Agent based modeling of viral infections: an investigation across several spatio-temporal scales

Mancini, E.

Publication date
2014

1. Introduction

The past century has been characterized by the fastest growth of both scientific and technological knowledge in human history. This quest to understand the laws of nature and its complex phenomena has led, among other things, to the formulation of quantum physics, the birth of informatics, and the sequencing of the human genome. The scientific progress of the 20th century turned into technological applications that, for better or for worse, reshaped life on the planet.
Although the 20th century vastly increased the quality of life in many countries, it also brought us new and more challenging problems. The technological and scientific growth has strongly influenced our society and has been accompanied by dramatic changes, such as the exponential growth in the world population and energy consumption [1,2], the doubling of the average life expectancy [3], the growth of large metropolitan areas and world airline traffic, and globalization. The latest change in our society has been Internet and the widespread use of electronic devices that have exponentially increased the communications and connections among individuals across the planet [4].

The major problems we have to face today and the evolution of our society under the pressure of new technologies are strongly correlated [5]. Some of these challenges are closely correlated to the exponential growth of the population worldwide and of the growth of metropolis and megacities, e.g. traffic and pollution in cities, worldwide energy consumption and the increased risk of pandemics. It is not surprising that as a society becomes more complex also its problems grow in complexity. Luckily new knowledge and tools are developed to study and understand these new challenges as soon as they are identified.

1.1 Background and Motivation

Among all the complex problems of modern society, the study of infectious diseases is clearly one of the most important. Both ethical and economical issues demand solutions to cure and prevent the

devastating effects that viral infections have on a society so tightly connected and densely populated as ours. One of the most studied viruses is Human immunodeficiency virus (HIV), a virus of the Retroviridae family. In the past 10 years HIV-1 has caused about 2 million deaths per year and the latest UNAIDS report [6] estimates that about 35.3 million people are currently infected. Although the development of the antiretroviral therapy has greatly increased the life expectancy of HIV infected individuals, the number of AIDS related deaths was still 1.6 million in 2012. A huge amount of resources has been invested globally on studying all the different aspects of the HIV infection: from the molecular and intracellular level, to the dynamics of its spread in patients and in countries. Several other viruses have been studied due to their fast spreading mechanisms (airborne transmission) like the influenza viruses and the SARS coronavirus (SARS-CoV) responsible for the severe acute respiratory syndrome. The "Spanish flu" caused over 50 million deaths worldwide in 1918 and the H1N1 influenza pandemic of 2009 was a reminder of their continued health risk and significant economic impact [7]. In 2003 an outbreak of SARS caused 774 deaths over 8098 reported cases, with a high fatality rate of 9.6%. New research projects are being funded to investigate the different scales of threat, from the epidemiological spread of a virus in a population to its early phases of the infection in a cell culture.

The emergence of drug resistant viral strains and the high genetic variability of both HIV and influenza viruses have shown that these solutions need to be dynamic, since any static solution risks to be weakened and eventually nullified by the evolution of the viruses under the selective pressure of drugs. The knowledge that modern biology and medicine have of the complex interplay between viruses, our immune system and drugs has shown the need to understand this adaptive complex system in order to provide long-term dynamic solutions. An example of this complexity is the emergence of HIV mutants resistant to anti-retroviral drugs and the ability of the influenza virus to mutate and reassort its genome that force us to constantly work on drugs and new vaccines. Only a system able to

simulate the biological mechanisms of this complex system and predict how viruses react to our strategies against them will allow our society to minimize the risk of future pandemics.

1.2 Complex adaptive systems

Scientific research in the 20th century has proven that nature, despite its relatively simple fundamental laws, is too complex to be understood by only examining its individual constituents. The complexity hidden in many natural phenomena has shown the limitations of a purely reductionist approach and the rise of a new philosophical approach to science well summarized by a quote from Aristoteles: "the whole is more than the sum of its parts".

Despite more than 60 years have passed since Warren Weaver's definition of disorganized and organized complexity [8], the definition of complexity is still controversial. A universal measure of complexity has not been defined yet, and complexity is often well defined in relation to a system or a specific field. The reason is that when considering a specific measurable quantity it is easier to first intuitively grasp and later define when one system is more complex than another. All complex systems consist of a large set of interconnected, mutually (and typically nonlinearly) interacting parts [9] and their macroscopic behavior is an emergent property of the individual interactions. John Holland defined complex adaptive systems as "systems that have a large numbers of components, often called agents, that interact and adapt or learn" [10]. For each complex system it is correct to say that the whole is more than the sum of its parts: conscience cannot be understood by the electrical impulses of a single neuron, a single water molecule cannot be perceived as "wet" and the immune response cannot be described by the interaction of two individual immune cells. Thus emergent properties are possessed by the whole system and not by any of its individual parts. Due to this characteristic it is necessary to study the system both in terms of its components and in terms of their interactions.

1.2.1 The new challenge

In order to study complexity and to create tools for managing complex problems the reductionist approach used so far to study nature and its fundamental laws would not allow for an appropriate level of understanding. According to reductionists any system, no matter how complex, is nothing but the sum of its parts and by examining its individual constituents even emergent phenomena could be understood. Historically using a minimal set of simple principles has succeeded in understanding and predicting relatively simple systems. This top-down approach studies complex systems in terms of its microscopic constituents and was sufficient to drive our understanding of nature from the birth of science until the 20th century. Recently the need to understand more complex problems led to the development of new approaches: connectionism, a bottom-up, data-driven, approach for which a system is a synthesis of its constituent parts (e.g. artificial neural-networks) and collectivism which merges the understanding of the constituent parts of a system and their interactions that characterize the system as a whole [ref. book]. It is in this collectivist and deeply interdisciplinary approach that rests the key to understand complexity and complex systems. A good example of the connectionist approach is the work of Soren Brunak and Ole Lund on the use of artificial neural networks applied to the prediction of T-cell epitopes [11]. An artificial neural network is a computational model in which simple artificial nodes called "neurons" are connected together to form a network. The neurons are the processing elements and the strength of the connections between neurons is represented by a set of adaptive weights. Such weights are tuned using a learning algorithm and, together with the interconnection pattern between neurons, determine the machine learning and pattern recognition properties of artificial neural networks. In this approach neural networks are used as a black box representing the complexity of the binding of a T cell receptor with an antigen. Huge amount of data of known binding events between different sequences is used to train the model to make predictions on the binding affinity of new epitopes. The amino acid sequences are

used as input for the model with no attempt at understanding or representing the complexity of the protein folding that is instrumental to the binding of receptor and antigen. This approach does not aim at understanding the complexity of a given problem. Instead it maps that complexity into another complex system that is easier to handle, namely a complex set of neural connections. This new complex system is then treated as a black box to be trained with known data of inputs and outputs, initial conditions and outcomes. Once the training is successful and the mapping of a given complex problem is complete, this black box will be able to predict with good accuracy the behavior of that specific complex problem without the need to understand it.

1.3 Complexity of viral infections

Viral infections are caused by viruses that interact at molecular level with the receptors of a host cell. They hijack the molecular machinery of the infected cell, replicate, spread to other cells in the host organism, and then, depending on the specific type of virus, infect others organisms by different transmission methods (sexual transmission, direct contact, airborne transmission, etc.). Thus, the infection takes place at various spatio-temporal scales and it can be investigated at each of these scales.

The grand challenge is to understand infectious diseases across all these spatio-temporal scales: from the intracellular events happening within a few milliseconds at the molecular level up to the transmission of the infection within cities and countries at the epidemiological level. One of the most studied viruses is the human immunodeficiency virus (HIV). HIV is a positive-sense single-stranded RNA virus belonging to the Retroviridae family. HIV is a lentivirus responsible for the acquired immunodeficiency syndrome (AIDS). It damages the adaptive compartment of the immune system by slowly wiping out the $CD4^+$ T helper lymphocytes, disabling the adaptive immune response and leaving the infected patients without protection from opportunistic diseases. The nature of HIV infection, as that of all viral infections, is multiscalar, spanning from the molecular interactions

within the host cell to the replication dynamics in the infected individual immune system up to its transmission between individuals. HIV and its interactions with the immune system have the characteristics of a complex system at each spatio-temporal scale and as such it can be studied at each of these levels.

1.4 Modeling of viral infections

The human body is a complex adaptive system that includes an enormous amount of unique, discrete and hierarchical components: millions of molecules form a single cell, thousands of cells are combined into tissues, various tissues form organs, and all the organs function together to keep a human alive. Additionally human beings constantly interact with their environment and with each other in a society of more than 6 billion individuals. This hierarchical system of nonlinearly interacting components crosses many orders of magnitude in temporal and spatial scales. One of the greatest scientific challenges of 21st century is to understand, quantify and eventually replicate *in silico* this complex multiscale system and is the goal of the Virtual Physiological Human project. The progress done within this program on several organs, for example the heart and blood vessels, is now leading to clinical applications of the *in silico* models and to personalized, evidence-based medicine.

Despite the success in some areas, modeling of the immune system lags behind. Much still needs to be done, both in terms of biological experiments and computational modeling, before a realistic multi-scale *in silico* model of the immune system can be integrated in the clinical practice. There are several reasons for such delay: on one hand there is a general lack of experimental data on several components of the immune system. Much information is still missing on the interactions between all the cells and molecules of the immune system, no map of the lymphatic vessels is available, and little is known on the trafficking of immune cells between the complex network of lymph nodes, blood vessels and tissues; on the other hand the human immune system and its interactions with pathogens are, together with

the nervous system, the most complex system in our body, so even if all the information would be available, it would still be a challenge to develop a complete computational model.

The immune system is composed by a large number of different, nonlinearly interacting, discrete microscopic constituents that through self-organization lead to the emergence of a macroscopic behavior: the immune response. Agent-based modeling, mathematical modeling, medical research, and systems biology have to join their individual efforts to achieve a complete understanding of the immune response, and computational science is the tool needed to successfully integrate all these different fields.

1.4.1 From mathematical modeling to cellular automata

Mathematical modeling as a tool for the study of viral infections started in the early 1960s, if we neglect few exceptions. The first models focused on epidemics more than on the infection dynamics within the host with the hope to eradicate all infectious diseases with the use of the newly discovered antibiotics and vaccinations. [12]. Several years later the emergence of new infectious diseases like hepatitis and HIV and the progress in the understanding of molecular biology and immunology led to the development of models for the infection dynamics within host [13,14] and inside the infected cells [15].

The most commonly used models to investigate the disease dynamics are still based on ordinary differential equations. In models using ordinary differential equations discrete entities like cells and viruses and their individual properties are homogenized into parameters and variables that describe entire populations. They assume the system to be spatially homogeneous neglecting spatial characteristics that could play a nontrivial role in the disease dynamics. Models using partial differential equations take into account the spatial component but still represent discrete entities as a continuum [9].

The development of computer science and the diffusion of computers as tools for scientific investigation allowed cellular automata and agent-based models to become an established modeling paradigm to investigate viral infections at all spatio-temporal scales. Cellular automata models are spatially and temporally discrete. The first research on cellular automata dates back to 1948 when von Neumann introduced self-reproducing automata as an alternative approach to the study of biology. During the following 50 years research on CA slowly progressed until the late 80s and the foundation of the Santa Fe Institute, a center for interdisciplinary study of complex systems. Several different types of cellular automata have been developed with different characteristics depending on the problem in analysis. In the list below Ilachinski [9] reported the characteristic possessed by most CA models:

- *Discrete lattice of cells: the system substrate consists of a one-,two- or three-dimensional lattice of cells.*
- *Homogeneity: all cells are equivalent.*
- *Discrete states: each cell takes on one of a finite number of possible discrete states.*
- *Local interactions: each cell interacts only with cells that are in its local neighborhood.*
- *Discrete dynamics: at each discrete time unit, each cell updates its current state according to a transition rule taking into account the states of cells in its neighborhood.*

Among the applications of CA modeling to biology there are models on plant growth [16], on the propagation of infectious diseases [17], tumor growth [18], and on the dynamics of HIV infection [19]. A famous CA model from Pytte et al. [20] was used to model a region of the hippocampus with up to 10,000 neurons. This model gives qualitative and quantitative data that otherwise would require the numerical solution of about 250,000 coupled differential equations.

1.4.2 Agent-based models

Agent-based models (ABMs) are an extension of CA. The main characteristics that differentiate ABMs from CA are the potentially asynchronous interactions triggered by discrete events. Agents do not necessarily perform actions simultaneously as happens in CA. In addition in agent-based models the agents do not tile the simulated environment and are not necessarily grid-based.

Agent-based simulations aim at studying both quantitatively and qualitatively complex systems for which the microscopic, simpler rules of the interacting constituents are well known. ABMs are appropriate to investigate hypotheses on the sources of emergent behaviors in the macroscopic system and are an ideal framework for interdisciplinary research, allowing, for example, the integration of physical properties on the diffusion of cells in a media with biological knowledge on the intracellular biochemical processes. The possibility to take into account the detailed knowledge on specific cellular processes makes this methodology appropriate for the simulation of biological systems, where the heterogeneity of the interacting components are not safely reducible to a simplified categorization. For example modeling unique receptors on the surface of immune cells, a key element in the dynamics of the immune response, is relatively simple using agent-based modeling, whereas to simulate thousand of sub-species of the same cellular type would be impractical if not unattainable using differential equations.

In the past 20 years several agent-based models were developed to simulate the immune system as a complex adaptive system whose global behavior emerged from the collective microscopic interactions of its constituent agents. The first attempt at modeling the immune response using ABM is the famous IMMSIM by Franco Celada and Philip Seiden published in 1992 [19]. This initial attempt paved the way to the development of several other models simulating the complex interplay between pathogens and specific mechanisms of our immune system. Some focused on simulating the immune dynamics of

specific infectious diseases [21] or tumor growth [22]. Other agent-based models simulated specific processes involved in viral infections, like the dynamics of the diffusion in tissues [23] or intracellular chemical reactions [24]. Agent-based models have also been extensively used to simulate the immune dynamics of other infectious diseases [25,26]. One of the most complex immune system simulators is C-ImmSim, an evolution of the original Celada-Seiden model. C-ImmSim and its extension to cancer simulation SimTriplex have been used to run *in silico* experiments in virtual patients.

1.4.3 C-ImmSim

Thanks to the scientific research in Immunology and Biology there is enough knowledge on the elementary components and their interactions to develop *in silico* frameworks and models of the immune response.

C-ImmSim is one of the most complex simulators of the immune response. This model has been developed over more than 10 years and its latest version is able to simulate in detail both the cellular and humoral immune response in a three-dimensional volume of a simplified lymph node [27]. It belongs to the class of bit-string models [28]. Most of the entities in the system are characterized by specific strings of bits that represent the binding sites of their receptors. Interactions among the cells and pathogens occur as a function of the matching between such binary strings. C-ImmSim is able to simulate several microliters of lymphatic tissue and all the main cellular entities normally present in it. The cellular entities modeled are macrophages and dendritic cells, responsible of presenting the antigens to activate the adaptive immune response, T helper ($CD4^+$) lymphocytes, the immune cell playing a central role in both humoral and cellular immune response, cytotoxic lymphocytes able to kill infected cells, B cells responsible for the humoral response by the production of pathogen-specific antibodies and epithelial cells. Each cell in the system can be considered as a stochastic finite state machine that changes its state according to the stochastic interactions with other

cells and molecules in the system.

Another key characteristic of C-ImmSim is the possibility to run experiments on different virtual patients. The use of bit-strings to simulate the major histocompatibility complexes (MHC) class I and class II expressed on the cells surface allows the simulation of the variability of immune systems present in different individuals. Thus the parameter that describes the MHC defines the specific virtual patient whose immune response will be simulated in the model.

C-ImmSim has several free parameters that were tuned over the years to replicate the biological functions of the immune system. The basic mechanisms of the immune response, such as the timing of primary and secondary immune response to bacteria and inert antigens, were used to tune the parameters that regulate the formation of immunological memory and the binding affinity thresholds for the triggering of the immune response. Disease specific parameters, for example mutation rates of HIV epitopes and HIV slow damage to the hematopoiesis mechanism, and the variability of immune response to HIV in different individual were tuned using clinical data on the time-to-AIDS of untreated patients and on the proportion of rapid progressors, long-term non progressors and normal progressors.
The latest C-ImmSim version also models the effect of antiretroviral treatment on the HIV infection. The parameters that control the effect of combined antiretroviral treatment have been tuned using clinical data from twenty-two patients from the Spallanzani hospital in Rome [29].

In terms of spatio-temporal scales C-ImmSim simulates the immune response with a time step of 8 hours, although in one experiment the model parameters were tuned to work with a time step of 10 minutes [30]. Spatially it models individual cells in a volume of several microliters of lymphatic tissues, whereas temporally it can simulate the whole HIV infection, which lasts several years.

1.4.4 Modeling the different spatio-temporal scales

The complex interplay between the immune system, pathogens, and drug therapy is a hierarchical complex system and as such it happens at several spatio-temporal scales: the intracellular processes that lead to the infection of host cells and the release of new viral particles; the diffusion of viral particles and cells in tissues and the cellular interactions that constitute the immune response; the transmission of infectious viral particles within cities and countries. From the modeling point-of-view the challenge lies in the complexity of the model, which might require extensive programming work, the computational resources necessary to simulate the system in analysis and more often the lack of biological knowledge or experimental data for modeling key processes and validating the results. Agent-based models rely on experimental data to validate their results. Since ABM simulates the microscopic dynamics and its spatial properties, it often requires additional experimental data compared to traditional ODE models. Such additional data usually refers to specific biological processes that are only indirectly related to the problem in analysis, for example the diffusion coefficients of all the entities involved.

1.4.4.1 Microscale, intracellular processes

Although HIV life cycle might appear relatively simple, the complexity of the HIV infection at this spatio-temporal scale lies in its interaction with the host cell machinery and its ability to evade intracellular immune response. These complex phenomena happen at a temporal scale of a few hours within the host cell, about 8 μm in diameter, and the basic constituents are the different molecules of the virus and of the host cell. HIV infects the cells of the immune system, mostly T lymphocytes and macrophages, by binding its viral surface protein gp-120 to the CD4 membrane protein present on both cell types. After the fusion of the viral membrane with the cellular one, the HIV capsid is released in the cell cytoplasm. HIV relies on the reverse transcriptase, a viral enzyme that reverse transcribes the RNA genome into a double-stranded DNA. During this stage, viral protein reverse

transcriptase is likely to make mistakes in the genome transcription and is responsible for the high mutation rate of the HIV and its ability to develop resistance to antiretroviral drugs. After the viral DNA is transported into the host cell nucleus, another viral enzyme, the integrase, integrates the viral genome in the DNA of the infected cell. The integrated DNA may either lie dormant in the host cell genome or actively produce new copies of the virus using the host cell machinery. The intracellular processes that lead to the viral infection of a host cell have been studied using different modeling techniques from deterministic to stochastic [31] and from mathematical to cellular automata and agent-based models [14, 32-37]. In a mathematical model based on a system of coupled ordinary differential equations (ODEs) [14] each sub-process of the viral replication cycle is represented by one or more equations. Although mathematical models represent the overall behavior of the system, they tend to neglect the role of spatial and topological dependencies. The advantage of ABM at this scale is that proteins and molecules are modeled as individual agents and the complex behavior of the system emerges from their interactions. ABM has been widely used in modeling phenomena at cellular and molecular levels [32-37]. Some models include cellular details such as membrane with lipid bilayers, substrate molecules and enzymes with reaction rules and metabolic pathways.

1.4.4.2 Mesoscale, intercellular dynamics

At this spatio-temporal scale we can observe the dynamics of the diffusion of viral particles from infected cells and their interactions with neighboring cells. The temporal scale of the events observed ranges from minutes to days. The spatial scale is in the order of a few hundreds micrometers or a few hundreds of cells. The microscopic interactions are those of the infectious viral particles with the cells, including diffusion of the virus, receptor binding ad viral entry in the cells, while the macroscopic effect is the outcome of the infection, usually the time to the death of the whole population of cells. At this scale the physical processes like the virus diffusion and cellular growth play a major role and can be studied in detail.

The first models studying the infection dynamics at intercellular level were ODE mathematical models that aimed at extracting key parameters of the infections dynamics [12,38,39]. As the importance of the spatial component on the infection dynamics became evident the modelling approaches shifted to cellular automata (CA) models [26,40-42]. A good example of the advantages of CA modeling is the comparison between Bocharov's [39] and Beauchemin's models [26] for the influenza infection. Bocharov uses an ODE model with 13 variables and 60 parameters to fit .the experimental data, while Beauchemin's model manages to reproduce the infection dynamics using only 7 variables and 12 parameters.

1.4.4.3 Macroscale, individual patients

The dynamics of the complex interaction between HIV and the immune adaptive response can be studied at larger temporal (weeks, months and years) and spatial (lymph nodes of a few microliter, several thousands of cells) scales. At this spatio-temporal scales the basic constituents are the immune cells and viral particles that diffuse within lymph nodes of several microliters. Scientific research aims to understand how the complex dynamics of the adaptive immune response emerges from the simple interactions among the different cells of the immune system and their interactions with an antigen. At this spatial scale there are several phenomena that take place at different temporal scales: the primary immune response, which usually lasts between a few days to a few weeks, following the first exposure to a foreign antigen, the secondary immune response, lasting just a few days, due to the presence of memory cells [43] and slower phenomena like the emergence of drug resistant strains of HIV over period of years due to the selective pressure of antiretroviral treatment [44]. Its ability to mutate, that takes place at intracellular level, leads to macroscopic effects such as evading the immune response and developing resistance to antiretroviral drugs. The effect of HIV reservoirs on host survivability and the effect of combined

antiretroviral treatment (cART) on the infection dynamics also occur at this scale.

On first approximation the minimum subset of cells involved in the adaptive immune response are T helper lymphocytes, macrophages, dendritic cells, cytotoxic lymphocytes, and B cells. For each cell type there are sub-populations characterized by their specificity toward given antigens. Each specific sub-population of cells in an immune system co-evolves with sub-populations of a different cell type according to a fitness function that is, in part, itself a function of the emerging immune system. For example a B cell with the right specificity toward a given antigen is activated by a T helper cells able to recognize that antigen. The activation of an immune cell is often followed by its clonal expansion and the formation of memory cells that permanently change the immune system.

The specific sub-populations of cells define the co-evolving immune system, which, in turn, determines the fitness function according to which its sub-populations of immune cells, the constituent parts, evolve. The uniqueness of the entire immune system is the result of its temporal evolution. Such evolution is determined by the nonlinear feedback between the sub-population of immune cells (the microscopic level) and the immune system defined collectively by those sub-populations (the macroscopic level).

A number of mathematical models describe the HIV infection dynamics and the related immune response [refs]. The most common mathematical models represent only the dynamics of the HIV virus and T cells, making only a distinction between T cells in different states (Healthy, Infected, Dead). A few mathematical models considered the role of additional immune cell types in the infection dynamics adding either macrophages [45] or cytotoxic T lymphocytes [46]. Some ODE based models take into account the use of cART [47,48]. Still results of mathematical models tend to be generic, giving high-level assessments of the infection dynamics. The study of HIV through cellular automata and agent-based models is also common

[49-51] due to the discrete and stochastic nature of the biological entities involved in the phenomena.

1.4.4.4 Population scale

Although in this thesis we have not studied this spatio-temporal scale it is worth mentioning that this is the largest scale at which viral infections can be studied, neglecting the even larger effects of evolution across centuries. The infectiousness of the virus, the types of transmission (droplet contact, direct contact, vector borne, sexual transmission etc.), and the fatality rate represent and approximate the complexity of the lower scales. At this scale the challenge is hidden in both the complexity of the social network that links all the individuals in a community and in the dynamics of the spread of infections between individuals. Epidemics have been studied extensively using mathematical models, the first model dating back to since 189x [52]. After the discovery of the properties of complex networks, epidemiological studies have been performed using this new modeling approach [53-55]. The most evolved models of epidemics are the complex agent networks, which combine characteristics from both agent-based and complex network modeling.

1.5 Thesis

The challenges our society will face in the coming century have surpassed the need of descriptive models or simple predictive models. The next scientific challenge is not just to predict the behavior of a single virus under specific conditions; it is to anticipate their reaction to our strategies and actions against them. A multi scale system able to simulate *in silico* the evolution of the complex interplay between viruses, the immune system and the eventual antiviral drugs is necessary in order to develop such a dynamically predictive model.

Our thesis is that agent-based modeling offers an appropriate tool for a real multi-scale *in silico* experimental environment, a "virtual patient", and that critical holes in our understanding of both the

modeling paradigm and the biology at each spatio-temporal scale need to be addressed. We argue that this challenging goal requires a collectivist and interdisciplinary approach and that agent-based modeling (ABM), due to its characteristics, provides one of the most appropriate tools for this new scientific approach.

This thesis consists of two parts:

- The models developed in this thesis provide information about the missing biological knowledge needed to reproduce *in silico* the natural interaction between viruses and the immune system at each spatio-temporal scale;
- The new insights on the dynamics of HIV infection prove that ABM is more suitable to predict the dynamics of virus-immune system-drugs interactions than other traditional descriptive models.

1.6 Outline of the Dissertation

In Chapter 2 we investigate the dynamics of the first cycle of HIV intracellular replication at the level of a single cell. A cell is modeled as an individual entity with certain states and properties. The model is stochastic and keeps track of the main viral proteins and genetic materials inside the cell. We evaluate and compare two different simulation approaches: the first is a rate-based approach in which, like in an ordinary differential equations model, discrete viral proteins are treated as continuous entities and rates dictate the physical transformations of the viral proteins during HIV life cycle; the second is a classic agent-based modeling approach in which discrete viral proteins and genome are represented as entities moving in a two-dimensional space representing the cell cytoplasm and nucleus. In this agent-based modeling approach events are driven by the interactions between the agents when they collide in the simulated space. The results of the simulations are compared based on the number of integrated viral cDNA and the number of viral mRNA transcribed after a single round of replication. Both approaches are validated with

available experimental data. Simulation results and their validation give insights about the details of HIV replication dynamics inside the cell at the protein level. In addition the model highlights the strengths and weaknesses of the two different approaches.

In Chapter 3 we focus on the complexity of an in vitro infection of SARS-CoV in a population of cells, increasing the spatial scale by two orders of magnitude. We present a cellular automata model describing critical aspects of in vitro viral infections taking into account spatial characteristics of virus spreading within a culture well. The aim of the model is to understand the key mechanisms of SARS-CoV infection dynamics during the first 24 h post infection. We use data from the infection of human lung epithelial cells to tune the free parameters in the model. We also perform a sensitivity analysis of key parameters using a Latin Hypercube sampling to identify which mechanisms are critical to the observed infection of host cells and the release of measured virus particles.

In Chapter 4 we increase both spatial and temporal scales of several orders of magnitude to simulate HIV infection in a cohort of 250 virtual patients. At this level we model the complex interplay between the human immune system, antiretroviral drugs and HIV using C-ImmSim, an agent-based model that tracks individual immune cells and viral particles inside a single lymph node of 4 microliters. We use this complex model to find the main cause of the failure of structured treatment interruptions observed in several clinical trials. Although continuous antiretroviral therapy is currently the most effective way to treat HIV infection, unstructured interruptions are quite common and cannot be prevented. Several attempts to structure these interruptions failed due to an increased morbidity compared to continuous treatment but its causes are poorly understood and often attributed to the emergence of drug resistance. Here we show that structured treatment interruptions would fail regardless of the emergence of drug resistance. Our computational model of the HIV infection dynamics in lymphoid tissue inside lymph nodes, demonstrates that HIV reservoirs and evasion from immune surveillance themselves are

sufficient to cause the failure of structured interruptions. We validate our model with data from a clinical trial and show that it is possible to optimize the schedule of interruptions to perform as well as the continuous treatment in the absence of drug resistance. Our methodology enables studying the problem of treatment optimization without having impact on human beings.

In Chapter 5 we investigate the effect that temporary antiretroviral therapy has during primary HIV infection on the disease progression in individual patients. The current evidence is insufficient and the understanding of the dynamics of temporary combination antiretroviral treatment (cART) during primary HIV infection is still controversial. Here we quantitatively predict the sustained response of the immune system to a temporary cART based on the temporal correlations between $CD4^+$ cells and HIV viral particles during the natural course of the HIV infection. We calculate the temporal correlations as a function of time post-infection from *in silico* generated patient data. The *in silico* experiments are performed with the same clinically validated agent-based model used in the previous chapter. Our analysis predicts that after the acute phase, during which cART is known to have a long-term beneficial effect on the immune system, there is a secondary phase of about ten months during which temporary cART still has a relatively long-lasting effect on a patient's immune system compared to the asymptomatic phase. We validate our results using recent clinical trial data from Primo-SHM and provide an explanation to the unexpected beneficial effect of cART observed in the Primo-SHM study. The discovery of a secondary phase of sustained response may lead to significant changes to the HIV treatment guidelines.

References:

1. Malthus T (1978) Population: the first essay. Ann Arbor, MI: University of Michigan Press.
2. Kremer M (1993) Population Growth and Technological Change: One Million B.C. to 1990. *The Quarterly Journal of Economics* 108 (3): 681-716.
3. Zhang J, Zhang J, Lee R (2001) Mortality decline and long-run economic growth. *Journal of Public Economics* 80 (3): 485–507
4. Havick J (2000) The impact of the Internet on a television-based society. *Technology in Society*. 22 (2): 273–287
5. Hannan K (2009) We the Peoples - the United Nations in the 21st Century. *United Nations Department of Public Information*
6. Global report: UNAIDS report on the global AIDS epidemic 2012. (2012) WHO Library Cataloguing in Publication Data, Joint United Nations Programme on HIV / AIDS (UNAIDS).
7. Baccam P, et al. (2006) Kinetics of Influenza A Virus Infection in Humans. *Journal of Virology* Vol. 80, No. 15, 7590-7599.
8. Weaver W (1948) Science and complexity. *American scientist* 36 (4): 536-44.
9. Ilachinski A (2001) Cellular automata: a discrete universe. *World scientific.*
10. Holland JH (2006) Studying Complex Adaptive Systems. *Journal of Systems Science and Complexity* 19 (1): 1-8.
11. Nielsen M et al (2003) Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci.* 1(5): 1007-17.
12. Hethcote HW (2000) The mathematics of infectious diseases. *SIAM Rev.* 42(4), 599–653.
13. Perelson AS (2002) Modelling viral and immune system dynamics. *Nature Rev. Immunol*. 2 (1), 28–36.
14. another mathematical model of viral infection
15. J. Yin B. Reddy (1999) Quantitative Intracellular Kinetics of HIV Type 1. *AIDS Research And Human Retroviruses* vol. 15, pp. 273-283.

16. Lindemayer A, Prusinkiewicz P (1987) Developmental models of multicellular organisms: a computer graphics perspective. *ALIFE, volume 6 of Santa Fe Institute Studies in the Sciences of Complexity* 221-250.
17. Segel LA, and Perelson AS (1989) Shape space: an approach to the evaluation of cross-reactivity effects, stability and controllability in the immune system. *Immunology letters* 22(2), 91-99.
18. Moreira J and Deutsch A (2002). Cellular automaton models of tumor development: a critical review. *Advances in Complex Systems* 5(02n03), 247-267.
19. Seiden P, Celada F (1992) A model for simulating cognate recognition and response in the immune system. *Journal of Theoretical Biology* 158 235–270.
20. Pytte E, Grinstein G and Traub RD (1991) Cellular automaton models of the CA3 region of the hippocampus. *Network: Computation in Neural Systems* Vol. 2, No. 2 : 149-167.
21. Beauchemin C, Forrest S, Koster FT (2006) Modeling influenza viral dynamics in tissue. *Artificial Immune Systems* Springer Berlin Heidelberg. 23-36.
22. Mansury Y, Kimura M, Lobo J, Deisboeck TS (2002) Emerging patterns in tumor systems: simulating the dynamics of multicellular clusters with an agent-based spatial agglomeration model. *Journal of Theoretical Biology* 219: 343–370
23. Azimi M, Jamali Y, Mofrad MRK (2011) Accounting for Diffusion in Agent Based Models of Reaction-Diffusion Systems with Application to Cytoskeletal Diffusion. *PLoS ONE* 6(9): e25306.
24. Pogsona M, Smallwooda R, Qwarnstromb E, Holcombe M (2006) Formal agent-based modelling of intracellular chemical interactions. *Biosystems* 85: 37–45.
25.Zorzenon RMdS, Coutinho S (2001) Dynamics of HIV Infection: A Cellular Automata Approach. *Phys. Rev. Lett.* 87(16): 168102.
26. Beauchemin C, Samuel J, Tuszynski J (2005) A simple cellular automaton model for influenza A viral infections. *Journal of Theoretical Biology* 232 (2): 223-234.
27. M. Bernaschi, F. Castiglione (2001) Design and implementation of an immune system simulator. *Comput. Biol. Med.* 31: 303-331.
28. Farmer JD, Packard NH, Perelson AS (1986) The immune system, adaptation and machine learning. *Physica D* 22: 187

29. Paci P et al. (2011) Earlier is better: a timely HAART initiation may pave the way for best controllers. *BMC Infectious Diseases* 11:56.
30. Baldazzi V, Castiglione F and Bernaschi M (2006) An Enhanced Agent Based Model of the Immune System Response. *Cellular Immunol* 244: 77-79.
31. Summersy J, Srivastavawz R, Youw L and Yin J (2002) Stochastic vs. deterministic modeling of intracellular viral kinetics. *Journal of Theoretical Biology* vol. 218, p. 309321.
32. White T, Webb, K (2004) Cell modeling using agent-based formalisms, *AAMAS'04*, New York, USA.
33. Qwarnstrom E et al. (2006) Formal agent-based modelling of intracellular chemical interactions, BioSystems, vol. 85, p. 3745.
34. Makkena R et al. (2003) Multi-agent system for the quantitative simulation of biological networks. *AAMAS 03*, Melbourne, Australia 385-392.
35. Litorco J., Lee L, Jacob C (2004) Immunity through swarms: Agent-based simulations of the human immune system. *In Proceedings of 3rd International Conference on Artificial Immune Systems*, Vol. 3239. Springer-Verlag GmbH, Catania, Sicily, Italy (2004) 400-412, Committee Draft
36. Han HK, Tay JC, Guo Z (2005) Sufficiency verification of hiv-1 pathogenesis based on multi-agent simulation. *In Proceedings of the Genetic and Evolutionary Conference*, Vol. I. Washington D.C. 305-312
37. Sumen C et al. (2005) Agent-based modeling of the context dependency in t cell recognition, Journal of Theoretical Biology, vol. 236, pp. 376-391.
38. Perelson AS et al. (1996) HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* 271, 1582–1586.
39. Bocharov GA, Romanyukha AA (1994) Mathematical model of antiviral immune response III. Influenza A virus infection. *J. Theor. Biol.* 167 (4), 323–360.
40. Beauchemin C (2006) Probing the effects of the well-mixed assumption on viral infection dynamics. *Journal of Theoretical Biology* 242 (2): 464-477.
41. Sidorenko R (2004) Structured model of influenza virus replication in MDCK cells. *Biotechnology and Bioengineering* 88 (1): 1-14.
42 Beauchemin C, Forrest S, Koster FT (2006) Modeling influenza viral dynamics in tissue. *Artificial Immune Systems, 5th International Conference,* ICARIS 2006, Oeiras, Portugal, Proceedings, Lecture Notes in Computer Science 4163 Springer, 23-36.
43 Abbas AK , Lichmann AH (2004) Basic immunology. Philadelphia, Saunders/Elsevier.

44 Coffin, J. M. (1995) HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science* 267:483-489.
45 Hernandez-Vargas, EA and Middleton RH (2013) Modeling the three stages in HIV infection. *J Theor Biol.* 320:33–40
46 . Ganusov VV, Neher RA and Perelson AS (2013) Mathematical modeling of escape of HIV from cytotoxic T lymphocyte responses. J. *Stat. Mech.* P01010.
47. Perelson AS, P. W. Nelson (1999) Mathematical Analysis of HIV-I: Dynamics in Vivo. *SIAM Rev.* 41: 3-44.
48. Simon V, Ho DD (2003) HIV-1 dynamics in vivo: implications for therapy. *Nat Rev Micro.* 1: 181-190.
49. Pandey RB, Stauffer D (1990) Metastability with probabilistic cellular automata in an HIV infection. *J. Stat. Phys*, 61: 235-240
50. Burkead EG, Hawkins JM, Molinek DK (2009) A dynamical study of a cellular automata model of the spread of HIV in a lymph node. *Bull Math Biol.* 71(1): 25-74.
51. Sloot PMA, Chen F and Boucher CAB (2002) Cellular Automata Model of Drug Therapy for HIV Infection. in S. Bandini; B. Chopard and M. Tomassini, editors, *5th International Conference on Cellular Automata for Research and Industry*, ACRI 2002, Geneva, Switzerland, October 9-11, 2002. Proceedings, in series Lecture Notes in Computer Science. 2493: 282-293.
52. Kermack WO and McKendrick AG (1927) *Proc. R. Soc. Lond. A* 115 (772): 700-721
53. Sloot PMA et al. (2008) Stochastic simulation of HIV population dynamics through complex network modeling. *International Journal of Computer Mathematics* Vol: 85(8): 1175-1187.
54. Quax R, Bader DA, Sloot PMA (2009) Simulating Individual-Based Models of Epidemics in Hierarchical Networks. Computational Science – ICCS 2009, *Lecture Notes in Computer Science* 5544: 725-734.
55. Shan Mei A et al. (2011) Increasing risk behaviour can outweigh the benefits of antiretroviral drug treatment on the HIV incidence among men-having-sex-with-men in Amsterdam. *BMC Infectious Diseases* 11:118.