A multi-scale approach for deciphering HIV infection

Ertaylan, G.

Citation for published version (APA):
Chapter 5

HIV-1 CO-RECEPTOR TROPISM

G. Ertaylan¹, Peter MA Sloot¹

¹ Computational Science, University of Amsterdam, Amsterdam, The Netherlands

A complex automata model of HIV-1 co-receptor tropism: Understanding mutation rate pressure, in Reviews in Antiretroviral Therapy, Washington D.C., USA, December 2007
Background

For attaching to its target cells, Human immunodeficiency virus (HIV) primarily binds to a cell surface molecule called CD4. In addition to CD4, the virus requires a secondary receptor for successful target entry. The most important secondary receptors in vivo are the chemokine receptors CCR5 and CXCR4 [1]. The virus variants that are able to use the CCR5 co-receptor in entry are called CCR5-tropic or R5 tropic and variants that can use the CXCR4 co-receptor are called CXCR4-tropic or simply X4 tropic. (Figure 5.1)

Early infection with HIV is characterized by the predominance of CCR5-tropic (R5) virus. However, over the course of infection CXCR4-tropic (X4) virus appears in the later stage of the infection in approximately 50% of the infected individuals and usually precedes an accelerated CD4+ T cell depletion with rapid disease progression [2], [3]. Regardless of endless efforts, the reason for this phenotypic switch and its effect on the disease progression is still not clear.

Current advances in entry inhibitors and the recent availability of co-receptor antagonist maraviroc initiated new studies on questions like: which patient is suitable for this co-receptor antagonist; does the blockage of CCR5 receptor by an antagonist actually favor X4 variants; do our essays sensitive enough to rule out the possible early emergence of X4 variants accompanying maraviroc usage [4]; can maraviroc be used safely in combination with other drug classes; when is the best time on the disease course to include it in the treatment? All the answers to these questions rely on our understanding of the longitudinal dynamics of co-receptor tropism of HIV populations.

Figure 5.1 Schema showing the HIV co-receptor tropism. The CD4 receptor and co-receptors are indicated on the CD4+ T cell surface. For successful cell entry R5 and X4 tropic viruses have to bind to CD4 and their corresponding preferred receptor consecutively.
Determining the Tropism

The third hypervariable region (V3) of HIV-1 gp120 is known to be involved in cell tropism [5], [6] co-receptor tropism [7], [8], [9], in-vitro cytopathogenicity, and viral replication competence [10]. There are various techniques such as 11/25 rule and bioinformatics tools to determine the HIV tropism (geno2pheno [11], wetcat [12] and webPSSM [13]), all these techniques depend on the amino acid sequence of the V3 loop.

The V3 loop has also been estimated to have the highest evolutionary rate for HIV-1 Group O [14]. There is no other reason to think otherwise for other groups of HIV-1. This combined with the finding that as few as one or two amino acid changes in V3 loop was sufficient for co-receptor switching in various studies [15], [10], [16], [17], [18], nominated the V3 loop as the key player in HIV co-receptor tropism.

Recently it has been argued that slower progression to AIDS is strongly correlated with a slower rate of synonymous substitution, indicating slower replication rate and longer viral generation times. Therefore the synonymous/non-synonymous (ds/dns) substitution rate ratio has been proposed as a marker for evaluating in-vivo HIV evolution [19]. Related to that, an important observation made by Pastore and colleagues was the extraordinarily low ds/dns ratio of mutations on the V3 loop recently [20] from in-vitro studies where the immune system was not involved. Since then, we learned form clinical data that once neutralizing antibodies are involved, there will be positive selection for the non-synonymous variants and the balance on the ds/dns ratio shifts towards zero. Therefore we hypothesize that mutation rate for HIV-1 can be observed from the substitution rates in V3 loop.

It has been shown by independent studies that, substitution of the V3 loop of X4 tropic clone with R5 tropic counterpart alone was sufficient to alter its co-receptor preference in vitro and its replication characteristics in vivo [7] [21] [22]. Furthermore, HAART treatment argued not to influence viral evolution within HIV-1 V3 region [23]. These findings once more underline the key role of V3 in HIV tropism also in the clinical field.

Theories and Models on Tropism Dynamics

It is observed from clinical data that the R5 tropic virus dominates the early infection, and R5X4 or X4 tropic viruses emerges later in the disease. The dominance of R5 tropic viruses early in the infection can be explained by preferential selection in either donor or the recipient [24]. Although it has been shown that differential distribution of R5 and X4 viruses can be observed in the donor genital secretions of women [25] but not in men [26] no consensus was established whether this differential distribution can lead to preferred transmission of R5 viruses. On the other hand on the recipient side it has been demonstrated that the selection can be dependent on early events during transmission. CCR5 rich intestinal epithelial cells selectively transfer R5 viruses to their target cells in oral-genital and mother-child transmission [27]. Dendritic cells that transport virus particles have been found to bind preferentially to R5 viruses [28] [29]. Finally, high-levels of stromal cell derived factor 1 (SDF-1), – a CXCR4 blocking ligand, is shown to be expressed by mucosal tissues [30]. It seems that several mechanisms can partially account for the preference of R5 viruses early in the infection.

Another possibility could be the competition between R5 and X4 viruses during primary infection. There is increasing support that R5 viruses predominate the early infection while other tropic variants coexist as a minor subpopulation that remains undetected by currently available methods [31] [32] [33].

Also, it is common to assume that considering the rapid turnover of ~10^{10} -10^{22} virions per day [34] combined with the mutation rate of HIV [35], [36] should result in the emergence of X4
variants fairly early during the infection. In certain cases as few as 2-3 mutations are thought to be necessary for the full co-receptor switch [20], [37].

Several models proposed aimed to explain the tropism dynamics since the discovery of the co-receptors CCR5 and CXCR4 in HIV cell entry. A model proposed by Rodrigo [38] assumed that X4 variants were more cytopathic and therefore the cells infected with X4 variants had a shorter life span. As a result even for the initial inoculum assumed to contain a 1:1 ratio of R5 and X4 viruses, the R5 tropic variant has dominated over the X4 tropic variant over time. If we take a closer look at their assumption; indeed increased cytopathicity of HIV from late infection compared to isolates from early infection from the same individual has been reported [39]. However the assumption of X4 viruses being more cytopathic than R5 variants have not been confirmed yet. Arien et al. reported no difference in replicative capacity of R5 and X4 using clones [59]. Finally it is also not clear how the X4 virus can reappear over time.

Regoes and Bonhoeffer have constructed a simple HIV evolution model where four mutations are necessary to switch from initial R5 to final X4 with intermediate mutants having a fitness disadvantage compared to R5 and X4 variants and the final X4 virus is fitter than the initial R5 virus.

They have observed a realistic fraction of infected individuals (about 50%) that switch to X4 virus as well as the time it takes for the switch to occur was in the clinically observed range [40]. Nevertheless the fitness criteria assumed in their model is too generic to have conclusive remarks from this study.

Problems of Modeling Co-receptor Tropism
Although the models have shown to be useful for simulating the simple qualitative behavior, their applicability to experimental data remains limited due to the innate properties of the modeling scheme used as well as the undetermined dynamical parameters. The reliability of models based on partial differential equations (PDE) and ordinary differential equations (ODE) is hampered by their assumption of homogeneity. Immune system interactions are mainly local cell-cell and virion-cell interactions, where little differences define the dynamics. Therefore assumption of the immune system as a big bowl of soup might not be sufficient to explain certain characteristics of the dynamics. For instance a high fitness virus burst from an isolated infected cell may leave no progeny whereas a low fitness virus burst having many infected cells in proximity may dominate the infection. Furthermore, these events occurring at the right time and the right place can characterize the course of the disease.

In this chapter we introduce an individual based model of HIV target cell entry to examine the population dynamics between virus strains within a single infected host. We have designed the model with emphasis on spatial interactions between target cells and R5, R5X4 and X4 tropic HIV-1 strains.

We have imposed a co-receptor change scheme, based on stepwise multiple point mutations. We allow for biologically natural trade-off between the viral mutation rate and the overall reproduction capacity of the viral quasispecies. We show that there is an optimum range of mutation rate for HIV co-receptor tropism and this coincides with the experimental range. Further we will discuss its implication for medical biology.

Materials and Methods

Here we illustrate the basic concepts of Transmission & Mutation Model of HIV-1 co-receptor tropism. First we introduce the conceptual model, later we explain implementation of the
conceptual model as a computational model, its parameters and assumptions, in the third subsection we present the experiments conducted with a) larger and smaller grid size, b) different burst size, c) different lethal, neutral and forward mutation rates and d) the standard model parameters. Finally in the last subsection we give an overview about the applications used and the model performance.

**Conceptual Model**

For studying the Transmission & Mutation hypothesis for the co-receptor switch we have defined two types of entities: T cells and HIV, where HIV particles are able to infect (leading to infection) T cells with varying efficiency. The efficiency of entry and infection is based on the type of T cell the virion is targeting and the virion’s tropism for this specific T cell. The sub-populations of T cells and HIV defined in the model are shown in Table 5.1.

<table>
<thead>
<tr>
<th>HIV</th>
<th>Memory T cells (T_M)</th>
<th>Naïve T cells (T_N)</th>
<th>R5 tropic</th>
<th>R5x4 tropic</th>
<th>r5X4 tropic</th>
<th>X4 tropic</th>
</tr>
</thead>
<tbody>
<tr>
<td>T cells</td>
<td>Exclusive CCR5 co-receptor. Target for R5, R5x4 and r5X4 tropic HIV</td>
<td>Exclusive CXCR4 co-receptor. Target for X4, r5X4 and R5x4 tropic HIV</td>
<td>Can only infect T_M</td>
<td>Can infect both T_M and T_N with preference for T_M</td>
<td>Can infect both T_M and T_N with preference for T_N</td>
<td>Can only infect T_N</td>
</tr>
</tbody>
</table>

Table 5.1. Sub-populations of T cells and HIV defined in the model with their specifications.

It has been described in the literature that CCR5 and CXCR4 are preferentially expressed in memory and naïve T cells, in that order [60]. The tropism of R5 virus for memory cells and X4 virus for naïve cells, is reported independently [61]. Hence we have made the simplifying assumption that T cell population consists of Memory T cells (T_M) and Naïve T cells (T_N) that have exclusive CCR5 or CXCR4 co-receptors respectively. HIV population is made of four subpopulations: R5, R5x4, r5X4 and X4, with decreasing affinity for CCR5 and increasing affinity for CXCR4 from R5 through X4. Therefore T_M and T_N are targets for different sub-populations of HIV.

Different rates of successful infection for each of the four virus variants (R5, R5x4, r5X4 and X4) while infecting a T_M or a T_N are determined by the co-receptor utilization efficiency (ε) of each variant.

The dynamics of the model are depicted in figure 5.2. T_N and T_M reproduce with the reproduction rates of r_N and r_M from existing cells. There is a constant influx of new cells from the thymus and bone marrow represented by the born rates of λ_N and λ_M.

A healthy T cell has a lifespan of d_N and d_M, for naïve and memory T cells respectively. When infected, they have a shorter lifespan denoted as d_Ni and d_Mi. For simplicity, the above-mentioned parameters have been taken equal for both T cell subpopulations T_N and T_M (Hence, r_N = r_M, λ_N = λ_M, d_N = d_M and d_Ni = d_Mi). At first glance, this assumption is rather simplistic. Nevertheless, considering the T cell homeostasis, including naïve T cells being the precursors of memory T cells would skew the dynamics toward a definitive co-receptor switch. This is due to the early infection and killing of naïve cells by X4 tropic viruses preventing them from maturing to become memory T cells. Thereby leaving R5 tropic viruses with less target cells.

V_{X4}, V_{R5x4}, V_{r5X4} and V_{R5} represent the subpopulations of HIV corresponding to X4 tropic, R5x4 (dual) tropic for preference for CCR5, r5X4 (dual) tropic for preference for CXCR4 and R5 tropic in that order.
The number of successful infections determines the dynamics of the virus population. Therefore all virus variants (except for R5 tropic and X4 tropic due to their exclusive target cell range) are competing for target cells to infect. Once the infection takes place there is a probability of mutation with the mutation rate (µ). The change of tropism is only possible in the progeny, if there is a mutation and that is non-lethal and non-neutral during the infection of the target cell. If it is lethal no successful progeny can be produced and if it is neutral the progeny is produced as if no mutation has occurred. Hence after a successful infection the possible four outcomes are:

\[
P_{\text{No-mutation}} = (1-\mu) \quad \text{(Probability of a normal infection without mutation)}
\]

\[
P_{\text{Neutral mutation}} = (\mu_{\text{Neutral}})\mu \quad \text{(Probability of a neutral mutation)}
\]

\[
P_{\text{Lethal mutation}} = (\mu_{\text{Lethal}})\mu \quad \text{(Probability of a lethal mutation)}
\]

\[
P_{\text{Tropism change}} = (1-(\mu_{\text{Lethal}}+\mu_{\text{Neutral}}))\mu \quad \text{(Probability of a mutation with tropism change)}
\]

The virions are cleared at a constant rate, and have lifespan of c. The infection rate \( \beta_{R5, X4} \) is dependent on \( V_{R5, R5X4, X4} \) and the infectiousness of the virions.

![Diagram of Transmission & Mutation model schema.](image)

Figure 5.2. Transmission & Mutation model schema. N and M represents TN and TM, VX4, VR5X4 and VR5 represents X4, R5X4 and R5 tropic HIV respectively. On each side r is the reproduction rate, \( \lambda \) is the born rate, d is the lifespan, \( \beta \) is the infection rate and p is the burst size. Generic parameters are mutation rate \( \mu \) and HIV lifespan c.

The conceptual model is a simplified version of the in-vivo immune system dynamics where \( T_N \) and \( T_M \) are decoupled with same parameters with the exception of having different co-receptors, therefore infectibility. The population dynamics of \( T_N \) and \( T_M \) are highly dependent on virus variants present around them. In this conceptual model the effect of the T cells that are latently infected but not producing HIV particles (latent reservoir in acute HIV infection) and double infections are not taken into account because of their debatable bound to the viral tropism.
Computational Model and Implementation

The implementation of the conceptual model explained above, is realized by using NetLogo© (ver.4.0.3) programming language [63].

Space and time are represented as discrete entities, and simulations are implemented in a two dimensional Cellular Automata (CA) that is divided into grid of pixels in a Cartesian coordinate system with periodic boundary conditions. The size of this grid was chosen as a representative portion of a complete lymph node and if desired, can be scaled up to any size. The topography of the grid is a closed torus. One time step in the simulation corresponds to six hours in real time.

We use a “cells-as-agents” approach, wherein each individual agent represents one biological cell or a virus. Agents (cells or viruses) are either motile or static within this coordinate grid and are able to survey and interact with their surrounding environment, including interactions with neighboring agents and the space in which they move. Agents exhibit different behaviors (e.g., migration, division, death), and their attributes at any given time point can be described by an array of different state variables.

As the simulation progresses in time, individual agents evolve by deciding which behaviors to execute based on their states and those of their neighbors. Agents integrate these states using a literature based rule-set to determine what behaviors they should carry out. The state of each cell and virus is followed over time, and their actions, interactions, and fate/lineage are recorded for each agent. Interactions between agents and their environment give rise to altered system level properties over time, such as co-receptor switch or clearance of the virus from the system.

Since the model is a discrete event computational method, the updating of the variables occurs via a series of iterations as the model executes. Therefore there are no kinetic equations per se for the protein interactions or the metabolic pathways modeled by the agent rules. Rather, the metabolic rules consist of arithmetical relationships based on the prior state (value) of a particular variable used to calculate the current value.

While the implementation of the conceptual model as a CA experiences the lack of "clarity" in interpreting the results compared to partial differential equations, several advantages outweigh the shortcomings in this context. First is the strength in exploiting the spatial interactions, which are crucial while modeling virus-host interactions. The system we are interested in has a distinct spatial component. Viruses can infect cells that are close-by and otherwise they are cleared by various components of the immune system. In-fact, recent estimates of the clearance rates in lymphoid tissue is very high (10–100 day−1) [56] for typical situations where most of virus in lymphoid tissue is associated with follicular dendritic cells. Therefore it is easy to recognize the importance of finding “a right target cell” within its short lifespan from the viral standpoint.

Second is the inherent stochasticity in CAs, which lends itself to a natural modeling paradigm for describing apparent variation in the real life phenomena. Every person has a different immune system. We all have different concentrations of specialized T cells, different health backgrounds, age and sex differences which all play a role in the infection outcome. Hence, a modeling paradigm allowing appropriate stochasticity is certainly more suitable for the purpose of analysis than assuming an average individual (as in deterministic modeling such as with ODEs and PDEs) representing the whole. Furthermore, possibility of direct implementation of the little details in the system (rather than the effects), make CA an attractive paradigm.

As with any computer simulation based on rules obtained from independent experimental literature, careful attention must be paid to the consistency and correctness of reported parameter
values that will be incorporated into the simulation. For this reason, when possible, multiple publications were referenced for each variable. We then evaluated important aspects of the reported data and experimental design from each publication in order to arrive at a value that accurately reflects the field’s understanding, to the best of our knowledge.

**Entities: Cells and Viruses**

The two types of agents defined are T cells and HIV. T cells are defined as static entities of size one (squares object with a surface area equal to 1). They are either naïve T cells ($T_N$) or memory ($T_M$). Both populations of $T_N$ and $T_M$ are defined with identical parameters and variables except for their co-receptor designation (CXCR4 for $T_N$ and CCR5 for $T_M$) and their color for easy interpretation in the visualization (See figure 5.5). The co-receptor designation of each T cell is defined as an internal parameter but not as separate entities.

The size of the grid is ~2x10^4 units and approximately 10^4 T cells (half of them are $T_N$ and the other half $T_M$) reside on the grid in equilibrium without infection in the standard model.

There is a constant influx of $T_N$ and $T_M$ into the grid, and they are able to reproduce. Their rates of proliferation are $r_N$ and $r_M$ and the rates of influx (born rate) are $\lambda_N$ and $\lambda_M$ for naïve and memory subsets respectively. They have a lifespan of approximately $d_N$ and $d_M$. When infected, their lifespan is significantly shorter and denoted in $d_{Ni}$ and $d_{Mi}$ (See table 5.2).
HIV virions are defined as mobile point particles that are released from infected cells and dying T cells (from $T_N$ or from the $T_M$). Their motion is Brownian with step-size 1 per-time-step. They are mobile until they bind to or infecting a healthy T cell. (Multiple infection of cells is not allowed since its longitudinal effect on HIV tropism is not yet clear according to our knowledge.) Once they are bound, they check the probability of successful infection by the co-receptor designation of the cell they are bound to and their own receptor utilization efficiency ($\varepsilon$) parameter. For instance when a R5 tropic ($V_{R5}$) virus binds a memory T cell ($T_m$), the probability of infection at a single time step is 10% (See figure 5.4). If there is no successful infection, they stay bound to the cell and keep checking for successful infection until they are cleared from the system. Their lifespan ($c$) is 3-4 time steps (0.75-1 day).

Note that the transmission of infection is possible only to the target cells in close proximity. This ensures the interactions are local and the global dynamics are dictated by local interactions.

If there is a successful infection, simulation proceeds with the mutation step. The infecting virion can have a mutation with probability $\mu$ times the length of the V3 loop (~100 base pairs) in the RNA chain. It is assumed that V3 loop determines the co-receptor tropism and the changes in

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range / Value</th>
<th>Reference and Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_N$ &amp; $r_m$ (T cell proliferation rate)</td>
<td>0.04</td>
<td>[1], [2].</td>
</tr>
<tr>
<td>$d_N$ &amp; $d_M$ (Activated Healthy T cell &amp; lifespan)</td>
<td>$\sim$ 8-10 days</td>
<td>[3], [4]. Naive and Memory T cell life span approximated to be equal for simplicity.</td>
</tr>
<tr>
<td>$d_N$ &amp; $d_M$ (Infected T cell lifespan)</td>
<td>$\sim$ 1-2 days</td>
<td>[3], [4].</td>
</tr>
<tr>
<td>$c$ (HIV lifespan in lymph node)</td>
<td>0.75 – 1 day</td>
<td>Estimated in plasma values from [5], [6] with neutralization sensitivities of R5 and X4 variants [7].</td>
</tr>
<tr>
<td>$p$ (Infective HIV burst size)</td>
<td>400-1000 virions per cell per day</td>
<td>[8], [9]</td>
</tr>
<tr>
<td>$\mu$ (Mut. rate)</td>
<td>$(10^{-8} - 5\times10^{-3})^*$</td>
<td>Model parameter in the range $^*$.</td>
</tr>
</tbody>
</table>

Table 5.2. Important parameters from the model with references and explanations.
the V3 loop are reflected in the tropism of the progeny. So, the probability of having a point mutation in the V3 loop equals to the probability of having a mutation in the tropism. The mutation can be forward (resulting in tropism change), neutral or lethal with probability 0.3, 0.3 and 0.4 correspondingly [43], [57]. The change of co-receptor tropism by forward mutation is defined in a straightforward manner adopted from the literature [58]. R5 tropic variants (V\textsubscript{R5}) able to change their tropism only for CCR5 preferring dual tropic R5x4 (V\textsubscript{R5x4}) and V\textsubscript{R5x4} variants for both V\textsubscript{R5} (backward mutation) and V\textsubscript{R5x4}, V\textsubscript{R5x4} variants are able to change to V\textsubscript{X4} and V\textsubscript{R5x4} (See figure 5.3).

![Figure 5.3 The implementation of changing co-receptor tropism.](image)

The infecting virion disappears and the infected cell changes its state from healthy to infected. As seen from table 5.2, infected T cells have five times shorter lifespan than their healthy counterparts and they release $p$ virions approximately once per day and when they are dying. Their parameter for tropism is set to the infecting virus’ tropism when they are being released from the infected T cells, unless there is a co-receptor changing mutation during the infection. Finally the cycle completes by free viruses moving and exploring the space around them looking for another healthy cell to bind to. The simulation take place 7,200 time steps (equal to ~ 5 human years)

The above-mentioned algorithm of model dynamics is expressed in the pseudocode as:

```plaintext
to go
1. ask hiv [ infect-tcell () ]
2. ask-concurrent hiv [
   move-hiv ()
   bind-cells ()
   death-hiv () ]
3. create-tcell 250  [...set cell-type "ccr5"...]
   create-tcell 250  [...set cell-type "cxcr4"...]
4. ask-concurrent tcell [
   advance-infection ()
   produce-hivs ()
   death-tcell ()
   reproduce-tcell ()]
5. if (ticks = 7200) [ stop ]
6. tick;
end
```

![Figure 5.4. The pseudocode showing the iteration of events in the model.](image)
The pseudocode shown in figure 5.4 demonstrates the execution of one cycle in the model. In the first step all HIV particles are calling the `infect-tcell` function. If the virus particle is bound to a T cell, `infect-tcell` function changes the state of an healthy T cell to infected with a certain probability based on the co-receptor tropism of the virus particle and the type of the cell.

In the second step, all HIV particles are concurrently calling `move-hiv`, `bind-cells` and `death-hiv` functions in this order. These functions are executed only if the calling virus is of the appropriate type for this function. Hence `move-hiv` function is executed for free virus particles only. If the virus is bound to a cell or infecting it is not executed. The functions `bind-cells` and `death-hiv` functions are executed only if the virus is non-bound (free).

In the third step, the functions responsible for the influx of naïve and memory T cells are executed.

The fourth step is similar to second step but for T cell subsets. It executes concurrent functions of naïve and memory T cell populations. The function `advance-infection` decreases the energy of the infected cells where as `produce-hivs` function simply creates virus particles with a certain probability from the infected cells. `Death-tcell` function checks the energy of each cell and terminates the cell if its energy is 0 or below. `Reproduce-tcell` function is healthy T cell specific and produces new healthy T cells of the same type from the existing T cells with a certain probability.

Fifth step is the timer of the model checking the simulation clock. If it is 7.200, the simulation stops. The last step is the clock of the simulation. It is incremented by one at each turn of the cycle.

**Co-receptor Utilization Efficiency (ε)**

Once bound to a healthy T cell, the infection probability is given by a probabilistic function (See figure 5.5) depending on the co-receptor designation of the target cell and the tropism of the virus bound to it. The parameters for the function are used as in [41] and [42]. The CCR5 co-receptor utilization efficiency of an R5 tropic virus is denoted by $\varepsilon_{R5-CCR5}$ where the “R5” marks the tropism of the virus and CCR5 stands for the type of co-receptor we are interested in.

![Co-receptor Utilization Efficiency Graph](image)

Figure 5.5. The graph of the probability function for co-receptor utilization efficiency ($\varepsilon$). It can be seen that R5 tropic variants can maximum utilize CCR5 co-receptor (hence, $\varepsilon_{R5-CCR5} = 0.1$) where they cannot use CXCR4 ($\varepsilon_{R5-CXCR4} = 0$). Opposite is true for the X4 tropic HIV (($\varepsilon_{X4-CXCR4} = 0.1$, $\varepsilon_{X4-CCR5} = 0$). R5x4 and r5X4 variants can use both co-receptors but with reduced efficiency ($\varepsilon_{R5x4-CCR5} = \varepsilon_{r5X4-CXCR4} = 0.4$ and $\varepsilon_{R5x4-CXCR4} = \varepsilon_{r5X4-CCR5} = 0.25$).
It is evident from figure 5.5 that even for the best-case scenario (R5 tropic virus infecting a memory T cell or X4 tropic virus infecting a naïve T cell), the probability of infection is only 10%. This effect is compensated for by having a multitude of virions around the T cell competing for the infection.

**Interface and User Interaction**

The graphical user interface (GUI) of the implementation is shown in figure 5.6. Control buttons are placed on the left, graphical interfaces of the simulation is in the bottom left and bottom right. The visualization of the current state of the model is placed in the top right corner.

![Graphical User Interface](image)

Figure 5.6 The screenshot from the implementation of the model. The model parameters can be adjusted from the control switches on the upper left corner. The visualization of the model is shown on top right. The graph for infection dynamics showing total number of T cells and types of viruses and the graph for the T cell dynamics showing types and states of T cells are shown on the left bottom and right bottom respectively.
One of the advantages of programming with NetLogo is the strength of the visualization of the simulation and ease of export as a Java applet. The screenshot seen in figure 5.6 has been taken from an instance of the model exported as a Java applet.

**Experiment Design**

Four experiments are designed (three other experiments in addition to the original model explained above) for exploring the parameter space for fuzzy parameters. Each experiment has been conducted fifty times for each of eight different mutation rates in the range of $10^{-8}$ - $5 \times 10^{-3}$. Distinct random generator seeds have been used for every run therefore all experiments are reproducible.

1. **Effect of the Mesh Size**

   It has been described in [52], [53] that the T cell density in a lymph node is approximately 1:2 as employed in our original model. Since this is a fuzzy parameter which might be subject to change, we explored the parameter space by designing two sets of experiments: first with compact grid size of $\sim 10^4$ (density 1:1) and the second is with spacious grid size of $\sim 3 \times 10^4$ (density 1:3).

2. **Effect of the Burst Size**

   Viral burst size $p$ is another ambiguous, yet important parameter. Latest in-vivo estimates are $4 \times 10^4$-$5.5 \times 10^4$ by [50]. We have employed this value in the light of [51] in the standard model as 250 virus particles per step per infected cell. In this experiment we have explored the parameter space for two burst sizes: 100 and 500 virions per step per infected cell.

3. **Effect of Different Configuration of Mutation Outcomes**

   In the model design it has been defined that a mutation can i) be lethal and leaves no progeny, ii) be neutral and does not change the tropism and iii) be forward and changes the tropism with probabilities 0.4, 0.3 and 0.3 correspondingly. In this experiment we explore the parameter space for penalizing more for mutation with values 0.6, 0.2 and 0.2 for lethal, neutral and forward mutations.

4. **Original Configuration**

   The original model is explained in the implementation section with the parameters from table 5.2.

   Analysis of the results and generation of the figures are accomplished using R programming language and MS Excel-2008.

**Resources Used and Performance**

The National Compute Cluster LISA has been used for the parallel execution of the experiments. We ran the all simulations on clusters of 40 nodes (50 simulations per mutation rate x 8 different mutation rates)/(5 serial jobs per CPU x 2 Intel® quad-core Xeon® L5520 with 2.26 GHz clock cycle and 8 MB cache per node) for an average of approximately 24.5 hours (~5 hours per simulation). For each experiment we have used 40 nodes.

The parallelization of the code for single simulation run is possible. However, considering the relatively small size of the simulation and the theoretical nature of the topic (not implemented for daily use) parallelization is beyond the scope of this work.
Results

VALIDATION

Increase in Selection Pressure Results in Increased Stochasticity
As mentioned in Experiment Design section, we have conducted several experiments to explore the parameter space. The mutation rate versus HIV fitness graph for the effect of the Grid size experiment is shown in figure 5.7. Decreasing the size of the Grid, hence increasing the density of the target cell population has only little effect on the system dynamics; on the other hand an increase in the Grid size changes the dynamics. Introducing extra space between target cells increases selection pressure by further reinforcing local infections and rigorously reducing the viral peak at the acute phase. As a result the bottleneck effect is avoided and stochasticity in the system is greatly boosted which is reflected in the standard deviation. The tropism change also became stochastic (Data not shown). In turn the optimum range of the mutation rate is also truncated to $5 \times 10^{-5}$-$5 \times 10^{-4}$. This result was expected: in total fewer viruses were able to reach the next healthy cell in proximity, the viral population needs to generate more mutants to achieve the switch.

Figure 5.7. The Mutation rate vs HIV & T cell Fitness graphs for different grid sizes experiment. a) Number of T cells and b) number of HIVs are shown vs mutation rate in log scale on x-axis. T cell and HIV population values are demonstrated on y-axes. The graphs for standard model configuration (1:2), dense model configuration (1:1) and sparse configuration (1:3) are shown above. The shaded area represents the values that are one sd away from the mean.

The second set of experiments is conducted for exploring the burst size parameter space and the third set is for examining different configurations of mutation outcomes. The results of these experiments are shown in figure 5.8 and figure 5.9.

In figure 5.8, first, decreasing the burst size from 250 virions per step to 100 virions per step disrupts the competition between HIV variants and the selection halts. The system stays in its equilibrium around its initial conditions. Secondly, as expected,
increasing the burst size to 500 virions per step increases the competition between HIV variants and sharpens the selection.

In experiment four, we have changed configurations of the mutation outcomes: lethal, neutral, forward (0.4, 0.3, 0.3) with alternatives (0.6, 0.2, 0.2). The resulting graph is shown in figure 5.9. Surprisingly this parameter swap did not change the behaviour of the model from the standard model as seen from the graph. We have concluded that the system dynamics are not sensitive to the changes in the configuration of the mutation outcome.

Figure 5.8. The Mutation rate vs a) T cell and b) HIV fitness graph for the different burst size experiment. Graphs representing the standard (250 virions per step), low burst size (100 virions per step) and high burst size (500 virions per step) are shown on top.

Figure 5.9 The Mutation rate vs a) T cell & b) HIV fitness graph for the different configuration of mutation outcomes experiment. Graphs representing the standard lethal mutation rate (LMR-Standard) and increased lethal mutation rate (LTR-Inc.) are shown on top.
There Exists an Optimum Range in the Mutation Rate Spectrum for HIV-1 Co-receptor Switch

We have studied the evolution of the HIV-host tropism dynamics under selection pressure for various mutation rates. To find out empirically whether there is an optimal rate of mutation that can maximize adaptation, we performed a series of simulations. In each simulation, we have started with a basic healthy immune system consisting of $T_M$ and $T_N$ in equilibrium (~$5 \times 10^3$ each) and $5 \times 10^2$ R5 tropic HIV particles in a confined 2D space where it has run for 7200 steps (~5 human years) with constant mutation rate.

We have defined the number of HIV particles averaged over fifty replicate simulations at any single point as the fitness criteria for the HIV population at that time point. Figure 5.10 shows the HIV and T cell population dynamics over time from the original model with a mutation rate ($\mu$) of $10^{-5}$. It is clear from the graph that co-receptor switch occurs in most of the replicate runs before step 2000. By step 3000, HIV population consists of mainly R5 and X4 tropic variants. R5X4 variants, which appear earlier, disappear when X4 tropic variants come into play. Later on, the system reaches an equilibrium where the turnover of $T_m$ and $T_N$ by R5 and X4 tropic HIV equals the replenish rate of T cell subsets.

Figure 5.10 The graphs for a) the number of HIV and b) T cells in the model as a function of time with $\mu=10^{-5}$. The x-axis is the time step (ticks) in the model and the y-axis is the number of particles. The results are averaged over 50 simulations and the grey lines represent standard deviation (sd).
One remarking observation (in figure 5.10 as well as in other simulations) is that, after the first time R5x4 emerges in the system, R5 tropic variants follow a decreasing trend due to the competition for T_M. It is only after the appearance of X4, they regain the increasing trend and R5 and X4 together outcompete the dual-tropic variants. This dynamics strongly suggests the possibility of **sympatric speciation** of HIV-1 within its host.

Another observation is the prominent **effects of a bottleneck** in the target cell population on HIV population dynamics. In the acute phase of the infection (~first sixty steps or 15 days) the HIV population is at its all-time high and predominantly consists of R5 tropic variants. Then, if a timely mutation towards R5x4 occurs and manages to stay in the viral quasispecies albeit the severe decline in T_M, it drastically increases its density in the population after the bottleneck. This indicates even though R5x4 variant exists during the viral peak it cannot compete with the massive number of R5 tropic variants. However, right after the bottleneck in the target cell population, when T_M population is severely depleted, R5x4 appears as a competitive variant against R5 in the viral quasispecies.

Figure 5.11 depicts the HIV fitness for all of the nine explored mutation rates spanning five orders of magnitude (10^{-8}-5x10^{-3} mutation per genome per replication) in the standard model. It is clear that the **optimum range of mutation rate** is 3x10^{-5}-5x10^{-4} maximizing HIV fitness and T cell depletion.

It is important to notice the decreasing standard deviation with increasing \( \mu \). This is due to the fact that for small values of \( \mu \), the system is stochastic (may or may not produce a timely co-receptor switch and increase its fitness), however for high mutation rates it converges to its chaotic attractor and becomes deterministic (constantly mutating). For \( \mu = 5x10^{-3} \) and higher values, every infection results in at least one mutation in V3 which also translates to having increased mortality because of the lethal mutation rate of 0.4 from mutation rate outcomes.

We have observed an indirect link between co-receptor switch and developing AIDS since the T cell count was at lowest when the mutation rate was in the optimum range. Note that developing AIDS was not possible in our model because of the constant influx of T cells into the system.
We also report the intervals and the instances (in percentages) of the co-receptor switch in those intervals in table 5.3. Interestingly, the co-receptor switch occurs most of the time before step 2000 (15 months in real time) and even as early as time step 120 (1 month) when the mutation rate is over $10^{-6}$. This scenario, although not realistic in medical terms, indicates the importance of the fluctuations in the target cell range such as the viral peak around time step 60. When the dominant R5-tropic viruses are abundant and there are plenty of memory cells ($T_M$) to infect around them, it seems to be very difficult for newly emerging dual-tropic viruses to compete with them. Therefore when the virus population is going over a bottleneck, due to shortage of memory cells in our case, dual tropic viruses have more chance to survive and dominate afterwards. The early-or-never emergence of the co-receptor switch is probably due to the lack of antibodies, CD8+ T cells and the effect of medication in the model.
<table>
<thead>
<tr>
<th>Mutation rate $\mu$</th>
<th>Time intervals and the percentage of co-receptor switch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-120 steps</td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td>-</td>
</tr>
<tr>
<td>$10^{-7}$</td>
<td>-</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td>20% (10)</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td>42% (21)</td>
</tr>
<tr>
<td>$5\times10^{-5}$</td>
<td>52% (26)</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>60% (30)</td>
</tr>
<tr>
<td>$5\times10^{-4}$</td>
<td>64% (32)</td>
</tr>
<tr>
<td>$3\times10^{-4}$</td>
<td>82% (41)</td>
</tr>
<tr>
<td>$5\times10^{-3}$</td>
<td>100% (50)</td>
</tr>
</tbody>
</table>

Table 5.3 Table showing the timing of the co-receptor switch in the model.

Finally, our original model is consistent with experimental evidence. The optimum range of mutation rate determined from the model coincides perfectly with the experimental estimates of HIV-1 mutation rate of $2.2\times10^{-5}$, $3.4\times10^{-5}$ and $5.4\times10^{-5}$ from [36], [35] and [54] respectively.

**Discussion**

In this chapter we have investigated the Transmission & Mutation hypothesis for HIV-1 co-receptor switch on an individual based model. We have shown that:

1-) there is an optimum band of mutation rate, which maximizes HIV fitness (in terms of viral survival) in the long run. All the experimental estimates of HIV-1 mutation rate falls into this band validating the adequate level of modeling in our study.

2-) for the cases where we observed co-receptor switch the time of the co-receptor switch was dependent on the fluctuations on the HIV population size. Therefore the bottlenecks (stress conditions such as target cell depletion) in the viral population size increased the chance of occurrence of the co-receptor switch therefore reducing the time required for the co-receptor switch.

In line with our simulations, studies with asymptomatic CCR5Δ32 heterozygous individuals with reduced levels of CCR5 expression showed that expanded co-receptor usage of HIV-1 can occur without rapid progression to AIDS as a consequence of limited target cell availability and stronger selection [55].

Consequently, we hypothesize that the target cell range has major influence on determining the dominant viral phenotype in terms of co-receptor tropism. Hence its size, constitution, internal dynamics and structural properties are of crucial importance for understanding the longitudinal dynamics of co-receptor tropism of HIV-1.

3-) R5 tropic variants benefit from the emergence of X4 tropic variants and outcompete dual-tropic variants together leading to sympatric speciation of HIV-1 *in-vivo*. This we speculate is also linked to in-vivo compartmentalization of the HIV-1 while adapting to its host. Therefore we would like to underline the importance of modeling work describing the cell-cell interactions, compartmentalization of the lymph node and the regulation of cells migration (See Baldazzi et al. [62]).

It is important to note that in this model the parameters of the naïve and memory T cell subsets are assumed to be the same for simplicity. An important extension for this model would be the implementation of the immune system dynamics. This would allow for observing the effects of a healthy immune system over a compromised immune system on the co-receptor switch.
In summary, we have developed a model for the Transmission & Mutation hypothesis of HIV-1 co-receptor switch, tested it for its sensitivity for the key parameters and validated against experimental estimates of the HIV-1 mutation rate.

**Conclusion**

In this chapter we have introduced the temporal scale to our approach and explained an evolutionary model about two well-established co-receptors of HIV-1: CCR5 and CXCR4. We have shown that simply with random mutations and the fluctuations in the T cell population, it is possible to observe the co-receptor switch.

In the next chapter we will summarize our findings throughout the thesis and will finalize with concluding remarks.

**Chapter References**


HIV-1 Co-receptor Tropism


[34] Perelson, Alan S. Modelling viral and immune system dynamics. 1, Jan 1, 2002, Nature reviews Immunology, Vol. 2, pp. 28-36. 10.1038/nri700. 1474-1733.


[44] Hazebergh, Mette D, et al. T-cell division in human immunodeficiency virus (HIV)-1 infection is mainly due to immune activation: a longitudinal analysis in patients before
and during highly active antiretroviral therapy (HAART). 1, Jan 1, 2000, Blood, Vol. 95, pp. 249-55. 0006-4971.


