

transcribing new viral mRNA. We have measured the average copies of mRNA transcripts over 1000 cells at different MOIs and compared it with experimental data on genomic HIV-1 mRNA molecules [21]. In experiments conducted by Barbosa et al. [21], mRNA of 22 patients with different viral loads were measured and plotted in Figure 4.4 (a). Progressors are patients with high viral loads and Non progressors are patients with lower viral loads. We have chosen the MOI range from 0.01 to 100 and run the simulation for certain values in this range. Then we plotted the average amount of mRNA transcripts produced in each run versus the value of MOI. Each data point in Figure 4.4 (b) is an average over 1000 cells. From the plot we can see that the average amount of mRNA produced per cell in the simulation is within the same range as the data collected from patients with different viral loads.

Figure 4.5 shows the average of mRNA production in the simulation over time. Different colors of each line shows the average mRNA produced in 1000 cells and different colors represent different values of MOI (Green line is MOI = 0.1, blue line is MOI = 0.4,

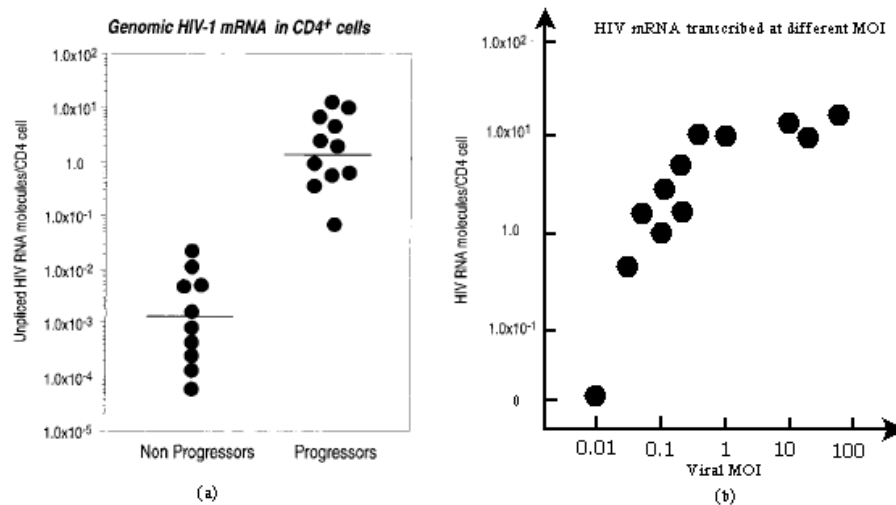


Figure 4.4: Expression of viral mRNA (a) Number of HIV mRNA molecules in peripheral blood CD cells for the nonprogressor and progressor patients. (b) copies of mRNA produced per cell after transcription at different MOIs. Each data point is an average over 1000 cells

red line is MOI = 1, and black line is MOI = 10). Although the transcription rate in the simulation is equal for all runs the initiation time of transcription differs for different MOIs. As shown in Figure 4.5 the average time required for cell activation is shorter in runs with higher MOI. Also, more cells are getting activated over time and produce more mRNA transcripts.

Translation

Viral mRNAs transcripts are exported with the export rate $k_{translation}$ specified in Table 3.1, from the cell nucleus to the cytoplasm. In the cytoplasm they are translated into viral proteins required to produce new viruses.

Figure 4.6 shows the production of viral proteins in the simulation over time. Each line represents the average of translated viral proteins over 1000 cells. Different colors represent different values of MOI (Green line is MOI = 0.1, blue line is MOI = 0.4, red line is MOI = 1, and black line is MOI = 10).

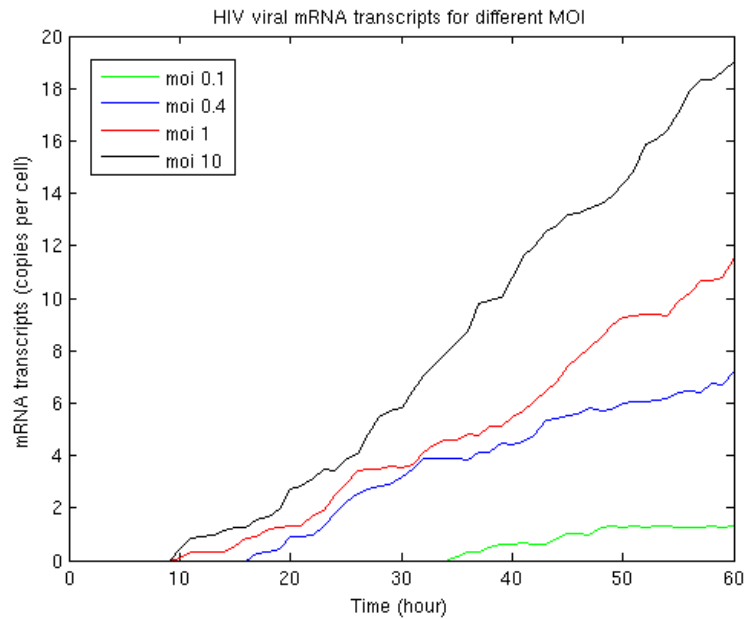


Figure 4.5: mRNA transcripts at different MOI. Each line represents the average over 1000 cells. Different colors represent different values of MOI (Green line is MOI = 0.1, blue line is MOI = 0.4, red line is MOI = 1, and black line is MOI = 10)

4.2 Sensitivity Analysis

In mathematical and computational models sensitivity analysis is performed to determine the output sensitivity or uncertainty due to small perturbation in the parameter values. In this model which simulates the intracellular replication of HIV, we applied sensitivity analysis to three key parameters of our model. For each parameter we analyzed the output of all the steps in replication that are influenced by its changes. For example, the integration parameter will influence the simulation outputs at integration, transcription and translation steps, while the transcription parameter will only affect the output of transcription and translation.

Results presented in this section are from the rate-limit approach as we found it computationally more efficient for doing the sensitivity analysis. This analysis helped us to determine which steps are more sensitive in the viral replication process. This results also

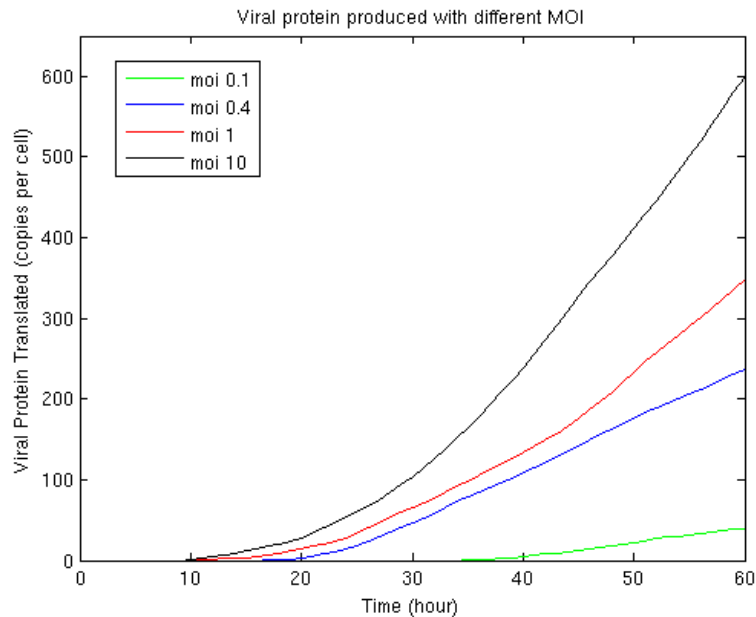


Figure 4.6: Viral protein translated at different MOI. Each line represents the average over 1000 cells. Different colors represent different values of MOI (Green line is MOI = 0.1, blue line is MOI = 0.4, red line is MOI = 1, and black line is MOI = 10)

indicate which parameters may be critical in the replication process. There is a total of 15 parameters in the model. We have chosen the following parameters to perform the sensitivity analysis:

- **Nuclear Transfer rate**
- **Integration rate**
- **Transcription rate**

Results of the simulations with the original parameter values, called *Basic case*, are presented in the section Primary Results. 4.1). For each selected parameter we specified a range by choosing a minimum and a maximum for the parameter value. Each parameter is then varied in uniform increments within the specified range. While a parameter is being varied, all other parameters and input data are kept unchanged at their literature values. At each variation we averaged the results over 1000 simulations and analyzed how the system behaves due to changes in the parameter. A noteworthy point in doing this sensitivity

analysis is that the statistical distribution and standard deviation of parameters are not required. Hence, only the mean value of the parameter and the minimum and maximum values are necessary for the sensitivity analysis.

Nuclear Transfer Rate

The first parameter we chose for sensitivity analysis was the nuclear transfer rate ($k_{transfer}$) which is the rate of viral DNA transferring from the cytoplasm to the nucleus. The experimental value for this parameter is 0.012 (copies/hour). We assumed the minimum and maximum value of this parameter to be 0.0012 (copies/hour) and 0.022 (copies/hour).

Figure 4.7, shows the variation in the trend of the output (integrated cDNA) for runs using different nuclear transport rates. We started with the minimum value of the parameter and used a constant increment of 0.002 until it reached its maximum value. The * line is the amount of integrated cDNA for the experimental value of the parameter. The dashed lines show the amount of integrated cDNA for parameter values lower than the mean, while the solid lines refer to higher values. Figure 4.7 clearly shows how for values lower than the experimental nuclear transfer rate the amount of integrated cDNA decreases significantly, while there is no significant change for the higher values.

This means that the output result of integrated cDNA is more sensitive to the decrease of the nuclear transfer parameter than to its increase. This conclusion is also visible in the next graph 4.8 which represents the sensitivity of integration to the percentage of changes in the nuclear transport parameter. Different colors in the graph shows the sensitivity of the output (integrated cDNA) at 3 time points 24, 48, and 60 hours post infection. The horizontal axis of this plot is in terms of Percent of Range. 0% of range represents the minimum value of the parameter, 100% of range represents its maximum value, and 50% of range represents its mean value (experimental value).

From sensitivity analysis we can see that the output of integration is not sensitive to an

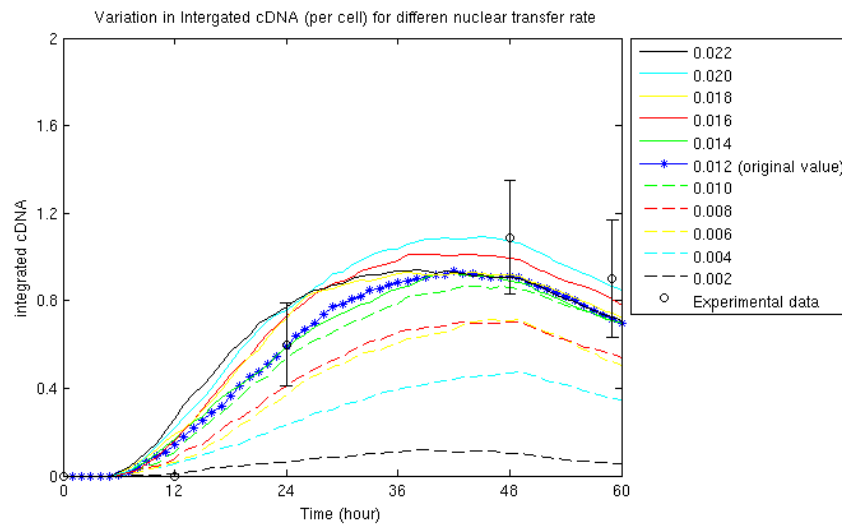


Figure 4.7: Variation in copies of integrated cDNA (per cell) for different nuclear transfer rate. The * line is the amount of integrated cDNA for the experimental value of the parameter. The dashed lines show the amount of integrated cDNA for parameter values lower than the mean, while the solid lines refer to higher values. For values lower than the experimental nuclear transfer rate the amount of integrated cDNA decreases significantly, while there is no significant change for the higher values

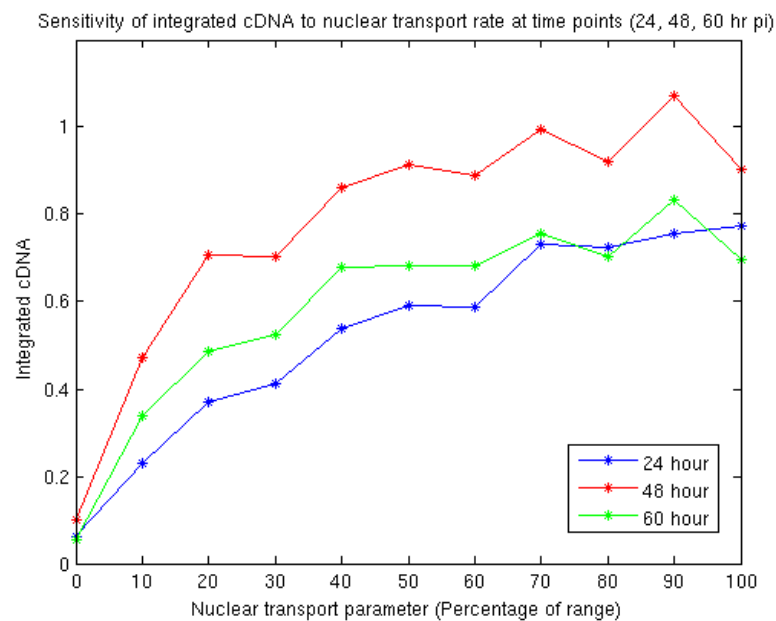


Figure 4.8: Change of Integrated cDNA at time points of 24, 48, 60 hour post infection. The percentage of parameter range is from 0% (the minimum value or parameter = 0.0012) to 100% (the maximum value = 0.022). 50% represents the original value of the parameter = 0.012.

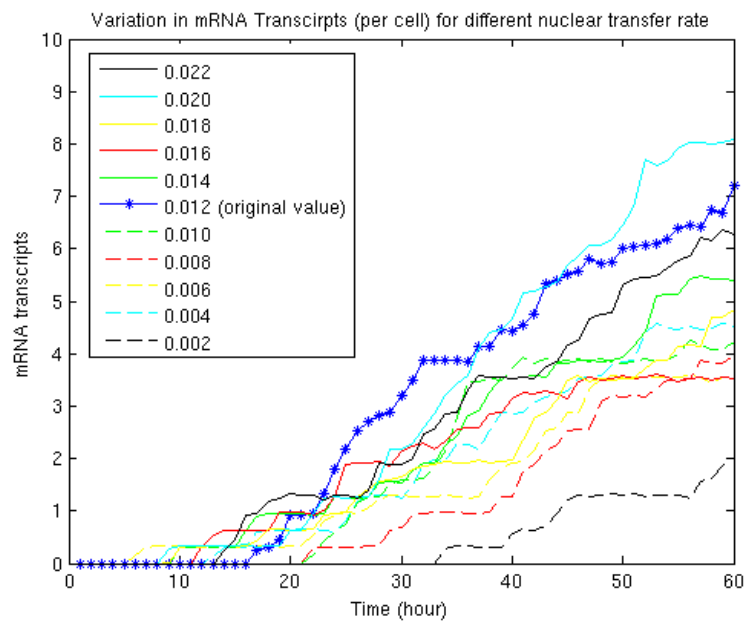


Figure 4.9: Variation in copies of mRNA transcribed (per cell) for different nuclear transport rates. The * line is the amount of mRNA transcripts for the experimental value of the parameter. The dashed lines show the amount of mRNA transcripts for parameter values lower than the mean, while the solid lines refer to higher values.

increase of the nuclear transport rate. The value that is used as the mean parameter value has almost the same effect as the maximum value we assumed for this parameter. Hence, this shows that the process of DNA transfer to the nucleus is a very efficient process in our simulation, because the output of this process is not sensitive to the increase of nuclear transfer rate. In other words, the nuclear transfer rate measured in the experiment is sufficiently high to transfer all the cDNA into the nucleus of the infected cell. This result is consistent with the experimental observations that the transport of viral DNA to the nucleus is active and rapid, and all the viral DNA molecules synthesized in cytoplasm are transported to the nucleus 24 hours post infection [21].

Figure 4.9 shows the change in the trend of mRNA production over time for variation in the nuclear transport parameter. The * line is the amount of mRNA transcripts for the original (mean) parameter value. The dashed lines show the amount of mRNA transcripts for parameter values less than the mean and the solid lines show the amount of mRNA transcripts for parameter values more than the mean value. As expected, the amount of mRNA transcripts in the simulation decreases for lower values of transport parameter, while there is no significant change in this quantity by increasing the transport parameter. This result confirms the previous conclusion that the output of simulation in different stages of replication cycle is sensitive to decrease of the transport parameter, while not sensitive to the increase of this parameter.

Integration Rate

Figure 4.10, shows the variation in the trend of the output (integrated cDNA) for runs using different integration rates. The mean value used for integration parameter is 4.6 copies per hour [3.1]. We started with the minimum value 0.6 (copies per hour) and at each variation there is a constant increment of 1 in the parameter value until it reaches the maximum value 9.6. The * line is the amount of integrated cDNA for the original (mean)

parameter value. The dashed lines show the amount of integrated cDNA for parameter values less than the mean and the solid lines show the amount of integrated cDNA for values more than the mean value. We see that variation in Figure 4.10 shows that variations in the integration parameter are not significantly changing the output of the model on integrated cDNA. Subsequently, variation in this parameter will not significantly change the output of the model on mRNA transcripts production and translated viral proteins (results not shown). Hence, we conclude that the replication process is not sensitive to the integration parameter rate, and significant changes in the output is only applicable if we change the integration parameter in the order of magnitude of 10.

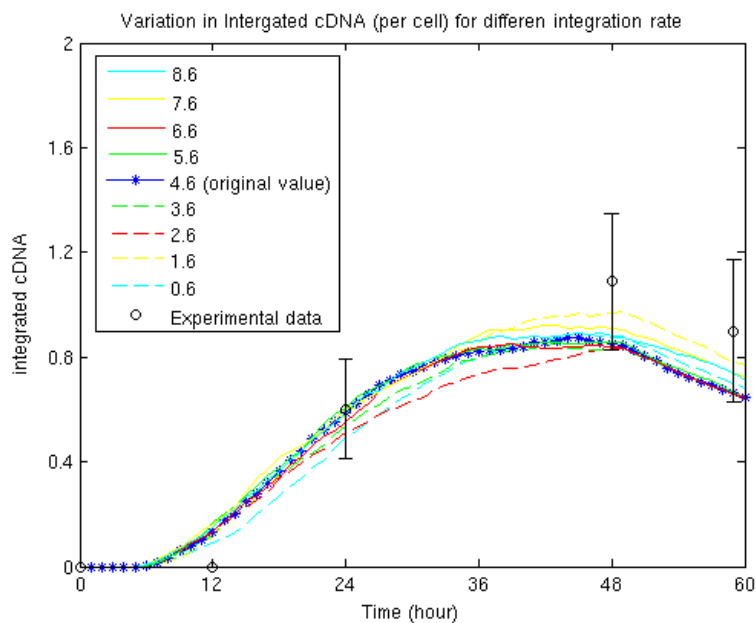


Figure 4.10: Variation in copies of integrated cDNA per cell for different integration rates. The * line is the amount of integrated cDNA for the experimental value of the parameter. The dashed lines show the amount of integrated cDNA for parameter values lower than the mean, while the solid lines refer to higher values.

The graph 4.11 represents the sensitivity of integration to the percentage of changes in the integration parameter. Different colors in the graph shows the sensitivity of the output (integrated cDNA) at 3 time points 24, 48, and 60 hours post infection. The horizontal

axis of this plot is in terms of Percent of Range. 0% of range represents the minimum value of the parameter, 100% of range represents the maximum value of each variable, and 50% of range represents the mean value (or the original value) of the parameter. The analysis of this figure also shows clearly that the model is not sensitive to changes in the integration rate.

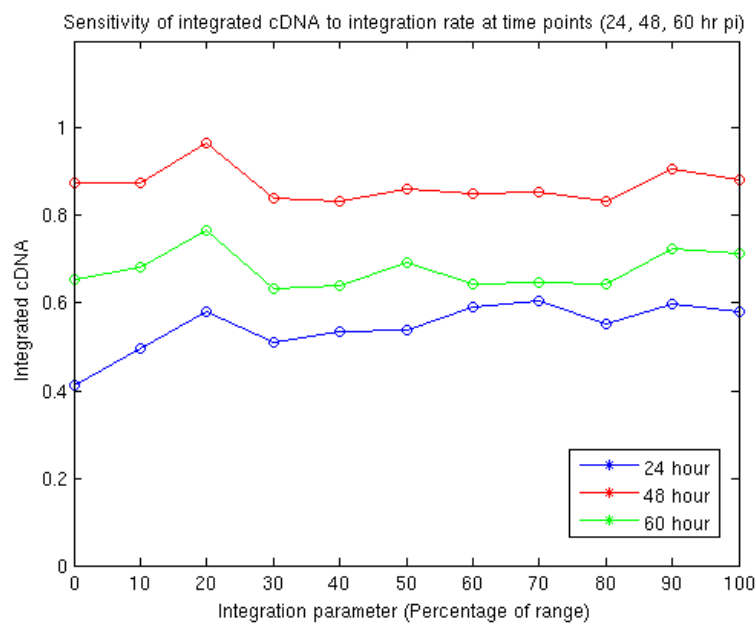


Figure 4.11: Change of Integrated cDNA at time points of 24, 48, 60 hour post infection. The percentage of parameter range is from 0% (the minimum value or parameter = 0.046) to 100% (the maximum value = 9.8). 50% represents the original value of the parameter = 4.6.

Transcription Rate

In the sensitivity analysis of transcription rate we have analyzed the outputs of mRNA transcripts and translated viral protein. mRNA transcripts are the direct outcome of transcription process and its quantity is obviously affected by variation in the transcription parameter (data not shown). We also analyzed the sensitivity of translated viral proteins to this parameter. Results are presented in Figure 4.12 show the variation in the value

and the trend of viral proteins translated in the simulation over time. The * line is the amount of viral proteins for the experimental value of transcription parameter, while the dashed lines and solid lines represent the same result for values lower and higher than the parameter's experimental value. transcription rates

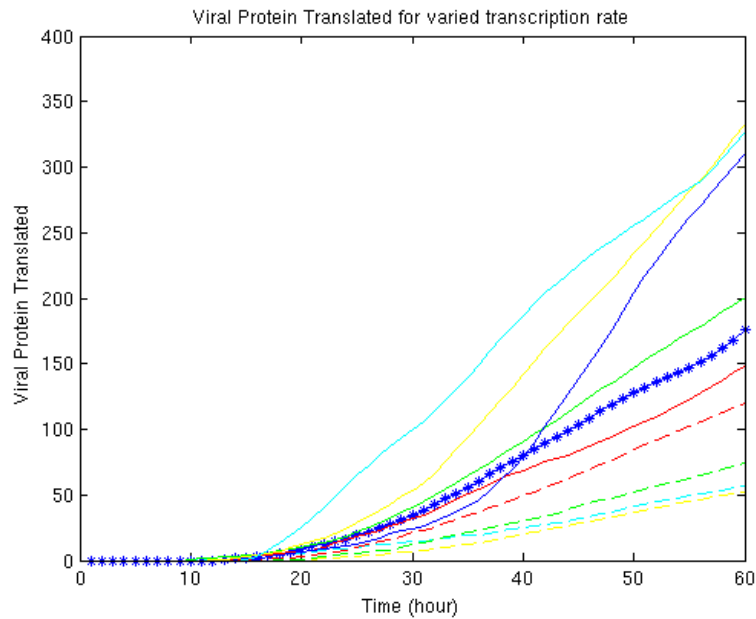


Figure 4.12: Variation in copies of viral proteins per cell for different transcription rates. The * line is the amount of viral protein for the original transcription parameter, while the dashed lines and solid lines represent the same result for values lower and higher than transcription rates

The mean experimental value used for transcription parameter is 3.3 transcripts per simulation time step (or 1000 transcripts per hour 3.1). We started with the minimum value 0.8 (transcripts/step) and at each variation there is a constant increment of 0.5 in the parameter value until it reaches the maximum value of 5.5 transcripts per step.

The graph 4.13 shows the sensitivity of viral proteins to the percentage of changes in the transcription parameter. Different colors in the graph shows translated viral proteins at 3 time points 24, 48, and 60 hours post infection. The horizontal axis of this plot is in terms of Percent of Range. 0% of range represents the minimum value of the parameter,

100% of range represents the maximum value of each variable, and 50% of range represents the experimental value of the parameter. As you can see in Figure 4.13, the blue line (24 hr pi) is smoother, while, the green line (60 hr pi) is the steepest. Hence, variation in viral proteins for changes in transcription rate is less in 24 hours post infection in compare to 48 and 60 hours post infection, indicating that viral protein synthesis is almost completely absent in the first 24 hours of infection.

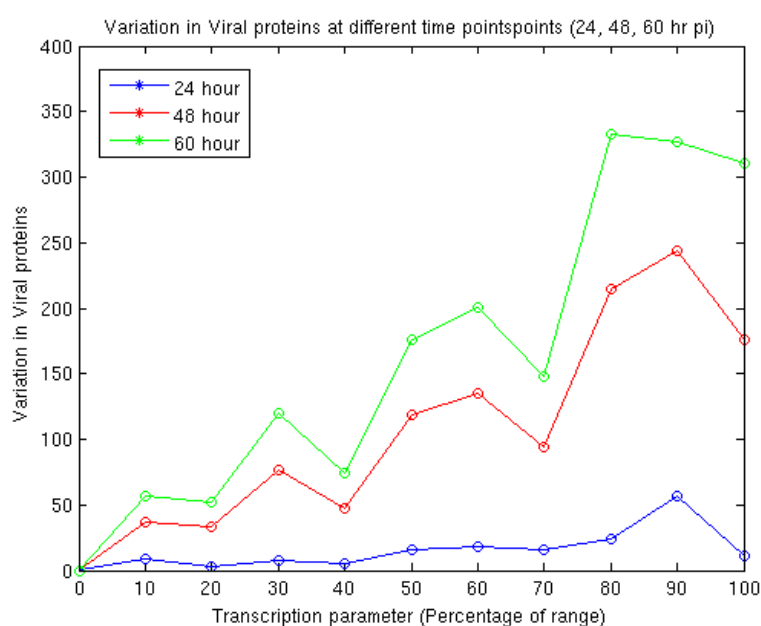


Figure 4.13: Change of viral proteins at time points of 24, 48, 60 hour post infection. The percentage of parameter range is from 0% (the minimum value or parameter = 0.046) to 100% (the maximum value = 9.8). 50% represents the original value of the parameter = 4.6.

4.3 Comparison of the Two Approaches

The diffusion-limited approach is based on the diffusion of particles and not just rates inferred from literature. The viral DNAs synthesized are transferred to the nucleus by random walk. The amount of integrated cDNA in this approach is illustrated in Figure

4.14. As it is seen, the diffusion-limit approach is slightly over estimating the amount of integrated cDNA. However, the mean value is still in the range of standard deviation of the experimental data.

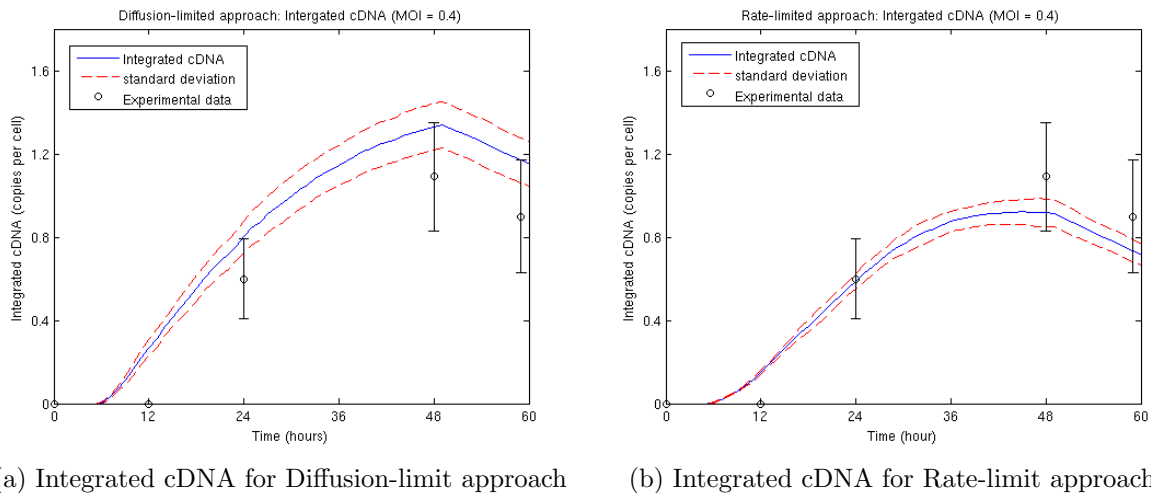


Figure 4.14: Simulation results of integrated cDNA for the two approaches: Diffusion-limited approach and Rate-limited

The same overestimation also happens in the range of mRNA transcripts produced in the diffusion-limited approach. This overestimation could be a consequence of modeling the movement of the entities inside the cell only based on diffusion. As mentioned earlier, the movement of viral particles inside the cell is more likely to be along micro-tubes together with diffusion [44]. Moving along tubes limits the interaction of entities in the cytoplasm as it is speeding up the transfer of particles to the nucleus. Since the effect of moving along micro-tubes is not modeled in our simulation more interactions between entities happen in the cytoplasm which leads to the production of more entities. This leads to the overestimation in the simulation results of the diffusion-limited approach.

The advantages of the diffusion-limit over the rate-limited approach was that fewer number of parameters were used and the spatial information was taken into account.

4.3.1 Simulation Performance

Results of the simulation are saved into history files over time. To improve the performance of the simulation we reduced the statistics collection or the amount of data that is collected and saved into the history files. This means that the data are collected and displayed only at certain time points that are interesting to the experiment. As mentioned earlier, in the results presented in this thesis the data are collected every one hour (or 300 time steps). To have simulation results for large amount of cells, we run the simulations on the Lisa cluster [46] requesting 2 nodes and 2 cores per node. In rate-limited approach, at $MOI = 0.4$, the total time required to run the simulation for 1000 cells is 3.383 minutes (or 230 sec). So, the time it takes to run a single cell is on average 0.203 seconds. The same times were measured in diffusion-limited approach with the same MOI . The time it takes to run 1000 cells is 8.866 minute (or 514 sec). Hence, the time required for running a single cell equals to 0.514 seconds on average. From this result, we can see that the rate-limited approach is approximately 2.532 times more computationally efficient.

Table 4.1 compares the two approaches used for implementing the model. Advantages and disadvantages of each approach are discussed as following:

Table 4.1: Comparison of Approaches: Rate-limited vs. Diffusion-limited

| | Rate-limited Approach | Diffusion-limited Approach |
|----------------------|--|---|
| Model | Based on rates inferred from literature | Based on Diffusion of particles |
| Spatial effects | No | Yes |
| Implementation | Each step is triggered by previous steps, time-driven (time step 0.2 minute) | Agent-collision rules, time-driven (time step 0.2 minute) |
| Number of parameters | 11 | 6 |
| Simulation time | 0.203 sec (average per cell) | 0.514 sec (average per cell) |

4.4 Summary

The results chapter contains three main sections: Primary Results and Validation 4.1, Sensitivity Analysis 4.2 and Comparison of the Two Approaches 4.3. Primary results section presents the simulation results of the rate-limited approach (Basi case with the rates from literature). At each run we computed the first replication cycle up to 60 hours and validated the results by comparing them with the experimental data. In sensitivity analysis section 4.2 we performed a sensitivity analysis for some key parameters in the rate-limit model: Nuclear transport, Integration and Transcription rates. Results of this analysis showed that replication process is more sensitive to some parameters such as nuclear transfer and transcription rates, while it is less sensitive to parameters such as integration rate. Section Comparison of the Two Approaches 4.3 compares the results of the diffusion-limited approach with those of the rate-limited and compares the two approaches we used for implementing the simulation. At the end we compare the performance and time of the simulation in each approach.

Chapter 5

Conclusions and Future Work

5.1 Summary of Contributions

In this thesis, we first described the HIV structure and the intracellular process of HIV infection and replication and introduced the existing models of the viral replication inside the cell. After that we presented our model of HIV intracellular replication and introduced the two approaches for its implementation: Rate-limited approach and diffusion-limited approach. In the rate-limited approach cells are defined as individual entities and the intracellular dynamics of each cell is based on rates from literature. In the diffusion-limited approach cells are individual entities as well but the dynamics is based on real quantities modeled as agents. Hence, in this approach the spatial information is taken into account. We presented the basic case of the simulation results for the rate-limited approach and validated them against the experimental data. Moreover, a sensitivity analysis for three key parameters was performed. From the sensitivity analysis results we concluded that the replication process is not so sensitive to the integration parameter, while it is sensitive to the transcription parameter and the decrease of nuclear transport rate. In the end, we compared the two approaches used in this model and discussed the advantages and

disadvantages of each approach.

Cell infection and viral replication is a stochastic process and it is modeled as such in both approaches. By looking at individual cells we get a more realistic description of the dynamics inside the cell. This is a more accurate description of the in-vivo and in-vitro experiments.

5.2 Recommendations for Future Work

For future directions, the model can be adapted or extended mainly in two different ways: single-cell level or population-level.

First the model can be extended to simulate reinfection and interaction between cells in a population. For example cell-cell transmission of the virus is due to virological synapse that forms between an infected and an uninfected T cell. The model can also be enhanced to simulate multiple rounds of replication. So far, our model focuses on the first round of replication, while it can be adapted to have a second, third or multiple rounds of replication. Both approaches we used to implement the model can be easily expanded to achieve these goals.

Moreover, the diffusion-limited approach can be extended at a single cell level to focus more on the behavior of the cell under viral replication. A more detailed description of the replication processes such as the movement of particles along micro-tubes or the regulatory effects of mRNA splicing will allow a better understanding of the intracellular mechanisms. Studying the spatial and finite size effects is also interesting in the modeling content and the diffusion-limited approach can be easily adapted to achieve these goals.

Both approaches in modeling HIV intracellular replication as a whole can be used for investigating the effect of various HIV inhibitors or new drug agents on the replication process. Effect of various drugs can be modeled either in the form of rates or real agents. This

will help doctors and pharmacologists for measuring the efficiency of drugs in inhibiting the viruses.

Bibliography

- [1] “Overview of the global aids epidemic,” Joint United Nations Programme on HIV/AIDS, Tech. Rep., 2006.

- [2] P. W. Nelson A. S. Perelson, “Mathematical analysis of hiv-1 dynamics in vivo,” *SIAM REVIEW*, vol. 41, p. 344, June 1999. [Online]. Available: http://www.math.lsa.umich.edu/~pwn/SIAM_review.pdf

- [3] P.M.A. Slood, “Virolab: A virtual laboratory for decision support in viral disease treatment.” [Online]. Available: <http://virolab.org/>

- [4] Parker C.E. Nowak M.A. Stekel, D.J., “A model of lymphocyte recirculation.” *Immunology Today*, vol. 18, pp. 216–221, 1997.

- [5] Sumen C. Reddy T.E. Alber M.S. Lee P.P Casal, A., “Agent-based modeling of the context dependency in t cell recognition.” *Journal of Theoretical Biology*, vol. 236, pp. 376–391, 2005.

- [6] Tay J.C. Guo, Z., “A comparative study of modeling strategies of immune system dynamics under hiv-1 infection,” in *Lecture Notes in Computer Science, Vol. 3627. Springer-Verlag, Banff, Alberta, Canada (2005) 220-233*, 2005.

- [7] C. Boucher Sloot P.M.A., F. Chen, “Cellular automata model of drug therapy for hiv infection,” *ACRI*, pp. 99–116, 2002. [Online]. Available: <http://www.science.uva.nl/research/scs/papers/archive/Sloot2002d.pdf>
- [8] Biaoru Li, “A strategy to identify genomic expression at single-cell level or a small number of cells,” *Electronic Journal of Biotechnology*, vol. 8-issue1, 2004.
- [9] David C. Chan and Peter S. Kim, “Hiv entry and its inhibition,” *Cell Press*, vol. 93, pp. 681–684, 1998.
- [10] McDougall BR Robinson WE Jr. Victoria JG, Lee DJ, “Replication kinetics for divergent type 1 human immunodeficiency viruses using quantitative sybr green i real-time polymerase chain reaction,” *AIDS Res Hum Retroviruses*, vol. 19(10), pp. 865–74, 2003.
- [11] Jean-Louis Sankale Seema Thakore Meloni Souleymane Mboup Adam MacNeil, Abdoulaye Dieng Sarr and Phyllis Kanki, “Direct evidence of lower viral replication rates in vivo in human immunodeficiency virus type 2 (hiv-2) infection than in hiv-1 infection,” *Virology*, vol. 81, p. 53255330, 2007.
- [12] [Online]. Available: <http://www.info.gov.hk/aids/pdf/g190htm/01.htm>
- [13] “Hiv and aids research.” [Online]. Available: <http://www.avert.org/hiv-types.htm>
- [14] HOWARD M. TEMIN LOUIS M. MANSKY, “Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase,” *Journal of Virology*, p. 50875094, 1995.
- [15] J J Rossi L Scherer and M S Weinberg, “Progress and prospects: Rna-based therapies for treatment of hiv infection,” *J Gene Therapy*, vol. 14, pp. 1057–64, 2007.

- [16] G. Ertaylan and P.M.A. Sloot, “A complex automata model of hiv-1 co-receptor tropism: Understanding mutation rate pressure,” in *Reviews in Antiretroviral Therapy*, Washington D.C., USA,, 2007.
- [17] A. Saib Nisole S., “Early steps of retrovirus replicative cycle,” *Retrovirology*, 2004.
- [18] Adam Davis-Christopher Burrell Peng Li David Warrilow, Luke Meredith and David Harrich, “Cell factors stimulate human immunodeficiency virus type 1 reverse transcription in vitrotriangledown,” *Journal of Virology*, pp. 1425–1437, 2008.
- [19] Burrell CJ Karageorgos L, Li P, “Stepwise analysis of reverse transcription in a cell-to-cell hiv infection model: kinetics and implications,” *General Virology*, vol. 76, pp. 1675–1686, 1995.
- [20] Berkhout B. Abbink TE, “Hiv-1 reverse transcription initiation: A potential target for novel antivirals?” *Virus Research*, vol. 134, pp. 4–18, 2008.
- [21] Dumey N-Clavel F Barbosa P, Charneau P, “Kinetic analysis of hiv-1 early replicative steps in a coculture system,” *AIDS Res Hum Retroviruses*, vol. 10(1), pp. 53–9., 1994.
- [22] Mark S.T. Hansen & Frederic D. Bushman Scott L. Butler, “A quantitative assay for hiv dna integration in vivo,” *Nature Medicine*, vol. 7, pp. 631 – 634, 2001.
- [23] J. Yin B. Reddy, “Quantitative intracellular kinetics of hiv type 1,” *AIDS RESEARCH AND HUMAN RETROVIRUSES*, vol. 15, pp. 273–283, 1999.
- [24] Anja Habermann Ina Allespach Jacomine Krijnse-Locker Hans-Georg Krusslich Sonja Welsch, Oliver T. Keppler, “Hiv-1 buds predominantly at the plasma membrane of primary human macrophages,” *PLoS Pathogens*, 2007.
- [25] P. Mendes, “Biochemistry by numbers: simulation of biochemical pathways with gepasi 3,” *Trends. Biochem. Sci*, vol. 22, pp. 361–363, 1997.

- [26] H. Sauro, “Jarnac: a system for interactive metabolic analysis. animating the cellular map 9th international bio thermo kinetics meeting,” *ISBN*, pp. 0–7972–0776–7, 2000.
- [27] Korin YD Scripture-Adams DD Zack JA-Kreisberg JF-Roederer M Sherman MP Chin PS Goldsmith MA Eckstein DA, Penn ML, “Hiv-1 actively replicates in nave cd4 t cells residing within human lymphoid tissues,” *Immunity*, vol. 15, no. 3, pp. 671–682, 2001.
- [28] [Online]. Available: <http://www.wisconsinlab.com/hiv.htm>
- [29] P.M.A. Sloot;P.V. Coveney;G. Ertaylan;V. Mller;C.A.B. Boucher and M.T. Bubak, “Hiv decision support: From molecule to man,” *Philosophical Transactions of the Royal Society*, vol. 367, pp. 2691–2703, 2009.
- [30] Ilkay Altintas Marian Bubak Charles Boucher Peter M.A. Sloot, Alfredo Tirado-Ramos, “From molecule to man: Decision support in individualized e-health,” IEEE Computer Society. [Online]. Available: <http://www.science.uva.nl/research/scs/papers/archive/Sloot2006b.pdf>
- [31] J. Summersy R. Srivastavawz, L.Youw and J.Yin, “Stochastic vs. deterministic modeling of intracellular viral kinetics,” *J. theoretical Biology*, vol. 218, p. 309321, 2002.
- [32] White T. Webb, K., “Cell modeling using agent-based formalisms.” in *AAMAS’04, New York, USA*, July 2004.
- [33] Eva Qwarnstrom Mike Holcombe Mark Pogson, Rod Smallwood, “Formal agent-based modelling of intracellular chemical interactions,” *BioSystems*, vol. 85, p. 3745, 2006.
- [34] Makkena R. McGeary F. Decker K. Grills W. Schmidt C. A Khan, S., “Multi-agent system for the quantitative simulation of biological networks,” in *AAMAS 03, Melbourne, Australia 385-392*, 2003.

- [35] Litorco J. Lee L. Jacob, C., “Immunity through swarms: Agent-based simulations of the human immune system.” in *In: Proceedings of 3rd International Conference on Artificial Immune Systems, Vol. 3239. Springer-Verlag GmbH, Catania, Sicily, Italy (2004) 400-412*, 2004, Committee Draft.
- [36] Han H.K. Tay J.C. Guo, Z., “Sufficiency verification of hiv-1 pathogenesis based on multi-agent simulation.” in *In: Proceedings of the Genetic and Evolutionary Conference, Vol. I. Washington D.C. (2005) 305-312*, 2005.
- [37] Jhavar Atul Tay Joc Cing, “Cafiss: A complex adaptive framework for immune system modeling,” in *Proceedings of ACM Symposium for Applied Computing Bioinformatics*, vol. vol 1, 2005, pp. 158–164.
- [38] “Multiplicity of infection (moi),” Stanley Maloy home page, College of Sciences, San Diego State University, November 2003. [Online]. Available: <http://www.sci.sdsu.edu/~smaloy/MicrobialGenetics/topics/phage/moi.html>
- [39] De Boer R Perelson AS, Kirschner DE, “Dynamics of hiv infection of cd4+ cells,” *Journal of Math Bioscience*, vol. 114, pp. 81–125, 1993.
- [40] Theo Theofanous Anh-Tuan Dinh and Samir Mitragotri, “A model for intracellular trafficking of adenoviral vectors,” *Biophysical Journal*, vol. 89, p. 15741588, 2005.
- [41] Takashi Okamoto and Flossie Wong-Staal, “Demonstration of virus-specific transcriptional activator(s) in cells infected with htlv-iii by an in vitro cell-free system,” *Cell Press*, vol. 74, pp. 29–35, 1986.
- [42] A L Hartman S H Swerdlow, P A Angermeier, “Nuclear morphology of follicular center cell-associated t-cells: an immunoultrastructural study of follicular hyperplasia and follicular center cell lymphomas,” *Modern pathology*, vol. Vol. 1 Issue 4, pp. 268–73, 1988.

- [43] P.M.A. Sloot and C.G. Figdor, “Elastic light scattering from nucleated bloodcells: Rapid numerical analysis,” *Applied Optics*, vol. 25, p. 3559, 1986.
- [44] Lucero G Svitkina TM Borisy GG Emerman M Hope TJ. McDonald D, Vodicka MA, “Visualization of the intracellular behavior of hiv in living cells,” *Journal of Cell Biology*, vol. 159(3), pp. 441–52, May 2002. [Online]. Available: <http://www.cs.wisc.edu/condor/manual/v6.8/>
- [45] [Online]. Available: <http://www.cs.gmu.edu/~eclab/projects/mason/>
- [46] “Lisa cluster.” [Online]. Available: http://www.sara.nl/index_eng.html
- [47] Martin Markowitz John M. Leonard David D. Ho Alan S. Perelson, Avidan U. Neumann, “Hiv-1 dynamics in vivo: Virion clearance rate, infected cell lifespan, and viral generation time,” *Science*, vol. 271. no. 5255, pp. 1582 – 1586, May 1996.
- [48] Dennis Bray, *Cell Movement: From Molecule to Motility*, 2nd ed. Garland, 2001. [Online]. Available: <http://www.citebase.org/abstract?id=oai:arXiv.org:cs/0103025>
- [49] Narendra M. Dixit Catherine Beauchemin and Alan S. Perelson, “Characterizing t cell movement within lymph nodes in the absence of antigen,” *The Journal of Immunology*, vol. 178, pp. 5505–5512, 2007. [Online]. Available: http://globus.org/toolkit/docs/development/4.2-drafts/key/GT4_Primer_0.6.pdf
- [50] A. S. Perelson Ho D. D., A. U. Neumann, “Rapid turnover of plasma virions and lymphocytes in hiv-1 infection,” *Nature* 373, vol. 373, p. 117, May 1995.
- [51] Seehra JS Richardson CC Huber HE, McCoy JM, “Human immunodeficiency virus 1 reverse transcriptase. template binding, processivity, strand displacement synthesis, and template switching,” *Journal of Biological Chemistry*, vol. 264(8):, pp. 4669–78, 1998. [Online]. Available: <http://gdp.globus.org/gt4-tutorial/>

- [52] S. Havlin H.E. Stanley G.H. Weiss Larralde H., P. Trunfio, “Number of distinct sites visited by n random walkers,” , *Phys. Rev. A* 45 (1992) 7128, vol. 45, pp. 7128 – 7138, 1992.
- [53] Theodoropoulos G Logan, B., “The distributed simulation of multi-agent systems,” in *Proceedings of the IEEE*, vol. 89, 2001, pp. 174–185.
- [54] Chiara Biancottoa Anna Knezevich Anna Kula Marina Lusic Mariacarolina Dieudonn, Paolo Maiuri and Alessandro Marcello, “Transcriptional competence of the integrated hiv-1 provirus at the nuclear periphery,” *The EMBO Journal*, vol. 10.1038, p. 141, 2009. [Online]. Available: http://en.wikipedia.org/wiki/Business_Process_Execution_Language
- [55] F. C. Neidhardt, *Escherichia coli and Salmonella: Cellular and Molecular Biology, Vol. 1. ASM Press 1996.*, Frederick C. Neidhardt, Ed. Washington, D.C. : ASM Press, 1996. [Online]. Available: citeseer.ist.psu.edu/foster02physiology.html
- [56] J. Nichols M. Mallen A. Pou D. Asmuth R. Pollard Nokta M. A., X. Lio, “Chemo-line/cd4 receptor density ratios correlate with hiv replication in lymph node and peripheral blood of hiv-infected individuals, aids 2001.” *AIDS*, pp. 161–169, July 2001.
- [57] A.S. Perelson and P.W. Nelson, “Mathematical analysis of hiv-1 dynamics in vivo,” *SIAM Review*, vol. 41, pp. 3–44, 1999.
- [58] Stafford A T Das and B Berkhout, “Modeling efficient extension of a misaligned trna-primer during replication of the hiv-1 retrovirus,” *Nucleic Acids Res*, vol. 23(8), p. 13191326, 1995.
- [59] Cao Y Daar ES Ho DD Perelson AS Stafford MA, Corey L, “Modeling plasma virus concentration during primary hiv infection,” *Theoretical Biology*, vol. 203(3), pp. 285–301, 2000.

- [60] Douglas Thain and Miron Livny, “Building reliable clients and servers,” in *The Grid: Blueprint for a New Computing Infrastructure*, Ian Foster and Carl Kesselman, Eds. Morgan Kaufmann, 2003. [Online]. Available: <http://www.cs.wisc.edu/condor/doc/grid2-ch19.pdf>
- [61] Ltd. Wisconsin Viral Research Group, “Physician’s corner: Hhv-6 and hiv,” October 2001. [Online]. Available: <http://www.wisconsinlab.com/hiv.htm>