A stochastic spatial model of HIV dynamics with an asymmetric battle between the virus and the immune system

A stochastic spatial model of HIV dynamics with an asymmetric battle between the virus and the immune system

Hai Lin[1] and J W Shuai[2,3]

[1] Department of Chemical Biology, Xiamen University, Xiamen 361005, People's Republic of China
[2] Department of Physics and Institute of Theoretical Physics and Astrophysics, Xiamen University, Xiamen 361005, People's Republic of China
E-mail: jianweishuai@xmu.edu.cn

Abstract. A stochastic spatial model based on the Monte Carlo approach is developed to study the dynamics of human immunodeficiency virus (HIV) infection. We aim to propose a more detailed and realistic simulation frame by incorporating many important features of HIV dynamics, which include infections, replications and mutations of viruses, antigen recognitions, activations and proliferations of lymphocytes, and diffusions, encounters and interactions of virions and lymphocytes. Our model successfully reproduces the three-phase pattern observed in HIV infection, and the simulation results for the time distribution from infection to AIDS onset are also in good agreement with the clinical data. The interactions of viruses and the immune system in all the three phases are investigated. We assess the relative importance of various immune system components in the acute phase. The dynamics of how the two important factors, namely the viral diversity and the asymmetric battle between HIV and the immune system, result in AIDS are investigated in detail with the model.

[3] Author to whom any correspondence should be addressed.

Contents

1. Introduction

Acquired immunodeficiency syndrome (AIDS) has rapidly emerged as one of the main global health problems. Since the time when the human immunodeficiency virus (HIV) was first isolated in 1983 [1] and was confirmed to be the cause of AIDS in 1984 [2, 3], intensive studies have been carried out to understand the complex interactions between viruses and the immune system and to unravel the dynamics of the ultimate failure of the immune system in fighting against HIV. There are two types of HIV: HIV-1 and HIV-2. In the present paper, we will only consider HIV-1 that is responsible for the great majority of infections worldwide.

The immune system can protect the host from viral infections. Most viruses, after infecting humans, are rapidly cleared during the initial acute stage of infection or establish a long-term asymptomatic chronic infection due to a mutual compromise of the virus and the immune system. But the pattern of disease progress in HIV infection is different. A typical HIV infection can be subdivided into three phases [4]. During the first 2–6 weeks after HIV entering the host, the acute phase occurs. Viral load (the number of virions in the peripheral blood) increases dramatically followed by a fall of the helper T lymphocytes (CD4⁺T cells). Patients usually show symptoms similar to flu in this phase. Then during the 9th to the 12th week, the viral load experiences a sharp decline and the number of CD4⁺T cells returns almost to their normal level. With the viral load falling to a very low level, the clinical symptoms disappear. However, the viruses cannot be completely eliminated after primary infection unlike most other viral infections, and patients enter the second phase, which is a long period of the asymptomatic stage (clinical latency) lasting for 1–10 years or more. During this phase, viral load remains

very low, while the CD4$^+$T cell population continues to decline slowly until it becomes lower than a critical value. Finally, the viral load climbs up again, resulting in an onset of AIDS. In such a final phase, the impaired immune system can no longer fight off infections, and patients usually die from a variety of opportunistic infections [5] that would normally be cleared.

What made the HIV different from the other pathogens? From more than 20 years of intensive studies, people have obtained an in-depth understanding of many aspects of HIV dynamics. Nowadays, it is well known that there are two prominent mechanisms of the HIV infection. Firstly, HIV preferentially infects and kills the CD4$^+$T cells [6, 7], which are one of the master regulators of the immune system required for the CD8$^+$ cytotoxic T cells and B cells to respond to viruses. Therefore, the progressive depletion of CD4$^+$T cells in the HIV infection eventually results in the loss of many cell-mediated immunological functions that are essential for effective defenses, and causes the patients to have a high susceptibility to opportunistic infections and malignancies. Secondly, HIV has a high mutation rate [8]–[11] because its reverse transcriptase lacks the proofreading mechanism. So the viruses can create highly diverse quasi-species [12] and escape recognition by the host's immune effector cells [10]. Although there are different perspectives on why the HIV infection displays a three-phase pattern [13]–[15], it is generally believed that the two distinguishing features mentioned above are the most likely causes of HIV's avoiding clearance and ultimate failure of the immune system. However, how these two mechanisms promote long-term co-evolution of viruses and the host immune system and induce the transition from the asymptomatic phase to development of AIDS still remains obscure and requires further investigation.

Nowadays, mathematical modeling has become an integral part of the research on the HIV infection. Elaborate mathematical models cannot only help further understanding of HIV dynamics [16]–[18], but also provide useful predictions that can be used as the basis for assessing the effects of chemotherapy or immunotherapy and for the design of optimum therapies and effective vaccines [16, 18].

Several mathematical models using ordinary differential equations (ODE) have been developed to investigate the dynamics of HIV infection [19]–[21]. Nowak *et al* [19, 20] proposed 'antigenic diversity threshold theory', suggesting that beyond a threshold of antigenic diversity, the immune system loses control over the viral population, triggering a high rebound in viral load and onset of AIDS. The Wang–Deem model [21] postulated that the competition among the different strains of T cells degrades immune control of viral infection and, finally, allows HIV to escape from the immune system. These models have successfully reproduced the three-state dynamics and contributed significantly to the understanding of the mechanisms of HIV infection. However, these mathematical models are limited to systems of ODE describing spatio-temporal averaged behavior and cannot capture the stochastic properties of HIV dynamics.

Recently, cellular automata (CA) models have been used in HIV modeling [22]–[24]. The interactions of HIV and the immune system are described by a set of simple local transition rules in CA models, and these models have shown great potential for simulating the temporal and spatial dynamics of HIV infection. But this kind of approach has its own limitations. Firstly, the inter-cellular interaction rules used in CA models are usually very simple, and it is hard to demonstrate the plentiful realistic virus–host interactions in HIV infection. Secondly, mobility is an important feature of virions and immune cells in the real immune system, whereas in most of the CA models [22, 24] the cells are unmovable.

In this paper, we proposed a stochastic spatial model with moderate complexity to model HIV dynamics. As opposed to the well-mixed population model, a spatially explicit model can take into account the local interactions and the spatial inhomogeneities. These features are of central importance since viral propagation and the immune response are fundamentally local processes [22, 23]. We propose a more detailed and realistic simulation frame by incorporating many important features of HIV dynamics, including infections, replications and mutations of viruses, antigen recognitions, activations and proliferations of lymphocytes, and diffusions, encounters and interactions of virions and lymphocytes. Compared to the ODE approaches, our model explicitly considers the stochastic characters of HIV infection. We also integrate more plentiful interactions between viruses and the immune system in our model than in previous CA models. Our simulation results are in good agreement with the clinical observations both at the individual level of three-phase HIV development and at the population level of AIDS onset time distribution.

2. Model

Simulations are carried out on a two-dimensional $L \times L$ square lattice. Three types of mobile individuals, CD4$^+$T cells, CD8$^+$T cells and HIV virions, are distributed on the lattice sites. Each lattice site can contain at most one individual of a certain type, but it can be simultaneously occupied by two or more individuals of different types.

2.1. The random diffusion of cells and virions

At each simulation step, which is equivalent to one day in real life, each individual can move randomly to a nearby site in the migration neighborhood of size $(2r_m + 1) \times (2r_m + 1)$ (the Moore neighborhood of range r_m). Hence, r_m stands for the maximum distance an individual can travel in a day. Such a stochastic diffusion process can be simulated as follows [26]: a virion or a lymphocyte selects randomly a site in the migration neighborhood r_m. If the selected site has been occupied by the same type of individual, the two individuals will swap their position; otherwise, the individual will move to the selected site.

2.2. Binary string model of virus and T cell

Lymphocytes recognize antigens through the surface receptors, such as B-cell receptors (BCRs) and T-cell receptors (TCRs) [27]. The receptors bind specifically to molecules from the antigen and elicit the immune response. In our model, both antigen epitopes on viruses and antigen-specific receptors on T cells are represented by binary strings of length l [28]. Hence, the strength of binding interaction between an antigen i and a T cell j is dependent on their Hamming distance [21]:

$$h_{ij} = \sum_{n=1}^{l} (1 - \delta_{a_i^n, a_j^n}),$$ (1)

where a_i^n and a_j^n denote the bit values at the nth site of the viral string a_i and the T cell's string a_j, and δ_{lm} is the Kronecker delta.

2.3. The infection of CD4$^+$T cells by viruses

The primary cellular receptor for HIV entry into a target cell is CD4, a molecule expressed on the surface of helper T lymphocytes and other immune cells [29]. The binding of the HIV envelope glycoprotein to CD4 and coreceptors initiates a series of conformational changes leading to membrane fusion and viral entry [30, 31]. Upon infection, viral RNA can be reverse transcribed into double-stranded DNA. During this step the mutations may occur because the process of reverse transcription is extremely error-prone [8]. Next, the viral DNA is transported to the nucleus and integrated into the host chromosomes. Hence, the virus can reproduce itself using the machinery of the host cell, and finally the cell dies and releases the viruses that can further infect other cells [32].

Accordingly, we model the infection process as follows: a virion will invade a healthy CD4$^+$T cell with a probability p_v if they meet at the same site. Following the assumption made in previous models [19, 21, 22], here we do not consider the multiple infections [33]; thus a virion cannot invade an infected cell. We also exclude the HIV recombination [34] and ignore the fitness cost [35] of the viral mutant in the model. All of these restrictions can be relaxed in principle, but doing so will lead to considerable complexity. Then during the process of reverse transcription the viral binary string will be mutated with a probability m_v. For simplicity, we assume that in each transcription step at most one bit site is to be mutated. Therefore, when mutation occurs, a random site of the binary string is selected to flip the bit value from '1' to '0' or from '0' to '1'.

An infected CD4$^+$T cell will produce a new virus with a probability r_v at each simulation step. The binary string of the progeny virus is the same as the viral DNA (either mutated or unmutated) carried by the infected cell. Once the number of viruses in a given cell exceeds a lysis threshold n_{lys}, the cell dies and the progeny virions are released and randomly dispersed on the lattice sites in the Moore neighborhood of range r_v. To mimic a finite carrying capacity of the system, only one virion selected at random can survive if two or more progeny virions occupy the same lattice site and the others are removed. This approach is also applied to model the finite carrying capacity of the system for the CD4$^+$T cells and CD8$^+$T cells.

2.4. The proliferation of CD4$^+$T cells

Clonal expansion of CD4$^+$T cells [36]–[38] is a necessary step in most adaptive immune responses. CD4$^+$T cells recognize antigen-derived peptides bound to major histocompatibility complex (MHC) class II molecules that are exposed on the surface of antigen presenting cells (APC). The CD4$^+$T cells will then become activated and begin to proliferate, secrete cytokines and differentiate into memory T cells.

For simplicity, in the model we do not distinguish different states of the CD4$^+$T cells, i.e. the naïve T cells, the activated T cells and the memory T cells. We do not take into account the APCs either. At each simulation step, the coincidence of a virion and a healthy CD4$^+$T cell leads to both infection (with probability p_v) and cell proliferation with a probability p_{ij}, where the subscripts denote the viral strain i and the CD4$^+$T cell clone j. Thus, the effects of APCs are represented implicitly via the parameter p_{ij}. Since the matching of an antigen with a TCR may not be exact but approximate [39]–[41], we assume that the activation probability depends exponentially on the Hamming distance between the virus and the cell [21]

$$p_{ij} = p_0 \exp(-\alpha h_{ij}), \tag{2}$$

where p_0 is the activation probability when the binary strings perfectly match between the HIV and the cell. The parameter α regulates to what degree p_{ij} depends on the Hamming distance. Upon clone extension, the activated CD4$^+$T cell will produce n_x offspring that are dispersed randomly in the Moore neighborhood of range r_x. Here we also consider the finite carrying capacity of the system described above.

2.5. The proliferation of CD8$^+$T cells and the killing of infected CD4$^+$T cells

Once inside the CD4$^+$T cell, viruses become hijackers, using the cell's machinery to produce more viruses. During this process some viral proteins synthesized by the cell are degraded into oligopeptides, a fraction of which are transported into the endoplasmic reticulum (ER). In the ER, these peptides bind to MHC class I molecules and the resulting complexes are transported to the cell surface for antigen presentation [42]–[44]. Thus CD8$^+$T cells can detect viral infections by monitoring the cell surface. On recognition of viral antigens, CD8$^+$T cells become activated and begin to proliferate, and most of its progenies become effector cells that can recognize and kill the infected cells [45, 46].

According to the way of interaction between CD8$^+$T cells and infected CD4$^+$T cells, we assume that when a CD8$^+$T cell meets an infected CD4$^+$T cell in which the virus has already begun to be reproduced, the CD4$^+$T cell will be killed with a probability k_{ij} and all of the viruses inside the cell are removed. Meanwhile, the CD8$^+$T cell will be triggered to proliferate with a probability p'_{ij} and produces n_y offspring that are dispersed randomly in the Moore neighborhood of range r_y. The subscripts of the p'_{ij} denote the viral strains i inside the infected cell and the CD8$^+$T cell clone j. In order to consider the cross-reactivities, we assume that the activation probability and the killing probability also depend on the Hamming distance

$$p'_{ij} = p'_0 \exp(-\alpha h_{ij}), \tag{3}$$

$$k_{ij} = k_0 \exp(-\alpha h_{ij}), \tag{4}$$

where p'_0 and k_0 are the corresponding probabilities when $h_{ij} = 0$, and α is a regulation parameter as defined in equation (2).

2.6. The help of CD4$^+$T cells to B cells

CD4$^+$T cells do not kill viruses directly. Their main function is to help the B cells and CD8$^+$T cells to develop fully to fight the infection. For a resting B cell to become activated and produce antibodies, signals provided by the CD4$^+$T cells are required. These signals enable the B cells to proceed to proliferate and differentiate into the antibody-secreting cells [47]. Therefore, with the help of CD4$^+$T cells, B cells can kill the virions more effectively, which means that the viral mortality rate depends indirectly on the number of virus-specific CD4$^+$T cells.

Based on the above considerations, in order to model the help of CD4$^+$T cells to B cells, we simply assume that the death rate of free virions is a function of the concentration of the virus-specific CD4$^+$T cells, and the B cells do not appear explicitly in our model for the sake of simplicity. Thus, the death probability of a free virion (strain i) in one day is given as

$$d_{v_i} = d_{v0} + (1 - d_{v0}) \times \frac{[CD4_i]}{\theta + [CD4_i]}. \tag{5}$$

The first term d_{v0} is the helper-independent natural death probability of a virion in one day, and the second term stands for the regulation of viral mortality probability by $[CD4_i]$, the density

of virus-specific CD4$^+$T cells. Here, we assume that the dependence has the form of the Hill function $f(x) = x^n/(\theta^n + x^n)$ with order $n = 1$, which is often used to represent an unknown regulation function in biology systems [50]. Parameter θ in the Hill function is the density or the concentration yielding half-maximal expression.

2.7. The help of CD4$^+$T cells to CD8$^+$T cells

The exact mechanisms of how the CD4$^+$T cells deliver help to the CD8$^+$T cells are still unclear. Many experiments [25, 51] suggest that the primary response of CD8$^+$T cells develops normally in the absence of help, while the efficient restimulation of memory CD8$^+$T cells appears to require some help. Therefore, with the help of CD4$^+$T cells, the CD8$^+$T cells can persist at elevated levels in the long term.

Accordingly, in our model we simply assume that the CD8$^+$T cells have a relatively high mortality rate without help, but have a long lifespan in an environment of plentiful virus-specific CD4$^+$T cells. Hence, the death probability of a CD8$^+$T cell (strain i) is given by

$$d_{y_i} = d_{y0} \left(1 - \frac{[\text{CD4}_i]}{\theta + [\text{CD4}_i]} \right). \tag{6}$$

Here, the regulation function of CD4$^+$T cells is also approximated by a Hill function, and parameter d_{y0} denotes the helper-independent natural death probability of a CD8$^+$T cell in one day.

2.8. Replenishment of T cells

In a healthy human body, the T cell loss can be fully compensated by replenishment of new naïve cells from the thymus [52]. Therefore, in our model we assume that the new T cells are generated at a constant rate λ. Since the concentration of the T cells must be kept at a steady state without viral infection, the replenishment rates of the CD4$^+$T cells and CD8$^+$T cells are given by $\lambda_x = x_0 d_{x0}$ and $\lambda_y = y_0 d_{y0}$, respectively, where x_0 and y_0 denote the initial concentration of CD4$^+$T cells and CD8$^+$T cells before viral infection. The newborn cells, whose binary strings are drawn randomly from a uniform distribution, are dispersed to the randomly chosen empty sites at each simulation step.

3. Results

3.1. Parameter settings

Simulations have been carried out on a lattice of size $N = 100 \times 100$ with periodic boundary conditions. At the beginning, we disperse randomly over the whole lattice CD4$^+$T cells and CD8$^+$T cells with random binary strings, and the initial virions with only one sequence of genes. Since the epitope of the antigen is 8–11 amino acids long [21, 53], we set the length of binary string to $l = 10$.

In our model, the death probability of a healthy CD4$^+$T cell in one day is $d_{x0} = 0.01$ [54]. Considering that the death rate of infected CD4$^+$T cells is $0.1 \, \text{day}^{-1}$ [10], which means that the average lifespan of an infected cell is 10 days, we set the lysis threshold $n_{\text{lys}} = 5$ and the probability that an infected cell generates a new virion in one day $b_v = 0.5$. Thus, the average time from infection to lysis of an infected CD4$^+$T cell is 10 days.

Table 1. The parameters used for simulations.

| Parameter	Value	Explanation
l	10	Length of binary string
v_0	0.03	Initial viral density
x_0	0.8	Initial density of CD4$^+$T cells
y_0	0.1	Initial density of CD8$^+$T cells
r_m, r_v, r_x, r_y	3	Migration range or dispersion range
p_v	0.8	Probability that a virion infects a healthy CD4$^+$T cell
b_v	0.5	Probability that an infected cell generates a new virion in one day
m_v	5.5×10^{-5}	Probability that viral mutation occurs during reverse transcription
n_{lys}	5	Lysis threshold of infected CD4$^+$T cells
p_0, p_0'	0.8	Baseline probability that a T cell is activated by a virion
n_x, n_y	5	Number of offspring produced during proliferation of T cells
k_0	0.8	Probability that a virion is killed by a CD8$^+$T cell
d_{v0}	0.01	Helper-independent natural death probability of a virion in one day
d_{x0}	0.01	Death probability of a healthy CD4$^+$T cell in one day
d_{y0}	0.1	Helper-independent death probability of a CD8$^+$T cell in one day
α	1.0	Regulation parameter in equations (2)–(4)
θ	0.5	Regulation parameter in equations (5) and (6)

We are only concerned with the dynamics of activated CD8$^+$T cells, which have a relatively high death rate ($0.1 \, \mathrm{day}^{-1}$) compared with the naïve cells. However, for simplicity, we do not distinguish between activated T cells and naïve T cells; therefore in our model, the density of the initial CD8$^+$T cells, which is used as a background to elicit primary cellular immune responses, is set to a very low level at $y_0 = 0.1$.

The other parameters used in the model are as given in table 1 unless specified otherwise.

3.2. Three-phase dynamics

The time evolution of HIV, CD4$^+$T cells and CD8$^+$T cells at the individual level for three samples and at the population level are given in figures 1(a)–(d), respectively. Figures 1(a)–(c) show three simulation samples corresponding to three individuals with AIDS onset occurring about 10, 2 and 15 years after infection. All the samples reproduce the three-phase pattern observed in the dynamics of HIV infection [4]. The results show a high level of viral load and a sudden drop of CD4$^+$T cells after infection, followed by a rapid increase of CD8$^+$T cells. With the viral number declining to a very low level in a few weeks, the CD8$^+$T cells go through a contraction process and then return to a relatively high level in asymptomatic phase. These results are consistent with clinical observation [59]. Different from CD8$^+$T cells, CD4$^+$T cells are depleted gradually during the whole incubation period. Finally, the exhaustion of the pool of CD4$^+$T cells results in two effects, i.e. the lack of stimulation by the infected CD4$^+$T cells and the lack of help provided by the healthy CD4$^+$T cells, which lead to a sudden drop in CD8$^+$T cells and a rebound of viruses.

Figure 1(d) shows the average evolution of HIV infection at the population level. The results are averaged over 10 000 simulation samples (corresponding to 10 000 HIV-infected individuals) with different initial configurations. The large fluctuation, i.e. the large error bars,
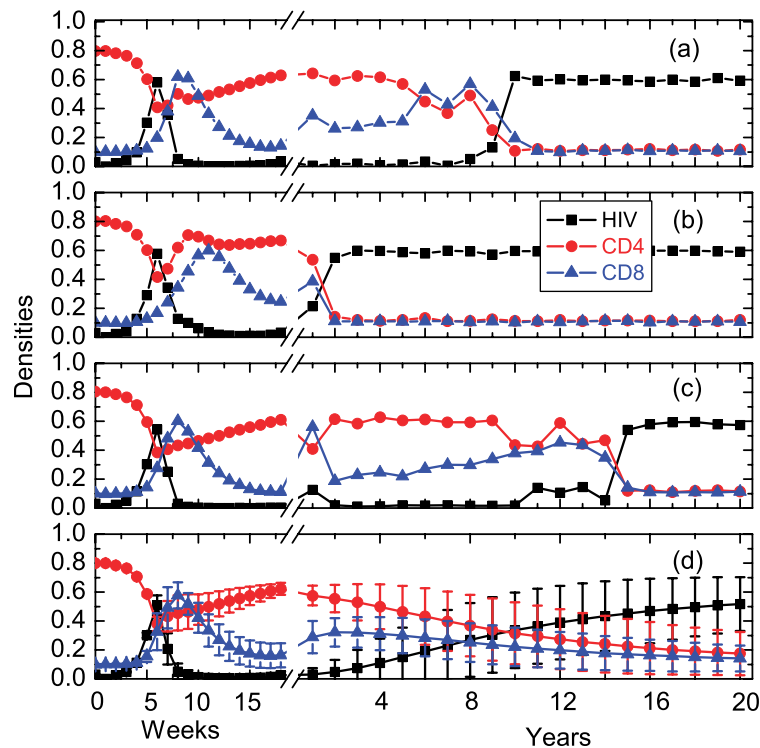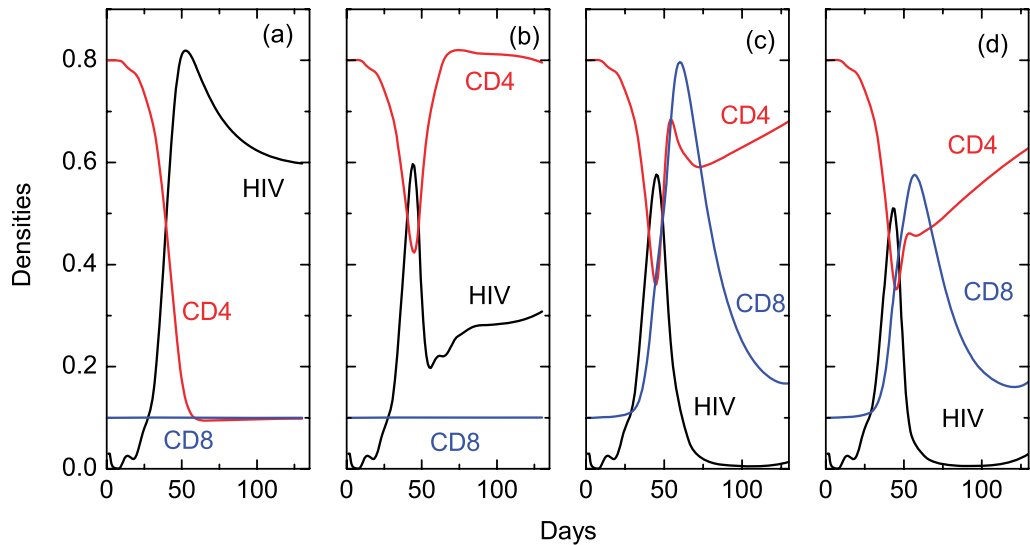
Figure 1. (a–c) Time evolution of HIV, CD4$^+$T cells and CD8$^+$T cells in different individuals with AIDS onset occurring about 10, 2 and 15 years after infection. (d) Average dynamics of HIV infection. The results are averaged over 10 000 samples. The error bars are for standard deviations.

of the results, especially shown in the incubation period, indicates the great diversity in disease development among different individuals. It was found that, on average, the growth of viruses and the decline of CD4$^+$T cells are approximately linear during the asymptomatic phase, while at the individual level the transition is abrupt, as shown in figures 1(a)–(c). Deterministic models based on ODE or well-mixed population models based on mean-field approximation can usually capture only the average behaviors, while spatial models, of which the localizations, the inhomogeneities and the stochasticities are natural ingredients, can yield qualitatively different behaviors from the averaged results.

3.3. The acute phase

A newly HIV-infected individual has plenty of uninfected cells that are suitable for the viruses to infect and reproduce, which induces a transient increase of the viral load in the acute phase. However, viral number will decline rapidly to a very low level in a few weeks. Most studies postulated that the fall in viral load is due to the host's immune response, including HIV-specific humoral (B cell) response [56] and the cellular (CD8$^+$T cell) immune response [57]–[59]. Some investigations have suggested that the decrease of viral number is simply a result of population dynamics [17], i.e. as more and more healthy CD4$^+$T cells become infected and die, the infection of the virions becomes more difficult, resulting in the decline of viruses. Therefore, it is still unclear to what extent the immune responses are involved in the reduction of viral load during primary HIV infection.

Figure 2. Time evolution of HIV, CD4$^+$T cells and CD8$^+$T cells in the acute phase (a) without immune response, (b) with only humoral response, (c) with only cellular response and (d) with fully immune response. The results are averaged over 10 000 simulations. The black lines are for HIV, red lines for CD4$^+$T cells and blue lines for CD8$^+$T cells.

In order to discuss this problem, we show in figure 2 the time evolution of our model in the acute phase in four different conditions, i.e. without immune response ($p_0 = 0$ and $p'_0 = 0$), with only humoral response ($p_0 = 0.8$ and $p'_0 = 0$), with only cellular response ($\theta \to \infty$ in equation (5)) and with fully immune response.

Our simulation results shown in figure 2 indicate that the three mechanisms, i.e. lack of target cells, and the humoral and cellular immune responses, contribute to the reduction of viral load in acute infection, but the cellular response plays a decisive role in suppressing the viral load. If the host's immune system does not respond to the infection ($p_0 = 0$ and $p'_0 = 0$), there is only a small decline in viral load, as shown in figure 2(a). Hence, the lack of target cells has no significant effects on HIV dynamics in the acute phase. When the absence of CD8$^+$T cells is simulated with $p_0 = 0.8$ and $p'_0 = 0$, we see a considerable degree of virus decline, but there still are relatively high levels of viruses in the host (figure 2(b)). However, it was found that the absence of neutralizing antibody ($\theta \to \infty$ in equations (5)) does not influence the viral dynamics in the acute phase, as shown in figure 2(c). Our results on the decisive role of the cellular response are consistent with the observation in experimental studies that the relatively poor clearance of virus occurred when the CD8$^+$T cell response was delayed [57].

It has been reported that the CA model proposed by dos Santos and Coutinho [22] is sensitive to the initial density of the infected cell. If the value of this parameter is lower than 0.05, then the acute phase disappears [60]. We also investigate the sensitivity of our presented model to the initial viral density v_0. The simulation results indicate that even if the initial viral density v_0 is set to 0.002, our model can reproduce the three-phase pattern successfully. However, the time needed for the viral load to reach the peak increases with decreasing v_0, while the peak value is insensitive to v_0 (figure 3). We also find that the global properties of the HIV dynamics are not influenced by the initial viral load (data not shown).
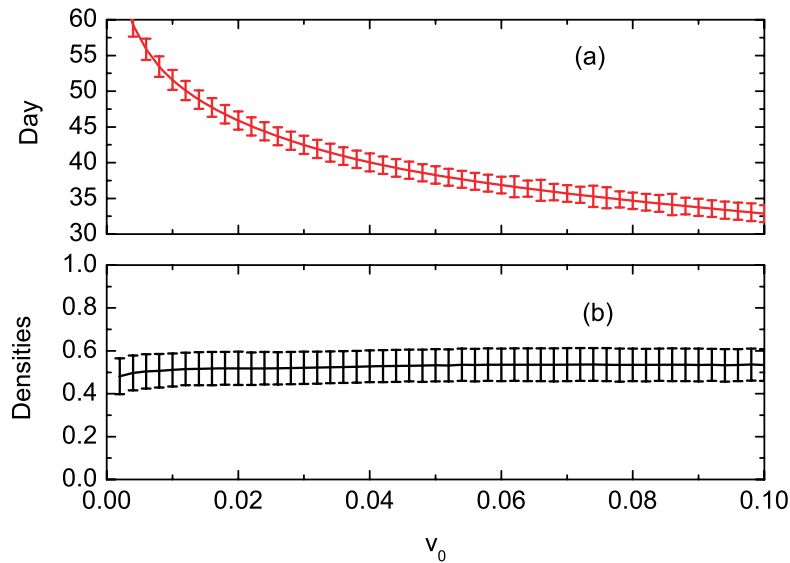
Figure 3. Sensitivity of the HIV dynamics to the initial viral density v_0 in the acute phase. (a) The day that the viral density reach the peak. (b) The peak value of viral density. The results are averages over 10 000 simulations. The error bars are for the standard deviations.

3.4. The asymptomatic phase

As shown in figures 1 and 2, HIV elicits a strong immune response, which effectively suppresses the viral load to a very low level. But it raises the question of why this response fails ultimately to eliminate the virus. In the following, we show that the answer may lie in the high mutation rate of HIV.

The emergence of viral mutants plays a major role in HIV infection for virus to escape immune responses [9]. To assess the effect of viral mutation on HIV disease progression, we study the viral and cellular dynamics in the absence of viral mutation ($m_v = 0$). The result shown in figure 4(a) indicates that the system converges to an equilibrium state that the viruses persist at a very low level and cannot be totally eliminated by the immune system. Then we increase the viral genetic diversity by setting the virus binary string length to $l = 1$, 2 and 3, which means that the total number of different strains in a viral quasi-species is $M = 2$, 4 and 8, respectively (figures 4(b)–(d)). It can be seen that by increasing the viral diversity, the system becomes more and more unstable. Each emergence of a new viral mutant induces an outbreak of viruses and an abrupt drop in CD4$^+$T cells, followed by a strong response of CD8$^+$T cells to suppress viral load to a low level again, and then the CD4$^+$T cells return to a normal level. Therefore, the system can tolerate some degree of viral diversity. However, once the viral diversity exceeds a threshold ($l = 4$, corresponding to $M = 16$ in our model) the immune system is no longer able to control HIV, resulting in the onset of AIDS, as shown in figure 4(e).

In order to further investigate the evolution of viral quasi-species and virus-specific immune response, we track the birth–death processes of viral mutant strains and virus-specific T-cell clones. Figure 5 demonstrates that the viral mutant appears in the early stage of the asymptomatic phase. Then, over a long period of time, only a small number of new strains emerge and they are then suppressed to a very low level by the specific immune response.
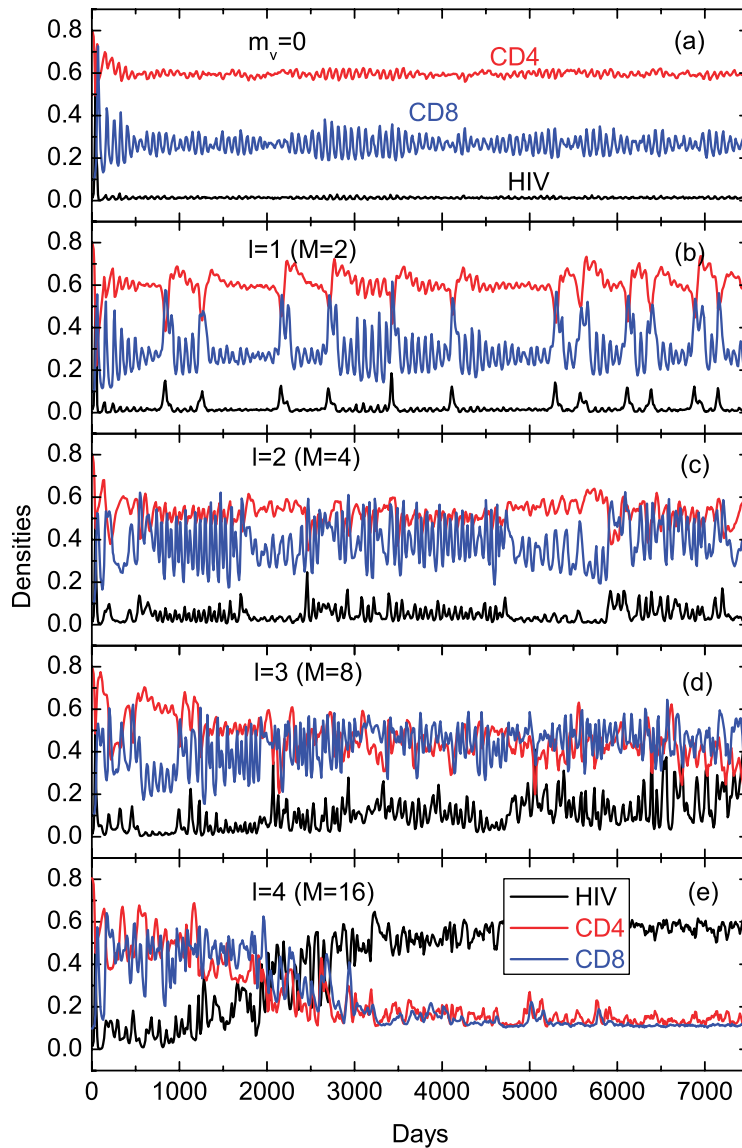
Figure 4. The time evolution of the HIV system at different viral mutation diversities. (a) The system converges to an equilibrium state if the viral mutation rate $m_v = 0$. (b)–(d) As the viral diversity increases, i.e. $M = 2$ for (b), $M = 4$ for (c) and $M = 8$ for (d), the system becomes more and more unstable. (e) Once the viral diversity exceeds a threshold ($l \geqslant 4$ with $M \geqslant 16$), the immune system is no longer able to control the HIV. The black lines are for HIV, red lines for CD4$^+$T cells and blue lines for CD8$^+$T cells.

Therefore, the immune system is totally dominant in this stage. However, in the late stage of the asymptomatic phase (from about the 2000th day to the 4000th day in figure 5), emergence of new viral strains become much more frequent. Although new types of specific T-cell clones will be generated to fight these mutants, the mutant viruses will have already produced new types of viral mutations before they are cleared. Consequently, with the accumulation of viral mutations, the immune system cannot function any more to respond to so many viral strains (after about the 4000th day in figure 5).
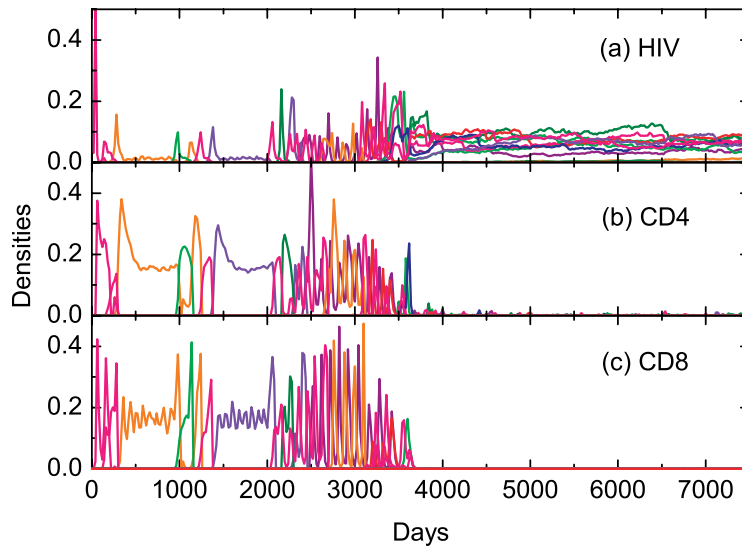
Figure 5. Evolution of (a) HIV, (b) CD4+T cells and (c) CD8+T cells. The lines with different colors are for different viral strains and corresponding virus-specific T-cell clones.

Why can the immune system not control HIV if there are too many viral mutants? The answer is that the fight between HIV and the immune system is asymmetry, i.e. a given CD4+T cell clone can only recognize a corresponding specific viral strain, while any viral strain can infect all susceptible CD4+T cells regardless of their specificity [25]. Therefore, if there are too many viral strains in the host, each T-cell clone can fight only a small fraction of the viruses, but has a lot of chances to be infected and killed by the whole viral population. Such an asymmetry battle leads to the continual depletion of CD4+T cells, which finally results in the destruction of the whole immune system. The simulation results shown in figure 6 demonstrate this process. It can be seen from an individual sample given in figures 6(a) and (b) that in the late stage of the asymptomatic phase the number of viral strains begin to overtake that of the virus-specific T-cell clones, which starts the collapse of the immune system. The results shown in figures 6(c) and (d) are averaged over 10 000 simulation samples. It was found that for each type of virus, its growth is followed by the expansion of the corresponding CD4+T and CD8+T cells, which in turn reduces virus growth, leading the system toward equilibrium. The repeat of such a process gives oscillations for virus and immune cells. This phenomenon can also be seen in other models [12, 25]. Note that in figure 6, we plot each day's data, whereas in figure 1, in order to facilitate comparison between the simulation results and the clinical data [4], we record only each year's data after the first 130 days; therefore the oscillations cannot be shown. However, as the number of viral strains increases more rapidly than the number of T cells, the equilibrium can no longer be maintained, and the T-cell population gradually loses the ability to keep up with the evolution of viral quasi-species. Finally, the immune system loses the fight against HIV, resulting in the onset of AIDS.

Besides the oscillation of virus and immune cells, the results shown in figure 6 also provide a prediction about how the viral diversities and divergences evolved in the HIV infection. We found that the virus strain number increases most of the time, followed by a growing increase of strain number before the onset of AIDS, which may, in principle, be tested by advanced sequencing technology.
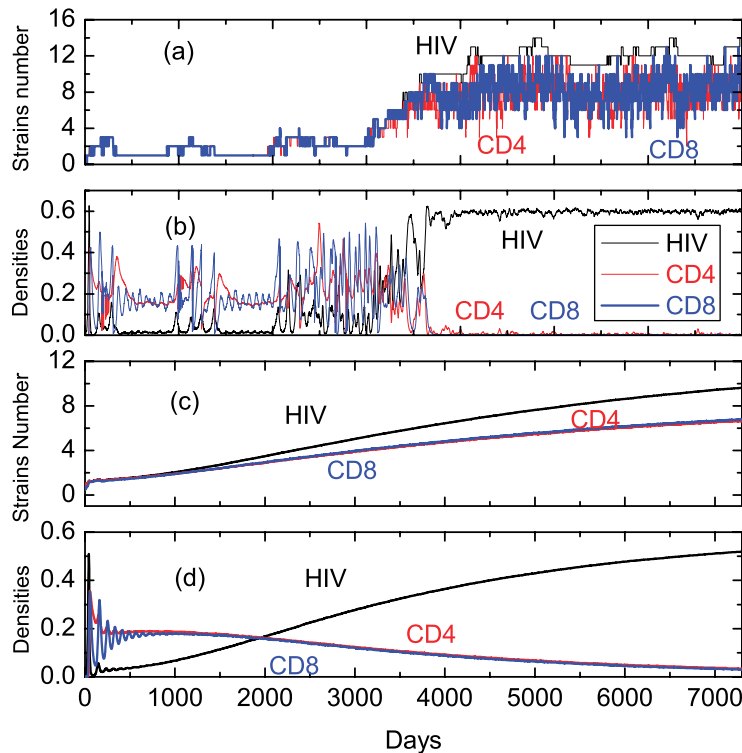
Figure 6. Time evolution of (a) the number of HIV strains and virus-specific T-cell clones and (b) their corresponding densities for an individual sample. (c, d) The corresponding results averaged over 10 000 samples. The black lines are for HIV, red lines for CD4$^+$T cells and blue lines for CD8$^+$T cells.

3.5. The onset of AIDS

There are a lot of statistical data from various studies about the incubation period distribution in the development of AIDS [61]. Much effort has been spent on parametric estimates of the waiting time from HIV infection (seroconversion) to the onset of AIDS [62], while some research works applied an alternative approach such as percolation theory to study the characteristics of the time course of HIV infection [63]. Here, we used our model to investigate the time distribution from HIV seroconversion to AIDS onset.

The CD4$^+$T cell count is usually adopted as a predictor of progression of HIV infection. It has been reported that the CD4$^+$T cell count, immediately after seroconversion, is about 800 cells μl^{-1} [64, 65]. When the CD4$^+$T cell count falls to 350 cells μl^{-1}, HIV-related opportunistic infections begin to appear, and at 200 cells μl^{-1} the person is classified as having AIDS regardless of the appearance of other opportunistic infections [65]. Consequently, in our model, we set the initial density of the CD4$^+$T cells to 0.8, and define the onset of AIDS as the CD4$^+$T cells density falls to 0.2. The simulation results shown in figure 7 describe the proportion developing AIDS with time from seroconversion. It was found that the time distribution from HIV infection to AIDS onset is greatly influenced by the viral mutation rate m_v. By setting $m_v = 5.5 \times 10^{-5}$, our simulation result is in good agreement with the clinical data from [61] on the AIDS development at population level.
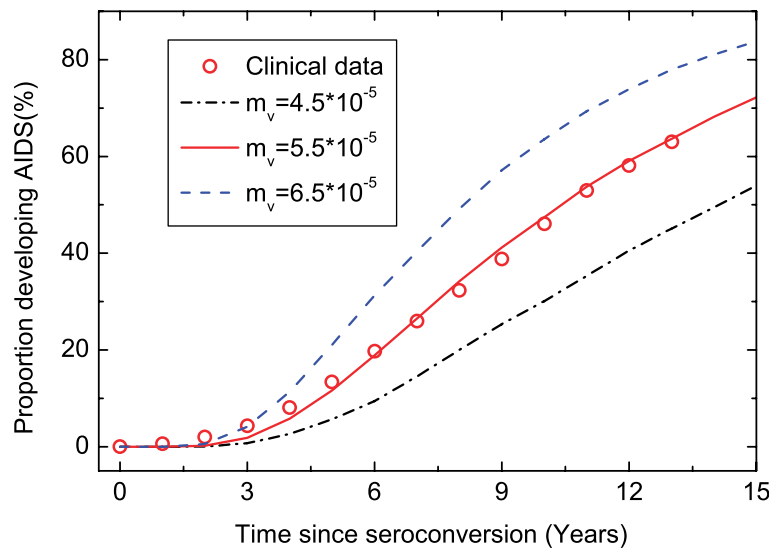
Figure 7. Development of AIDS from seroconversion. The results were obtained from 10 000 simulation samples. The lines represent the simulation results with the black dash-dot line for $m_v = 4.5 \times 10^{-5}$, the solid red line for $m_v = 5.5 \times 10^{-5}$ and the blue dash line for $m_v = 6.5 \times 10^{-5}$. The red circles are the clinical data from [61].

4. Discussion

We investigate the dynamics of HIV infection based on a stochastic spatial model, which incorporates the interactions between viruses and the immune system in a more detailed and realistic manner than previous models. Firstly, we consider the random mobilities of virions and cells, and define the individual interactions of HIV and the immune system components in terms of probabilities and rates. Therefore, our model explicitly incorporates the stochastic characters of the HIV dynamics, which enables us to study the great diversity of disease development among different individuals, and to investigate the time distribution from infection to AIDS onset. Secondly, we distinguish two different T-cells, namely, CD4$^+$T cells and CD8$^+$T cells, which are collectively treated as a single entity in previous models [19]–[22], [24]. We also consider the help of CD4$^+$T cells to B cells and CD8$^+$T cells. Hence, we can assess the relative importance of various immune system components in fighting against the viruses. Thirdly, the introduction of a binary string model facilitates our study of the evolution of viral quasi-species and virus-specific T-cell population.

The simulation results are in good agreement with the clinical observations, both at the individual level and at the population level. At the individual level, the three-phase pattern of HIV infection is reproduced successfully. The interactions of viruses and the immune system in all the three phases were investigated in detail. We assessed the relative importance of various immune system components in the acute phase and found that the CD8$^+$T cells play a decisive role in suppressing the viral load. This observation implies that stimulation of a CD8$^+$T cell response might be an important goal in the development of an effective vaccine against AIDS [57]. We have observed the viral diversity threshold phenomena postulated by Nowak *et al* [19]. However, in the Nowak model [19], the viruses will be cleared if

IOP Institute of Physics ⏀ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

the production rate of new viral mutants is too low, whereas in our model the immune response cannot completely eliminate the viruses even if in the absence of viral mutation. By distinguishing the different entries of lymphocytes, we can use a biologically more realistic model to simulate the asymmetric battle, a concept proposed in previous studies [19, 20, 25], between HIV mutant strains and immune cell clones. Furthermore, we are able to show how the viral mutation rate is largely responsible for the HIV dynamics, and we obtain, at the population level, incubation period distributions that are consistent with the clinical data [61].

Although our model has taken into account many of the important features of HIV infection, it is still simple compared with the real interactions between HIV and the immune system. For example, the B cell and the APC did not appear explicitly in our model, and we did not take into account the fact that the viral size is much smaller than the lymphocyte size. Thus, a lattice site can accommodate only one virion, as for the lymphocyte. For simplicity, we also assume that the migration and dispersion range of virions in one day is the same as that of cells. However, this restriction is not necessary, for the simulation results do not show any differences even for a diffusion velocity of virions 10 times faster than that of cells (data not shown). Many of the parameters of our relatively simple model are set quite arbitrarily; for example, the distances that T cells and virions move, the number of progeny virions released from the infected cell, and the mutation rate of virus do not strictly correspond to those *in vivo* [66, 67]. Therefore, our model needs to be further developed. More types of lymphocytes can be added to the system, and we can allow more than one virion to occupy the same lattice site so that multiple infection is permitted, and the sensitivity of the model to the choice of parameters should be tested, which will be the subject of our further research. However, in spite of all these simplifications in our present model, the simulation results capture many observed features of HIV infection quite well and may contribute to a better understanding of AIDS disease.

Acknowledgment

References

[1] Barré-Sinoussi F *et al* 1983 *Science* **220** 868
[2] Gallo R C *et al* 1984 *Science* **224** 500
[3] Levy J A *et al* 1984 *Science* **225** 840
[4] Pantaleo G, Graziosi C and Fauci A S 1993 *N. Engl. J. Med.* **328** 327
[5] Grant A D, Djomand G and De Cock K M 1997 *AIDS* **11** (Suppl B) S43
[6] Douek D C *et al* 2002 *Nature* **417** 95
[7] Brenchley J M *et al* 2006 *J. Virol.* **80** 6801
[8] Nowak M A 1990 *Nature* **347** 522
[9] Goulder P J and Watkins D I 2004 *Nat. Rev. Immunol.* **4** 630
[10] Althaus C L and De Boer R J 2008 *PLoS Comput. Biol.* **4** e1000103
[11] Carlson J M *et al* 2008 *PLoS. Comput. Biol.* **4** e1000225
[12] Nowak M A and May R M 2000 *Virus Dynamics: Mathematical Principles of Immnology and Virology* (Oxford: Oxford University Press)

[13] Brenchley J M, Price D A and Douek D C 2006 *Nat. Immunol.* **7** 235

[14] Grossman Z *et al* 2006 *Nat. Med.* **12** 289

[15] Hel Z, McGhee J R and Mestecky J 2006 *Trends Immunol.* **27** 274

[16] Coffin J M 1995 *Science* **267** 483

[17] Philips A N 1996 *Science* **471** 497

[18] Perelson A S *et al* 1997 *Nature* **387** 188

[19] Nowak M A, May R M and Anderson R M 1990 *AIDS* **4** 1095

[20] Nowak M A *et al* 1991 *Science* **254** 963

[21] Wang G and Deem M W 2006 *Phys. Rev. Lett.* **97** 188106

[22] dos Santos R M Z and Coutinho S 2001 *Phys. Rev. Lett.* **87** 168102

[23] Strain M C *et al* 2002 *J. Theor. Biol.* **218** 85

[24] Shi V, Tridane A and Kuang Y 2008 *J. Theor. Biol.* **253** 24

[25] Wodarz D 2007 *Kill Cell Dynamics: Mathematical and Computational Approaches to Immunology* (Berlin: Springer)

[26] Tobias R, Mauro M and Erwin F 2007 *Nature* **448** 1046

[27] Goldstein B, Faeder J R and Hlavacek W S 2004 *Nat. Rev. Immunol.* **4** 445

[28] Farmer J D, Packard N H and Perelson A S 1986 *Physica* D **22** 187

[29] Dalgleish A G *et al* 1984 *Nature* **312** 763

[30] Doms R W 2000 *Virology* **276** 229

[31] Dimitrov D S 2000 *Cell* **101** 697

[32] Pomerantz R J and Horn D L 2003 *Nat. Med.* **9** 867

[33] Dixit N M and Perelson A S 2005 *Proc. Natl Acad. Sci. USA* **102** 8198

[34] Levy D N *et al* 2004 *Proc. Natl Acad. Sci. USA* **101** 4204

[35] Ganusov V V and De Boer R J 2006 *PLoS Comput. Biol.* **2** e24

[36] Watts T H and McConnell H M 1987 *Annu. Rev. Immunol.* **5** 461

[37] Liu Y and Janeway C A Jr 1992 *Proc. Natl Acad. Sci. USA* **89** 3845

[38] Garcia S, DiSanto J and Stockinger B 1999 *Immunity* **11** 163

[39] Davis M M and Bjorkman P J 1988 *Nature* **334** 395

[40] Kersh G J and Allen P M 1996 *J. Exp. Med.* **184** 1259

[41] Rudolph M G and Wilson I A 2002 *Curr. Opin. Immunol.* **14** 52

[42] Moore M W, Carbone F R and Bevan M J 1988 *Cell* **54** 777

[43] York I A and Rock K I 1996 *Annu. Rev. Immunol.* **14** 369

[44] Rock K L 1996 *Immunol. Today* **17** 131

[45] Oehen S and Brduscha-Riem K 1998 *J. Immunol.* **161** 5338

[46] Veiga-Fernandes H *et al* 2000 *Nat. Immunol.* **1** 47

[47] Garside P *et al* 1998 *Science* **281** 96

[48] Parker D C 1993 *Annu. Rev. Immunol.* **11** 331

[49] Smith K M *et al* 2000 *J. Immunol.* **165** 3136

[50] Savageau M A 1976 *Biochemical Systems Analysis: A Study of Function and Design in Molecular Biology* (Reading, MA: Addison-Wesley)

[51] Bevan M J 2004 *Nat. Rev. Immunol.* **4** 595

[52] Jameson S C 2002 *Nat. Rev. Immunol.* **2** 547

[53] Goulder P J and Walker B D 1999 *Nat. Med.* **5** 1233

[54] De Boer R J *et al* 2003 *J. Immunol.* **170** 2479

[55] Wilson J D *et al* 2000 *AIDS* **14** 225

[56] Richman D D *et al* 2003 *Proc. Natl Acad. Sci. USA* **100** 4144

[57] Koup R A *et al* 1994 *J. Virol.* **68** 4650

[58] Borrow P *et al* 1994 *J. Virol.* **68** 6103

[59] Wilson J D *et al* 2000 *AIDS* **14** 225

[60] Strain M C and Levine H 2002 *Phys. Rev. Lett.* **89** 2198501
[61] CASCADE Collaboration 2000 *Lancet* **355** 1131
[62] Dangerfield B and Roberts C 1999 *Syst. Dyn. Rev.* **15** 273
[63] Kamp C and Bornholdt S 2002 *Proc. Biol. Sci.* **269** 2035
[64] Williams B G and Dye C 2003 *Science* **301** 1535
[65] Williams B G *et al* 2006 *J. Infect. Dis.* **194** 1450
[66] Beauchemin C *et al* 2007 *J. Immunol.* **178** 5505
[67] Mansky L M and Temin H M 1995 *J. Virol.* **69** 5087