Bulletin of Mathematical Biology (2002) **64**, 29–64
doi:10.1006/bulm.2001.0266
Available online at http://www.idealibrary.com on IDEAL®

HIV-1 Infection and Low Steady State Viral Loads

DUNCAN S. CALLAWAY
Department of Theoretical and Applied Mechanics,
Cornell University,
Ithaca,
NY 14853, U.S.A.
E-mail: dc52@cornell.edu

ALAN S. PERELSON*
Theoretical Biology and Biophysics,
Los Alamos National Laboratory,
Los Alamos,
NM 87545, U.S.A.
E-mail: asp@receptor.t10.lanl.gov

Highly active antiretroviral therapy (HAART) reduces the viral burden in human immunodeficiency virus type 1 (HIV-1) infected patients below the threshold of detectability. However, substantial evidence indicates that viral replication persists in these individuals. In this paper we examine the ability of several biologically motivated models of HIV-1 dynamics to explain sustained low viral loads. At or near drug efficacies that result in steady state viral loads below detectability, most models are extremely sensitive to small changes in drug efficacy. We argue that if these models reflect reality many patients should have cleared the virus, contrary to observation. We find that a model in which the infected cell death rate is dependent on the infected cell density does not suffer this shortcoming. The shortcoming is also overcome in two more conventional models that include small populations of cells in which the drug is less effective than in the main population, suggesting that difficulties with drug penetrance and maintenance of effective intracellular drug concentrations in all cells susceptible to HIV infection may underlie ongoing viral replication.

1. INTRODUCTION

Highly active antiretroviral therapy (HAART), containing a combination of drugs that inhibit the replication of human immunodeficiency virus type 1 (HIV-1), has proven to be extremely effective at reducing the amount of virus in the blood and tissues of infected patients. In many patients the viral load becomes undetectable, and the possibility of eradicating the virus from the human body has been recognized as a real goal. However, the currently available drugs still lack the potency to completely stop HIV-1 replication and allow for the total decay of latently infected

*Author to whom correspondence should be addressed.

0092-8240/02/010029 + 36 $35.00/0

lymphocytes (Chun *et al*., 1997; Finzi *et al*., 1997; Wong *et al*., 1997; Ramratnam *et al*., 2000; Siliciano and Siliciano, 2000) and HIV-1 trapped on follicular dendritic cells (Hlavacek *et al*., 1999, 2000a,b; Smith *et al*., 2001). The detection of continued viral evolution and replicative intermediates, such as intracellular HIV-1 mRNA and cDNA episomes, in infected patients with undetectable (i.e., below 50 copies of HIV-1 RNA per ml) plasma virus (Furtado *et al*., 1999; Lewin *et al*., 1999; Natarajan *et al*., 1999; Zhang *et al*., 1999; Sharkey *et al*., 2000), as well as detection of HIV-1 RNA in plasma by ultrasensitive assays (Dornadula *et al*., 1999; Yerly *et al*., 2000) in patients classified as having undetectable viral loads by standard assays indicate that replication is ongoing despite the high efficacy of HAART. Furthermore, intermittent episodes of detectable viremia, or 'blips' in viral load are often observed in these otherwise well suppressed patients, and the frequency of these episodes has been correlated with a slower decrease in the size of the latently infected pool (Ramratnam *et al*., 2000).

The aim of this paper is to gain insight into the mechanisms responsible for sustained, yet undetectable, viral loads. Bonhoeffer *et al*. (1997) set the precedent for this type of work by establishing that several models are insufficient to describe the long term effects of reverse transcriptase (RT) inhibitor monotherapy because the steady state viral load in these models is extremely sensitive to minor changes in drug efficacy.

We will argue that the constraints on models of HIV-1 dynamics are more stringent than those suggested by Bonhoeffer *et al*. (1997). This is due to the potency of HAART, which reduces viral load by as many as four orders of magnitude, substantially more than the drops of only one to two orders of magnitude due to single RT inhibitors studied by Bonhoeffer *et al*. After reviewing several models that are not suitable for simulating low steady states, we present three that are. In the first the rate that governs infected cell clearance is a function of infected cell density, rather than constant. Though we must choose the functional form to be slightly different than what empirical evidence suggests (Holte *et al*., 2001), this model's ability to simulate a low steady state viral load highlights the importance of subtle nonlinearities in the model form. The remaining two successful models make use of heterogeneities in drug efficacy (either spatial or phenotypic) and we argue on this basis that there is a growing need for further research to identify subcompartments of target cells in which the currently available drugs are ineffective.

2. THE BASIC MODEL

2.1. ***Steady state analysis.*** We begin by introducing a model that is the basis for many mathematical studies of HIV-1 dynamics, in particular for extensions that have been used to estimate virus and infected cell decay rates *in vivo* (Ho *et al*., 1995; Wei *et al*., 1995; Perelson *et al*., 1996). As in Bonhoeffer *et al*. (1997), we will explain that a shortcoming of this model is its inability to describe sustained, low level production of virus under antiretroviral pressure.

Let T represent CD4$^+$ cells that are susceptible to infection, T^* productively infected cells, V_I infectious virus, and V_{NI} virus made noninfectious by the action of protease inhibitors (Perelson *et al*., 1996), respectively. The following equations represent the rates of change of these populations under the effect of RT and protease inhibitors:

$$\frac{dT}{dt} = \lambda - dT - (1-\kappa)kV_I T \tag{2.1a}$$

$$\frac{dT^*}{dt} = (1-\kappa)kV_I T - \delta T^* \tag{2.1b}$$

$$\frac{dV_I}{dt} = (1-\eta)N_T \delta T^* - cV_I \tag{2.1c}$$

$$\frac{dV_{NI}}{dt} = \eta N_T \delta T^* - cV_{NI}. \tag{2.1d}$$

The constant λ represents a source of susceptible cells, and d is their death rate. k is the infection rate constant, and infection is assumed to occur at a rate proportional to the product of the concentration of virus and target cells, an assumption which is valid for a well-mixed system with relatively high concentrations of each product. δ is the infected cell death rate, N_T is the number of free viral particles produced during the average infected T cell life span, and c is the rate at which free virus is cleared. In the Appendix we provide motivation for the values used for these parameters in model simulations.

RT inhibitors prevent HIV RNA from being converted into DNA, a crucial part of the viral life cycle. Thus, the infectiousness of the virus, k, is reduced by the quantity $(1-\kappa)$, where κ represents the efficacy of RT inhibitors and $0 \leq \kappa \leq 1$. Protease inhibitors, on the other hand, do not directly inhibit the infectiousness of virus. Rather, they alter part of the viral assembly process in the final stage of the viral lifecycle, and as a result cause the production of defective, noninfectious virus. Here, η, the protease inhibitor efficacy, is the fraction of total virus produced that is noninfectious due to the action of the protease inhibitor, and $0 \leq \eta \leq 1$.

Expressions for the steady state quantity of each variable T, T^*, V_I and V_{NI}, can be found by setting the right-hand sides of equation (2.1) equal to zero and solving the resulting algebraic equations. In terms of the parameters above, the amount of virus present at this steady state is either

$$\overline{V} = 0, \tag{2.2}$$

or

$$\overline{V} = \frac{\lambda N_T}{c} - \frac{d}{k(1-(\kappa+\eta-\kappa\eta))}, \tag{2.3}$$

where $V \equiv V_I + V_{NI}$. We shall refer to the first case as the trivial equilibrium, and will focus on the behavior of the latter.

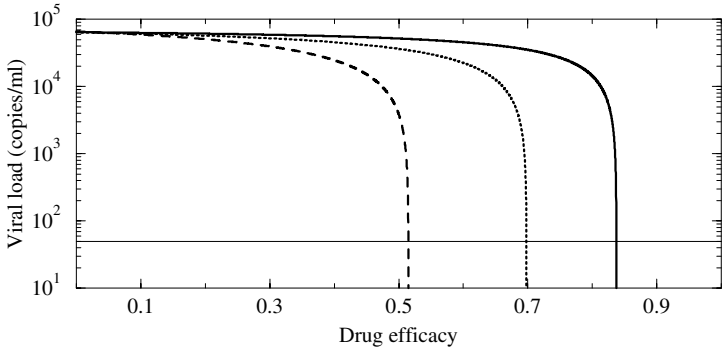

Figure 1. Virus load vs drug efficacy for the basic model [equation (2.1)]. Fixed parameters are defined in the Appendix. In addition, $k = 8\times10^{-7}$ day^{-1} and $N_T = 100$ (solid curve), $k = 3.6 \times 10^{-7}$ day^{-1} and $N_T = 119.6$ (dotted curve), and $k = 1.6 \times 10^{-7}$ day^{-1} and $N_T = 167.5$ (dashed curve).

The composite parameter $(\kappa+\eta-\kappa\eta)$ represents the total combined drug efficacy which we will henceforth refer to as ε. Note that $(1-\varepsilon) = (1-\kappa)(1-\eta)$; this rearrangement indicates that we are assuming the drugs act independently of one another. The basic reproductive ratio prior to drug therapy for this system is $R_0 = \lambda N_T k/cd$. This quantity represents the average number of secondary infected cells which will result from the introduction of a single infected cell into a population of completely uninfected CD4$^+$ cells (Anderson and May, 1991). When $0 < R_0 < 1$, the population is able to reproduce, but the death rate exceeds the growth rate and the population size will asymptotically approach zero. In terms of ε and R_0, the steady state viral load is

$$\overline{V} = \frac{d}{k}\left(R_0 - \frac{1}{1-\varepsilon}\right). \tag{2.4}$$

Figure 1 depicts the relationship between steady state viral load and total drug efficacy for three different choices of parameters. Increasing drug efficacy causes the infection rate to decrease, which in turn increases the number of available target cells. For efficacies far from the critical efficacy, this increase in the target cell pool size allows the virus to maintain a large population despite its reduced infectiousness. It is at least in part due to this predator–prey relationship that the steady state viral load curve in Fig. 1 is flat for small efficacies. However, the general shape of the curve is concave down; this results from the inverse relationship between $\overline{V}$ and ε.

It follows from the inverse relationship that a large range of the possible steady state viral loads during drug therapy occurs in only a small range of efficacies. This effect is apparent in Fig. 1, and can be shown more precisely as follows: let V_0 be the steady state viral load in the absence of therapy, and let f be the fraction of the pre-therapy steady state viral load present for a given drug efficacy, so that if

drug therapy reduces the viral load by 2 logs, f would be 10^{-2}, a 4 log drop would correspond to $f = 10^{-4}$, and so on. This can be written as

$$f = \frac{\overline{V}}{V_0} = \frac{R_0 - \frac{1}{1-\varepsilon}}{R_0 - 1}$$

which, when solved for ε, yields

$$\varepsilon = 1 - \frac{1}{R_0 - (R_0 - 1) f}. \tag{2.5}$$

We shall refer to the drug efficacy, which corresponds to $\overline{V} = 0$ or equivalently $f = 0$ as the 'critical efficacy' (ε_0). Thus,

$$\varepsilon_0 = 1 - \frac{1}{R_0}. \tag{2.6}$$

For $\varepsilon > \varepsilon_0$ the infected steady state becomes negative and exchanges stability with the uninfected steady state, i.e., the drug will extinguish the viral population, even if the virus continues to replicate.

Equation (2.5) can be used to obtain an upper bound on the difference between the critical efficacy and any other efficacy based on the factor of reduction of steady state viral load, as follows:

$$\varepsilon_0 - \varepsilon = \frac{1}{R_0}\left(\frac{1}{1 - \frac{R_0 - 1}{R_0} f} - 1\right) < \frac{1}{1 - f} - 1 \sim f \tag{2.7}$$

where we have assumed that $R_0 > 1$ and $f \ll 1$. This means that for any choice of R_0, as long as the drop in virus load is more than 1 log ($f < 0.1$), the difference between the critical efficacy and any other efficacy is smaller than the factor of reduction. For example, the difference between ε_0 and ε is less than 0.01 for a 2 log drop in steady state viral load. Drops in viral load of at least this size are commonly observed in clinical studies of patients on HAART, meaning that if this model is correct, many patients experience drug efficacies extremely close to the critical efficacy. However, because it appears that the viral population is not extinguished in any patients studied thus far (Chun *et al*., 1999; Dornadula *et al*., 1999; Furtado *et al*., 1999; Lewin *et al*., 1999; Natarajan *et al*., 1999; Zhang *et al*., 1999; Sharkey *et al*., 2000), one must conclude that no patients have experienced drug efficacies greater than or equal to the critical efficacy.

These observations are in conflict. The results of the model indicate that extinction *should* occur, because there must be a distribution of possible efficacies: the drug will not be equally potent in every patient, and because the probability of experiencing an efficacy extremely close to the critical efficacy is high, one would expect that the tail of the distribution should include efficacies greater than the

critical efficacy, and hence the virus should go extinct in at least some patients. However, clinical results indicate that extinction of the viral population does not occur under HAART. Thus, models with a concave down relationship between steady state viral load and efficacy contradict clinical results. This presents the opportunity to search for models that do not bear this shortcoming, and consequently gain insight into the mechanisms which do and do not play a role in maintaining low viral loads under drug therapy.

Bonhoeffer *et al*. (1997) argue that the shortcomings of this steady state relationship can be alleviated by using models with a linear relationship between steady state viral load and drug efficacy. Note, however, that a relationship very similar to equation (2.7) holds for models with a linear relationship between steady state viral load and drug efficacy. Let $\overline{V} = a - b\varepsilon$ where a and b are positive constants. Then at the critical efficacy $\overline{V} = a - b\varepsilon_0 = 0$ and so $\varepsilon_0 = a/b$. For this relationship, the fraction of steady state viral load remaining after therapy would be $f = \overline{V}/V_0 = (a - b\varepsilon)/a = 1 - \varepsilon/\varepsilon_0$, and thus $\varepsilon = (1 - f)\varepsilon_0$. So we can write $\varepsilon_0 - \varepsilon = \varepsilon_0 f < f$. The final condition is identical to equation (2.7), meaning that the constraints drug therapy places on models with a linear relationship between steady state viral load and drug efficacy, while potentially less severe, are nonetheless also poorly suited for modeling low steady state viral loads.

2.2. ***Decay rates as a function of drug efficacy in the basic model.*** If we turn from steady states to look at the dynamic behavior of the model given by equation (2.1), we are presented with another difficulty. Figure 2 shows therapy simulations for drug efficacies that would result in a 1 log drop and a 2 log drop in steady state viral load compared to a simulation with an efficacy that would extinguish the virus. The minimum value of these trajectories, even in cases where the populations ultimately recover, is well below one viral particle in a patient, and on this basis we would expect extinction in all three cases, not just when the efficacy is greater than the critical efficacy. Beyond this shortcoming, these dynamics suggest that steady state viral loads one order of magnitude lower than the set point would not be observed until long after (e.g., 3 years) the initiation of therapy. On the contrary, patients on RT inhibitor monotherapy typically show viral loads settling very quickly to steady states 1 log below the original viral load (Bonhoeffer *et al*., 1997). The two other curves show that over the first 2 years of therapy the dynamics that precede attaining a steady state with 2 log drop in viral load are practically indistinguishable from those for which the virus would become extinct, and that during the transient the viral load will likely remain below detectable levels for well over 10 years.

This could be an explanation of sustained reductions in viral load: the true steady state viral load could be detectable, but the time required to reach this state is extremely long. In the mean time, the viral load would remain undetectable. This would, however, still require that the drug efficacy be quite close to the critical efficacy for sustained long term undetectability, and the arguments

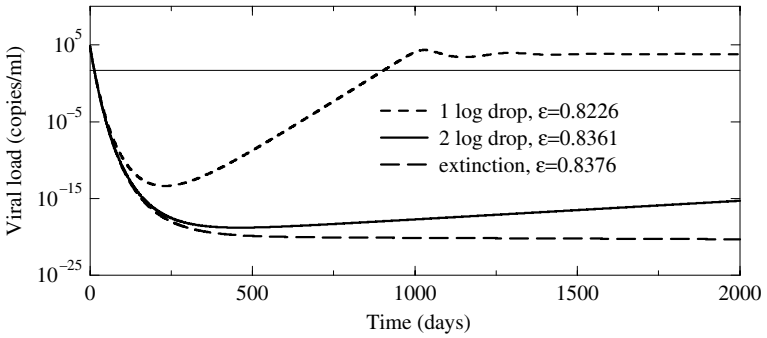

Figure 2. Decay of virus just above and below the critical efficacy, $\varepsilon_0 = 0.8375$, as predicted by the basic model [equation (2.1)]. Therapy starts at $t = 0$. The dashed curve follows the dynamics of a system that settles to a steady state viral load 1 log lower than the pre-drug amount ($\varepsilon = 0.82265$), the solid curve is for a system that will ultimately settle to a steady state 2 logs below the set point ($\varepsilon = 0.83613$), and the long-dashed line will continue to decay indefinitely ($\varepsilon = 0.8376$). Other parameters are defined in the Appendix. For these simulations it was assumed that drug efficacy was due solely to RT inhibitors, i.e., $\eta = 0$ and thus $\kappa = \varepsilon$. The thin dotted curve delineates 50 copies ml^{-1}, the detection threshold.

presented in Section 2.1 would also apply here. Furthermore, if one takes this behavior as correct, then it remains to be seen what mechanisms would lead to the rapid convergence to steady state in patients receiving RT monotherapy (Bonhoeffer *et al.*, 1997). Another difficulty we encounter with these dynamics is that the minimum value of the oscillation is well below one virion/patient $\approx \frac{1}{20l} \times 10^{-3}$ L ml^{-1} $= 5 \times 10^{-5}$ virions/ml, where we have used the fact that a typical patient has approximately 20 l of fluid in which HIV replicates. Thus if we interpret viral loads below this amount as viral extinction, this model predicts extinction would occur long before even a 1 log drop steady state is attained.

Though the parameters used in these simulations are derived directly from experimental data (see Appendix), changing some or all of the parameters could alleviate the problem of long convergence times. To understand in more detail the rates of convergence to steady state, we can simplify equation (2.1) to a two-dimensional system. Once T^* is determined, equation (2.1d) decouples from equation (2.1). Furthermore, the production and decay of virus is fast enough that we can assume viral load changes instantaneously with the number of infected cells, and thus substitute $V_I = (1-\eta)N_T\delta T^*/c$ and eliminate equation (2.1c) as well. This leads to

$$\frac{dT}{dt} = \lambda - dT - (1-\varepsilon)\beta T^* T \tag{2.8a}$$

$$\frac{dT^*}{dt} = (1-\varepsilon)\beta T^* T - \delta T^* \tag{2.8b}$$

where we have used $\beta = kN_T\delta/c$. The number of target and infected cells at the

nontrivial equilibrium is

$$\overline{T} = \frac{\delta}{\beta(1-\varepsilon)} \tag{2.9a}$$

$$\overline{T^*} = \frac{\lambda\beta(1-\varepsilon) - d\delta}{\delta\beta(1-\varepsilon)}. \tag{2.9b}$$

Because the system is two dimensional, the eigenvalues for this fixed point are readily obtainable. They are

$$\Lambda = \frac{-\alpha \pm \sqrt{\alpha^2 - 4\delta^2(\alpha - d\delta)}}{\delta} \tag{2.10}$$

where $\alpha = \lambda\beta(1-\varepsilon)$. If the eigenvalues are complex (as in the two curves that do not lead to extinction in Fig. 2), the real part of the eigenvalue simplifies to

$$\Re(\Lambda) = -\frac{\lambda\beta(1-\varepsilon)}{\delta}. \tag{2.11}$$

Clearly, significant changes in any one of these parameters from their accepted values would be necessary in order to alter the convergence time substantially. Thus, we need to either find new ways to estimate the parameters for this model, or revise the structure of the model itself. Given that the model has already been criticized on the basis of the sensitivity of steady state viral load to changes in drug efficacy, we will invest the remainder of this paper examining new model forms and whether they alleviate these problems.

3. Other Models with a Nonrobust Low Steady State Viral Load

3.1. ***Models with inverse relationships between $\overline{V}$ and drug efficacy.*** We will ultimately proceed to a model for which steady state viral load is less sensitive to changes in drug efficacy, but first we examine some familiar models and show that they also exhibit the inverse relationship between viral load and efficacy discussed in the basic model above.

Notice that the relationship between total drug efficacy and virus load in equation (2.4) remains the same if the system uses RT inhibitors only (i.e., $\eta = 0$) or protease inhibitors only ($\kappa = 0$). Thus, for simplicity, in the following models we will restrict the effect of the drug to reducing the infectiousness of the virus (as with RT inhibitors), rather than including the effect of protease inhibitors as well.

3.1.1. *Examining the mass action assumption.* The infection rates k and β in equations (2.1) and (2.8), respectively, are composite parameters accounting for three mechanisms: first, the rate at which virus and target cells collide, second, the fraction of target cells which are activated and hence susceptible to infection (Stevenson, 1996), and third, the fraction of interactions between activated target cells and virus which result in a productive infection. All of these quantities have been assumed constant, or in the case of virus–cell collision rates, dependent only on the product of the concentrations of those two populations. This last notion, the so-called 'mass action principle' is valid when the system is well mixed (i.e., there are no significant spatial concentration heterogeneities) and there are significant quantities of each reactant. However, interactions between two populations are not always this simple, and the form of such an expression can change, among other ways, as the relative and absolute population sizes vary (in Section 5 we will deal with the effect of spatial concentration heterogeneities using a compartmental model). Consider the following model:

$$\frac{dT}{dt} = \lambda - dT - \frac{(1-\varepsilon)kVT}{\alpha + \nu V + \tau T} \tag{3.1a}$$

$$\frac{dT^*}{dt} = \frac{(1-\varepsilon)kVT}{\alpha + \nu V + \tau T} - \delta T^* \tag{3.1b}$$

$$\frac{dV}{dt} = N_T \delta T^* - cV. \tag{3.1c}$$

The terms in the denominator of the infection term serve to 'saturate' the infection rate when either V or T gets large. For example, if $\nu V \gg \alpha + \tau T$, then $kVT/(\alpha + \nu V + \tau T) \approx kT/\nu$; the result is that the infection rate depends almost exclusively on the number of available $CD4^+$ cells when the amount of virus is relatively large. A similar situation arises when $\tau T \gg \alpha + \nu V$, but in this case the amount of virus present limits the infection rate. A similar 'competitive saturation' term has been used to describe cell growth in models of T cell proliferation and HIV-1 immune control (De Boer and Perelson, 1995, 1998).

As with the basic model, there is only one nontrivial equilibrium. The steady state viral load is

$$\overline{V} = \frac{(1-\varepsilon)k\frac{\lambda N_T}{c} - (\alpha d + \tau\lambda)}{(1-\varepsilon)k + \left(d\nu - \frac{c\tau}{N_T}\right)}. \tag{3.2}$$

This expression has linear multiples of ε in both the numerator and the denominator. Still, the exact nature of the $\overline{V}$–ε curve will depend on parameter choices. Because we have not specified ν, τ and α, this is not immediately obvious. However, we can find the slope and concavity of the $\overline{V}$–ε curve and compare them to Fig. 1 (which is sloping downward and concave down). The slope of the curve is

$$\frac{d\overline{V}}{d\varepsilon} = -\frac{kd\left(\frac{N_T\lambda\nu}{c} + \alpha\right)}{\left((1-\varepsilon)k + d\nu - \frac{c\tau}{N_T}\right)^2}$$

which is negative definite, assuming that α and ν are positive. This is to be expected, since increasing ε should serve to decrease the viral load. The second derivative of $\overline{V}$ with respect to ε is

$$\frac{d^2\overline{V}}{d\varepsilon^2} = -2\frac{k^2 d\frac{N_T\lambda\nu}{c} + \alpha}{\left(k + d\nu - \frac{c\tau}{N_T}\right)^3}.$$

The steady state quantity of productively infected cells, $\overline{T^*} = (\lambda\nu + c\alpha/N_T)/(k + d\nu - c\tau/N_T)$, is greater than zero at the infected steady state. Hence $(k + d\nu - c\tau/N_T) > 0$ and so $d^2\overline{V}/d\varepsilon^2 < 0$, meaning that the curve is concave down. So the relationship between steady state viral load and drug efficacy for equations (3.1a)–(3.1c) is similar to that in equation (2.1). Note that we could choose τ, α, and ν such that $|d^2\overline{V}/d\varepsilon^2| \ll 1$, meaning that the relationship between drug efficacy and viral load would be approximately linear. However, as explained in Section 2, even a linear relationship is not sufficient to explain the low virus loads which result from HAART .

Similar arguments can be made for infection terms of the form $kV^2T/(1+\alpha V)$, $kVT^2/(1+\alpha T)$, and $k[V/(\alpha+\nu V)][T/(\beta+\tau T)]$.

3.1.2. *Logistic growth of target cells.* There is evidence to suggest that target cell populations are governed by logistic growth, where the rate of production of cells is limited by the number of cells already present (Sachsenberg *et al.*, 1998). To incorporate this into the basic model, we can replace the first two terms of equation (2.1a), $\lambda - dT$, by $rT(1 - T/T_{\max})$ where r is the maximum growth rate and $T_{\max}$ is the maximum sustainable number of target cells. For this system of equations there is a single nontrivial equilibrium; the steady state viral load is

$$\overline{V} = \frac{r}{(1-\varepsilon)k}\left(1 - \frac{c}{(1-\varepsilon)kT_{\max}N_T}\right). \tag{3.3}$$

Near the critical efficacy, specifically for $\varepsilon > 1 - 2c/kT_{\max}N_T$, this function is a downward sloping, concave down function of ε. Hence the incorporation of a logistic growth term does not alleviate the sensitive dependence of viral load on drug efficacy at low viral numbers.

Interestingly, for $\varepsilon < 1 - 2c/kT_{\max}N_T$, this expression is an increasing function of ε; increasing drug efficacy actually increases steady state viral load. This may or may not compromise the validity of this model, since parameters could be chosen such that $1 - 2c/kT_{\max}N_T < 0$. Then the above condition on ε would never be satisfied, because $0 < \varepsilon < 1$.

Other forms of target cell growth and death do not improve the relationship between steady state viral load and drug efficacy (Bonhoeffer *et al.*, 1997). In particular, replacing $\lambda - dT$, in equation (2.1a) by $rT(1 - T/T_{\max}) - dT$, $\lambda + rT(1 - T/T_{\max})$ or $\lambda + rT(1 - T/T_{\max}) - dT$ still results in a concave down relationship.

Using $\lambda/(1+\alpha T) - dT$ as the source term generates a concave down relationship as well, however, as α becomes large, the relationship becomes nearly linear. In Section 3.2 we will discuss more models which exhibit this type of relationship.

3.1.3. *Quiescent cell populations.* As we explained in Section 3.1.1, the fraction of target cells which are activated and thus permissive to infection are typically considered to be constant and consequently absorbed into the infection rate constant k. Here we relax this assumption by including a population of resting cells from which activated cells are derived upon antigenic stimulation. Let the variable Q represent this quiescent population of cells. Then we may write

$$\frac{dQ}{dt} = \lambda - d_Q Q - \theta(V+B)Q \tag{3.4a}$$

$$\frac{dT}{dt} = s\theta(V+B)Q - dT - (1-\varepsilon)kVT \tag{3.4b}$$

$$\frac{dT^*}{dt} = (1-\varepsilon)kVT - \delta T^* \tag{3.4c}$$

$$\frac{dV}{dt} = N_T \delta T^* - cV \tag{3.4d}$$

where d_Q is the death rate of quiescent cells and s is the factor by which cells proliferate upon activation. The term $\theta(V+B)Q$ represents the rate at which activation occurs, and comes from assuming that HIV-1 and any other antigen (quantified by the parameter B) interact with quiescent cells according to mass action and activate these cells with a rate constant θ. An extended version of this model was used previously to study the effect of the active cell population on multiple strain competition and disease progression (Callaway *et al.*, 1999).

Only one nontrivial equilibrium point exists for this system. The steady state viral load is

$$\overline{V} = \frac{\lambda}{\gamma\theta} - \frac{(d_Q + \theta B)}{\theta}, \tag{3.5}$$

where

$$\gamma = \frac{1}{2}\left(\frac{\lambda}{d_Q} + \frac{cB}{sN_T d_Q} + \frac{c}{s\theta N_T} - \frac{dc}{(1-\varepsilon)ksN_T d_Q}\right)$$
$$-\frac{1}{2}\sqrt{\left(-\frac{\lambda}{d_Q} - \frac{cB}{sN_T d_Q} - \frac{c}{s\theta N_T} + \frac{dc}{(1-\varepsilon)ksN_T d_Q}\right)^2 - \frac{4c\lambda}{s\theta N_T d_Q}}.$$

The manner in which $\overline{V}$ depends on ε is not immediately obvious from this result. However, it is clear from Fig. 3 that the relationship is similar to that observed for equations (2.1a)–(2.1c), in particular, the slope of the curve continues to decrease

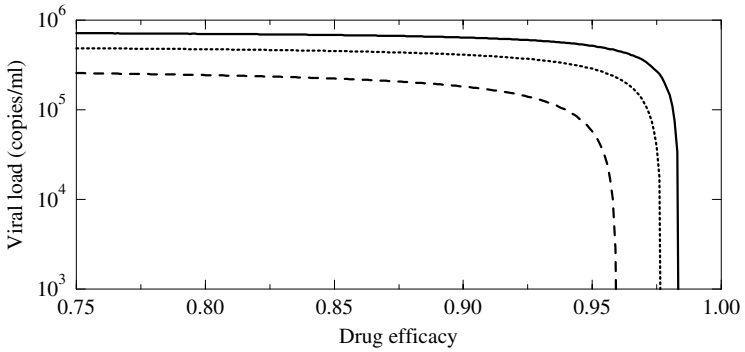

Figure 3. Virus load vs drug efficacy for the quiescent cell model [equation (3.4)]. Parameters are defined in the Appendix, and $B = 10$, $\theta = 10^{-3}$ and $s = 10$ (solid curve), $s = 7$ (dotted curve) and $s = 4$ (dashed curve). The curve was insensitive to changes in B and θ.

to the point of virus extinction, implying that the steady state viral load becomes increasingly sensitive to changes in drug efficacy. This is true for several parameter combinations, suggesting that the result is not specific to one set of parameters.

3.1.4. *Chronically infected cells.* Drug perturbation studies have found that as plasma HIV-1 decays under drug therapy, there are several observable phases of the decay. The most rapid occurs as free virus is cleared from the blood (due to the rate constant c) and establishes a quasi-steady state with the number of infected cells. The next phase is due to the decline to productively infected cells, and has been used to determine the rate constant δ. This phase is characterized by an exponential fall in V, and has been termed the 'first' (observable) phase. A later (second) phase has been observed, and this has been attributed to the decay of 'chronically' infected cells, or cells which produce much smaller amounts of virus than the main population of infected cells, and, perhaps as a consequence, die at a much slower rate. A standard representation (Perelson *et al.*, 1997) of a system with this population of infected cells included is

$$\frac{dT}{dt} = \lambda - dT - (1-\varepsilon)kVT \tag{3.6a}$$

$$\frac{dT^*}{dt} = (1-\alpha)(1-\varepsilon)kVT - \delta T^* \tag{3.6b}$$

$$\frac{dC^*}{dt} = \alpha(1-\varepsilon)kVT - \mu C^* \tag{3.6c}$$

$$\frac{dV}{dt} = N_T \delta T^* + N_C \mu C^* - cV \tag{3.6d}$$

where C^* is the population of chronically infected cells. A fraction α of infection events result in chronic infection, chronically infected cells die with a rate constant

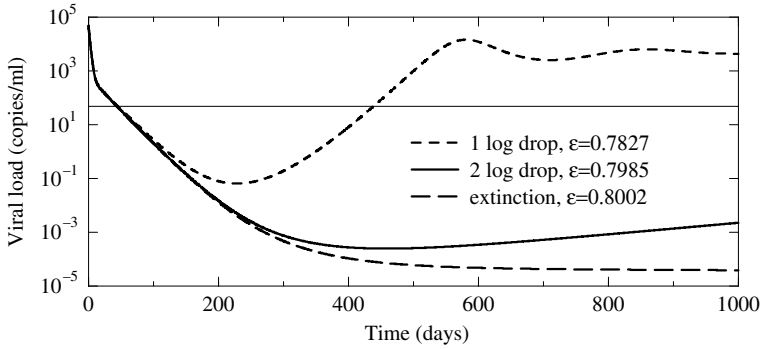

Figure 4. Virus load vs time for the chronically infected cell model [equation (3.6)]. Note that the minimum value of the oscillation is much higher than in the basic model. Parameters are defined in the Appendix, and the critical efficacy, ε_0, is 0.8001. The thin line represents 50 copies ml^{-1}, the detection threshold.

μ, and N_T and N_C are the average number of virions produced in the lifetime of short-lived and chronically infected cells, respectively.

As with the previous models, there is only one nontrivial fixed point. This gives a steady state viral load of

$$\overline{V} = \frac{(1-\alpha)\lambda N_T + \alpha\lambda N_C}{c} - \frac{d}{(1-\varepsilon)k}. \tag{3.7}$$

Again, there is an inverse relationship between steady state viral load and drug efficacy. However, the chronically infected population may play a different role in preventing extinction. Figure 4 shows that the addition of chronically infected cells serves to dampen the oscillations in viral load predicted by the basic model (see Fig. 2), and consequently prevents the minimum viral load from reaching values of less than one per patient.

If we consider extinction to occur when the viral load is less than one virion/patient, i.e., $V < 5 \times 10^{-5}$ ml^{-1} (see Section 2.2), Fig. 4 demonstrates that if the drug efficacy is only slightly greater than the critical efficacy extinction will occur within 1000 days, or approximately 3 years. Yet clinical trials have followed patients on HAART for more than 3 years and evidence of extinction has yet to be seen. Thus the results here do not support the possibility that the reason extinction has not been observed is simply because the viral load has yet to decay completely.

3.1.5. *Latently infected cells.* A fraction of CD4$^+$ cells with integrated HIV-1 provirus are in a latent state and do not produce virus until they have been activated. Because of their latent state, these cells are not likely to be subject to immunosurveillance by HIV-specific T cells. Perhaps as a consequence, these cells persist during drug therapy (Chun *et al.*, 1997; Finzi *et al.*, 1997; Wong *et al.*, 1997), either by replenishment from ongoing replication or an extremely long half-life. Latently

infected cells thus pose a serious obstacle to eradicating the virus, suggesting that this population of cells could play a crucial role in maintaining a low steady state viral load during therapy. The following equations describe this mechanism:

$$\frac{dT}{dt} = \lambda - dT - (1-\varepsilon)kVT \tag{3.8a}$$

$$\frac{dT^*}{dt} = (1-\alpha)(1-\varepsilon)kVT - \delta T^* + aL \tag{3.8b}$$

$$\frac{dL}{dt} = \alpha(1-\varepsilon)kVT - \delta_L L - aL \tag{3.8c}$$

$$\frac{dV}{dt} = N_T \delta T^* - cV \tag{3.8d}$$

where L is the population of latently infected cells. The parameter α, the fraction of infections which result in latency rather than the active production of HIV-1 particles, serves the same purpose as it did in the chronically infected cell model [equation (3.6)]. The new parameters in this system are a, the rate at which latently infected cells become activated and produce virus and δ_L, the death rate of latently infected cells. δ_L should be substantially smaller than δ, since latently infected cells are less susceptible to cell mediated killing and death due to viral cytopathicity.

There is a single nontrivial equilibrium, which gives a steady state viral load of

$$\overline{V} = \frac{\lambda N_T}{c}\left(1 - \frac{\delta_L \alpha}{\delta_L + a}\right) - \frac{d}{k(1-\varepsilon)}, \tag{3.9}$$

which bears the familiar inverse relationship between steady state viral load and drug efficacy. As with chronically infected cells, the steady state remains sensitive to the precise value of the drug efficacy. Yet because latently infected cells persist after years of infection (Chun *et al.*, 1997; Finzi *et al.*, 1997; Wong *et al.*, 1997), it is possible that the steady state viral load is never obtained prior to the patient's death or discontinuation of therapy. Rather, the system is in a state of extremely slow decay, suggesting that this model is not necessarily an incorrect representation of the dynamics of HIV-1 infection under drug therapy. However, Fig. 5 shows that, although this slow decay could lead to persistence of the viral population, the kinds of dynamics predicted by the model are also extremely sensitive to changes in drug efficacy. Furthermore, as with the basic model, the time needed to reach a steady state one order of magnitude below the set point is much longer than is observed in patients on RT inhibitor monotherapy (Bonhoeffer *et al.*, 1997).

If we examine duration of time required to reach extinction (that is, the time required for the viral load to decay below 5×10^{-5} ml^{-1}, see Section 2.2), then this model could provide an explanation for why extinction has not yet been observed in clinical studies. When the drug efficacy exceeds but is near the critical efficacy, more than 10 years can pass before the viral load diminishes below one

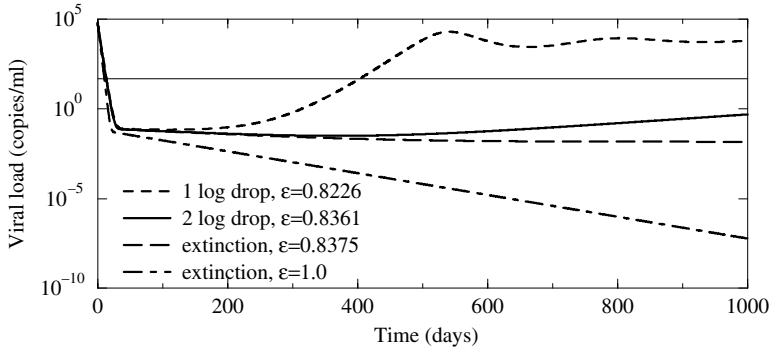

Figure 5. Virus load vs time for the latently infected cell model [equation (3.8)]. Parameters are defined in the Appendix, $a = 0.01$ day^{-1} and $\delta_L = 0.004$ day^{-1}. The thin line represents 50 copies ml^{-1}, the detection threshold.

virion/patient. Thus, this model predicts that it could take many years of observation before extinction is observed in a clinical setting.

3.1.6. *Follicular dendritic cells.* Follicular dendritic cells (FDC) reside in secondary lymphoid organs known as germinal centers and bind HIV–antibody complexes. HIV-1 trapped on FDC is thought to be a significant viral reservoir (Pantaleo *et al*., 1993, 1994) and binds up to 10^{11} copies of HIV-1 RNA in untreated patients (Haase *et al*., 1996; Cavert *et al*., 1997). Consequently, FDC have been cited as a possible impediment to HIV-1 eradication (Smith *et al*., 2001), and their effect on viral dynamics has been addressed (Hlavacek *et al*., 1999, 2000a,b). To understand what effect FDC could have on the steady state viral load, consider the following model:

$$\frac{dT}{dt} = \lambda - dT - (1-\varepsilon)kVT \tag{3.10a}$$

$$\frac{dT^*}{dt} = (1-\varepsilon)kVT - \delta T^* \tag{3.10b}$$

$$\frac{dV}{dt} = N_T \delta T^* - (c+b)VuV_b \tag{3.10c}$$

$$\frac{dV_b}{dt} = bV - uV_b - c_b V_b \tag{3.10d}$$

where V_b represents HIV-1 particles bound to FDC. FDC bind free virus at rate b, bound virus dissociates from FDC at rate u, and bound virus is cleared at rate c_b. Notice that clearance of free virus is now governed by two parameters, suggesting that the clearance parameter c in previous models represents a composite of at least two processes.

As we have seen with the preceding models, this system has only one nontrivial

fixed point. It gives a steady state viral load of

$$\overline{V} = \frac{\lambda N_T (u + c_b)}{(uc + c_b c + c_b b)} - \frac{d}{(1-\varepsilon)k}\,, \tag{3.11}$$

which is identical in its ε-dependence to equation (2.3), and thus has a sensitive dependence on drug efficacy. However, as with the latency model [equation (3.8)], if the decay from FDC is slow enough, the system may never reach steady state, and the virus load would undergo a long, slow decline.

3.1.7. *Differential drug effects.* RT and protease inhibitors work by entering $CD4^+$ cells and disrupting part of the viral lifecycle. Consequently these drugs need to work effectively within diverse cellular environments, since HIV-1 infects multiple cell types (Gartner *et al*., 1986; Koenig *et al*., 1986). In fact, protease inhibitors are known to be less effective in some $CD4^+$ subpopulations than in others (Kim *et al*., 1998; Perno *et al*., 1998; Puddu *et al*., 1999) since P-glycoprotein pumps on the surface of some cell types reduce protease inhibitor concentrations (and thus efficacy) within the cell. Reduced drug efficacy in a subpopulation of cells is among the factors implicated in the cause of ongoing replication in the face of HAART (Zhang *et al*., 1999), suggesting that accounting for this phenomenon in a model might produce robust low steady state viral loads. One such model is as follows:

$$\frac{dT}{dt} = \lambda - dT - (p(1-\varepsilon) + (1-p))kVT \tag{3.12a}$$

$$\frac{dT^*}{dt} = p(1-\varepsilon)kVT - \delta T^* \tag{3.12b}$$

$$\frac{dT_n^*}{dt} = (1-p)kVT - \delta T_n^* \tag{3.12c}$$

$$\frac{dV}{dt} = N_T \delta (T^* + T_n^*) - cV. \tag{3.12d}$$

Here, we assume the target cell population can be split into subpopulations, such that in a fraction p the drug is active with efficacy ε, while in the remaining fraction $(1-p)$ the drug has no effect (e.g., it is pumped out). T^* is the population of infected cells derived from target cells in which the drug was effective, and T_n^* is the population derived from target cells in which the drug had no effect. There is a single nontrivial equilibrium, which gives a viral load of

$$\overline{V} = \frac{\lambda N_T}{c} - \frac{d}{k(1-p\varepsilon)}, \tag{3.13}$$

which differs from equation (2.3) by a factor of p multiplied by the drug efficacy. Hence the relationship between drug efficacy and steady state viral load remains unchanged. In the following section we will discuss a model which groups cells into those affected by the drug and those not, with a slightly more favorable result.

3.2. ***Models with linear relationships between drug efficacy and steady state viral load.*** Though we showed in Section 2 that models with a linear relationship between drug efficacy and steady state viral load are not sufficient to describe the effects of drug therapy, such models still bring us a step closer to robustly modeling the effects of drug therapy. We discuss several models of this nature in the present section: one that describes the effect of RT inhibitors more explicitly than equation (2.1), one that characterizes the effect of cells such as cytotoxic T lymphocytes (CTLs) on HIV-1 infected cells, and another that includes a term for virus induced killing of uninfected cells.

3.2.1. *Reverse transcriptase inhibitors.* RT inhibitors prevent HIV-1 from infecting $CD4^+$ cells by hindering the reverse transcription of HIV-1 RNA into DNA. A more precise way to model RT inhibitors than we did in equation (2.1) is to consider the effect of the drug at the cellular level. The model in this section differs from equation (3.12) in that we explicitly group the uninfected cells, rather than the infected cells, into those which do and do not respond to the drug. The effect of the drug is then measured both by the rate at which target cells are transferred to a pool of infection resistant cells, and by the amount the viral infectivity is reduced in the pool of infection resistant cells. A stronger drug increases the rate of transfer to the pool of cells affected by the drug, decreases infectivity in cells affected by the drug, or both.

Nucleoside analog RT inhibitors, such as AZT, need to be phosphorylated within a cell before they become active. Thus, for concreteness, the rate of transfer between the two uninfected T cell populations can be thought of as the rate of phosphorylation of the RT inhibitor within the cell. Let T denote $CD4^+$ cells that remain susceptible to infection and T_d denote cells containing active drug with susceptibility to infection reduced by the factor $1-\varepsilon$. Then we may write:

$$\frac{dT}{dt} = \lambda - dT - kVT - rT \tag{3.14a}$$

$$\frac{dT_d}{dt} = rT - dT_d - (1-\varepsilon)kVT_d \tag{3.14b}$$

$$\frac{dT^*}{dt} = kV(T + (1-\varepsilon)T_d) - \delta T^* \tag{3.14c}$$

$$\frac{dV}{dt} = N_T \delta T^* - cV \tag{3.14d}$$

where r is the rate at which cells are transferred to the infection resistant population. If r is large enough it will outweigh the production of susceptible cells and leave the virus with a small number of target cells.

The expression for steady state viral load is too cumbersome to display here. However, we found that its dependence on ε is inverse as in the models of the

previous section. In order to determine how r affects the viral load, we set $\varepsilon = 1$, and find a single nontrivial equilibrium; the viral load is

$$\overline{V} = \frac{\lambda N_T}{c} - \frac{d+r}{k}. \tag{3.15}$$

As in the basic model, the viral infectiousness k appears in the denominator of the second term in equation (3.15). However, in this case the analog of drug efficacy, r, appears in the numerator, and this sets up a linear relationship between the strength of the drug and steady state viral load.

This linear relationship is more appropriate for modeling the 1–2 log drops in viral load observed during RT inhibitor monotherapy, because the steady state viral load changes less appreciably near the critical efficacy (Bonhoeffer *et al.*, 1997). Still, according to the argument in Section 2 a linear relationship is not good enough, particularly for the purposes of modeling HAART.

3.2.2. *Cell mediated immunity.* Cytotoxic T lymphocytes (CTLs) are T cells which are capable of recognizing and killing cells infected with HIV, and are usually not susceptible to infection, since they generally lack the $CD4^+$ receptor. A way to generate a nearly linear relationship is by modeling the effect of the CTL population on HIV-1 reproduction (Bonhoeffer *et al.*, 1997). Consider the following model:

$$\frac{dT}{dt} = \lambda - dT - (1-\varepsilon)kVT \tag{3.16a}$$

$$\frac{dT^*}{dt} = (1-\varepsilon)kVT - \delta T^* - mET^* \tag{3.16b}$$

$$\frac{dV}{dt} = N_T \delta T^* - cV \tag{3.16c}$$

$$\frac{dE}{dt} = \rho T^* - d_E E. \tag{3.16d}$$

Here, E represents the effector population of CTLs and m determines the rate of killing of productively infected cells. Effectors are generated in the presence of infected cells at rate ρT^*, and die at rate d_E per cell. As we will show in Section 3.3, the form of the effector generation term is crucial to the relationship between steady state viral load and drug efficacy.

As with the preceding models there is only one nontrivial fixed point. The steady state viral load is

$$\overline{V} = \frac{d}{2(1-\varepsilon)k}\left[\sqrt{(1-R_0(1-\varepsilon)\tau)^2 + 4R_0^2(1-\varepsilon)^2\tau} - (R_0(1-\varepsilon)\tau + 1)\right] \tag{3.17}$$

where $\tau = \delta^2 d_E/\lambda m\rho$ and $R_0 = \lambda N_T k/dc$ as before. We have previously used $\lambda = 10^4$ cells ml^{-1} day^{-1}, and if the rate constant governing effector cell production

is approximately equal to the rate of effector removal, then $\tau \approx \delta^2/m \times 10^{-4}$. Provided $\tau \ll 1$, one can expand equation (3.17) about $\tau = 0$. The leading order term is:

$$\overline{V} \approx \frac{R_0 d\tau(R_0(1-\varepsilon)-1)}{k} \tag{3.18}$$

which is linear in ε.

If the uninfected population is approximated to be constant, the relationship becomes precisely linear (Bonhoeffer *et al*., 1997):

$$\overline{V} = \frac{N\delta d_E}{cm\rho}\left(\frac{(1-\varepsilon)k\overline{T}N_T\delta}{c} - \delta\right). \tag{3.19}$$

Such linearity suggests that cell mediated killing does not play a role in the maintenance of low steady state viral load. However, as we will show in Section 4, if the death rate of infected cells is density dependent (a plausible effect of the immune response on the removal of infected cells) then robust low steady state viral loads are obtainable.

3.2.3. *Virus induced killing of uninfected cells.* Uninfected cells in the vicinity of infected cells are subject to killing by the formation of syncytia and interaction with gp120 shed by infected cells (Lifson *et al*., 1986; Sodroski *et al*., 1986; Yoffe *et al*., 1987). We can model these phenomena by appending a death term to the differential equation for target cells in the basic model, as follows (Bonhoeffer *et al*., 1997):

$$\frac{dT}{dt} = \lambda - dT - (1-\varepsilon)kVT - qT^*T \tag{3.20a}$$

$$\frac{dT^*}{dt} = (1-\varepsilon)kVT - \delta T^* \tag{3.20b}$$

$$\frac{dV}{dt} = N_T\delta T^* - cV. \tag{3.20c}$$

There is one nontrivial equilibrium, and the corresponding viral load is

$$\overline{V} = \frac{(1-\varepsilon)\lambda k N_T - dc}{((1-\varepsilon)k + \frac{qc}{N_T\delta})c}, \tag{3.21}$$

which, for $q \gg k$, is approximately linear in ε. Note that models for which the rate of killing of uninfected cells is proportional to V rather than T^* will have similar behavior because at steady state V is proportional to T^*, in particular $\overline{V} = N_T\delta\overline{T^*}/c$.

3.3. ***Models with constant steady state viral load.*** Here we will briefly mention a class of models for which drug efficacy does not have *any* effect on the steady state viral load. Consider the following model (Nowak and Bangham, 1996):

$$\frac{dT}{dt} = \lambda - dT - (1-\varepsilon)kVT \tag{3.22a}$$

$$\frac{dT^*}{dt} = (1-\varepsilon)kVT - \delta T^* - mET^* \tag{3.22b}$$

$$\frac{dV}{dt} = N_T \delta T^* - cV \tag{3.22c}$$

$$\frac{dE}{dt} = \rho T^* E - d_E E. \tag{3.22d}$$

This model differs from equations (3.16a) to (3.16d) only in the form of the growth term of the effector cells. Here the effector cells grow at a rate that depends upon the pre-existence of other effector cells and infected cells. This form is typical of predator–prey dynamics, and is motivated by the notion that precursor CTLs encounter infected cells and subsequently proliferate into mature effectors. Hence E represents both the precursor and mature effector cell populations.

There are two nontrivial equilibria for the model; this is the only system in this paper with more than one infected steady state. For the first, $E = 0$, and the viral load in this case is given by equation (2.3), corresponding to the basic model.

The second fixed point occurs when effector cells are present, i.e., when $E \neq 0$. In this case the steady state viral load is

$$\overline{V} = \frac{N\delta d_E}{c\rho}, \tag{3.23}$$

which is independent of drug efficacy, or viral infectiousness.

Figure 6 shows that for low drug efficacies, the $E \neq 0$ equilibrium is stable. As ε increases, steady state viral load remains constant until the $E = 0$ equilibrium becomes stable. As in the previous cases, ultimately the viral load goes to zero.

4. Using a Density Dependent Infected Cell Death Rate Can Lead to a Robust Low Steady State Viral Load

In most of the models discussed here, the death rate of infected cells, δ, was assumed to be constant. An exception to this occurred in Sections 3.2.2 and 3.3, in which the infected cells were cleared at a rate which depended on the density of effector cells. We can also model this by assuming that the size of the effector cell population is a function of the density of infected cells, and represent its effect by choosing the infected cell death rate to be a function of infected cell density.

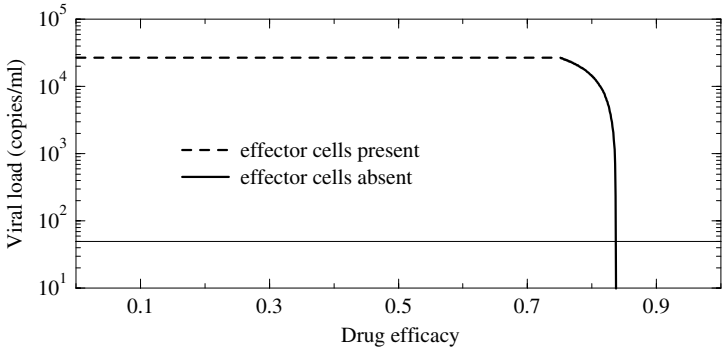

Figure 6. Steady state viral load vs drug efficacy for a model that includes a cell meditated immune response [equation (3.22)]. For low drug efficacies, $E \neq 0$ and the viral load is constant (dashed curve). However, as efficacy increases the steady state that gives the viral load in equation (3.23) loses stability; the viral load is then determined by equation (2.3) (solid curve), and $E = 0$. Parameters used in this figure are defined in the Appendix, except $d_e = 0.05$ day^{-1}, and $\rho = m = 10^{-5}$ ml day^{-1} [taken from Nowak and Bangham (1996)]. The value of ε at which the stability transition occurs increases with ρ, such that for these parameters, when $\rho > 5 \times 10^{-5}$, the $E \neq 0$ equilibrium is stable until viral extinction occurs.

As suggested by Holte *et al*. (2001), one of the simplest ways to do this is via a power law, and consequently, here we will replace δ in previous models by $\delta(T^*) = \delta' T^{*\omega}$:

$$\frac{dT}{dt} = \lambda - dT - (1-\varepsilon)kVT \tag{4.1a}$$

$$\frac{dT^*}{dt} = (1-\varepsilon)kVT - \delta' T^{*\omega} T^* \tag{4.1b}$$

$$\frac{dV}{dt} = pT^* - cV. \tag{4.1c}$$

In previous models, the rate of production of free virus from infected cells was $N_T\delta$, the average number of viral particles produced per infected cell times the natural death rate of infected cells. In addition to viral bursting, in this case $\delta(T^*)$ accounts for death of infected cells due to immune system clearance, meaning infected cells could be cleared prior to producing virus. Consequently $N_T\delta(T^*) = N_T\delta' T^{*\omega}$ overestimates the rate of production of free virus, and we use the parameter p instead.

There is a single nontrivial equilibrium. The associated steady state viral load is

$$\overline{V} = \frac{p}{c}\mathrm{e}^{\gamma}, \tag{4.2}$$

where γ satisfies

$$\omega\gamma + \ln\frac{\lambda kp(1-\varepsilon)}{\delta(cd + kp(1-\varepsilon)e^{\gamma})} = 0. \tag{4.3}$$

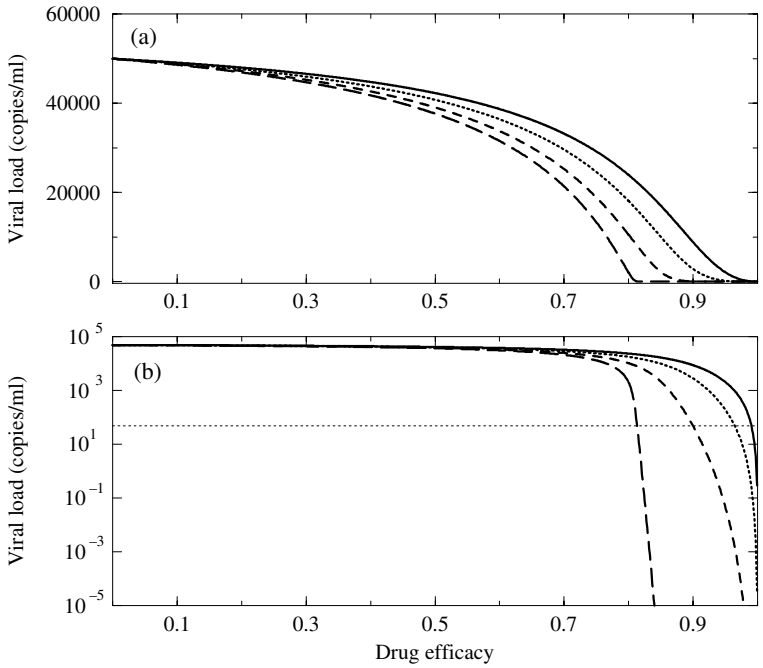

Figure 7. (a) Virus load vs drug efficacy for density dependent infected cell death [equation (4.1)]. Parameters are defined in the Appendix, and $\omega = 0.44$, $\delta' = 0.0155$ day^{-1}(ml cell^{-1})$^{\omega}$ (solid curve), $\omega = 0.25$, $\delta' = 0.0878$ day^{-1}(ml cell^{-1})$^{\omega}$ (dotted curve) $\omega = 0.1$, $\delta' = 0.3455$ day^{-1}(ml cell^{-1})$^{\omega}$ (dashed curve) and $\omega = 0.01$, $\delta' = 0.7863$ day^{-1}(ml cell^{-1})$^{\omega}$ (long dashed curve). The thin line delineates 50 copies ml^{-1}, the detection threshold. As ω decreases, the range of efficacies over which the virus persists below the detection threshold increases. (b) The same figure on a semilog scale, demonstrating that steady state viral load persists below 50 copies ml^{-1} for a broad range of efficacies.

If we choose $p = N_T\delta = 70$ day^{-1} (where δ is the constant death rate parameter used in previous models), $\omega = 0.44$ (Holte *et al.*, 2001), and keep the remainder of the parameters as defined in the Appendix, then δ' remains as the only unknown. If we use the constraint that $\overline{V}(\varepsilon = 0) = 5 \times 10^4$ copies ml^{-1}, then we find $\delta' = 0.0155$ day^{-1}(ml cell^{-1})$^{\omega}$. The upper curve in Fig. 7 shows how the steady state viral load varies with drug efficacy for these parameter choices. Figure 7(a) illustrates that the curve becomes concave up as the steady state viral load approaches zero, suggesting it is much less sensitive to changes in efficacy for very low viral loads than any of the other models we have discussed so far. However, in order for the steady state viral load to decline below 50 copies ml^{-1} for this combination of parameters, the efficacy must be $\varepsilon \geq 0.991$, which is an unrealistically high value (Louie *et al.*, 2001).

As we decrease ω toward zero, and continue choosing δ' via the constraint $\overline{V}(\varepsilon = 0) = 5 \times 10^4$ copies ml^{-1}, we find that the curve begins to appear more reasonable. In particular, the steady state viral load is substantially reduced for drug efficacies of 0.85 or even lower. Numerical calculations reveal that for the parameters used in this section, the infected steady state is positive and stable for all $\varepsilon < 1$. This

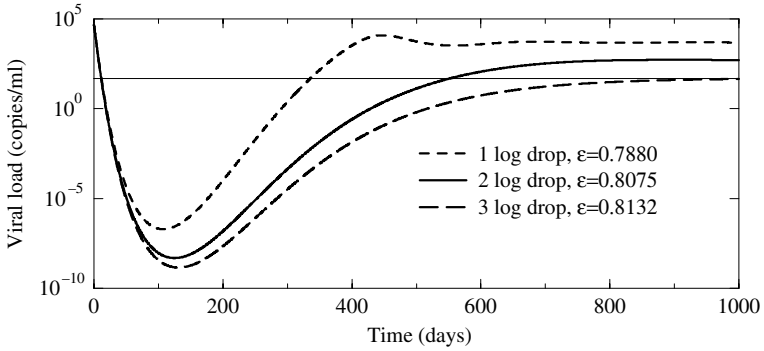

Figure 8. Decay in virus load predicted by equation (4.1) for different drug efficacies. The time required for convergence to low steady states is less than 1000 days, which is much faster than in any of the previous models. Parameters are defined in the Appendix, with $\omega = 0.01$ and $\delta' = 0.7863$ day^{-1}(ml cell^{-1})$^{\omega}$. The thin line delineates 50 copies ml^{-1}, the detection threshold.

indicates that unless the antiretroviral drug combination is 100% effective, viral extinction is not possible. Note, however, that we have exchanged sensitivity to drug efficacy with sensitivity to the size of ω as the steady state viral load becomes very low. For example, when $\omega = 0.01$, $\overline{V}(\varepsilon = 0.8132) = 50$ copies ml^{-1}. However, if we fix the drug efficacy and increase ω to 0.02, the steady state viral load is $\overline{V}(\varepsilon = 0.8132) = 864$ copies ml^{-1}, a 17-fold increase in the viral load.

Parametrization of the model from an empirical study in which it was assumed $\varepsilon = 1$ (Holte *et al.*, 2001) suggests that $0.4 < \omega < 0.47$, yet we find the model behaves most reasonably by our criteria when $\omega = 0.01$. Relaxing the assumption $\varepsilon = 1$ could result in an estimate of ω that is closer to what we find reasonable here. However, because a reduction in efficacy will reduce the model's predicted rate of virus decay, to fit the model to data one would expect that the exponent ω would need to increase (thus increasing the rate of infected cell removal) to compensate for a reduction in ε. Thus, allowing $\varepsilon < 1$ should not lead to the low values of ω that would make this model serve as a robust descriptor of low viral loads. However, only one data set has been analyzed with this model, and estimates of ω from other data sets need to be done before a more informed decision about the relevance of this model can be made.

Despite these complicating factors, the dynamics of drug therapy simulation are substantially improved from previously discussed models. In Fig. 8 we show the simulation of therapy for different values of ω and δ' and the drug efficacy which corresponds to a steady state of 50 copies ml^{-1} for each parameter combination. Even for very low steady state viral loads, the time to convergence, though still long, is much shorter than the times seen previously (see, for example, results in Figs 2, 4 and 5).

5. COMPARTMENTAL MODELS WITH A ROBUST LOW STEADY STATE VIRAL LOAD

In Sections 3.1.7 and 3.2.1 we introduced models in which infected cells were distinguished according to whether or not they were affected by the drug. This type of model is justified for two reasons. First, there is evidence that drug concentrations, and therefore efficacies, are reduced in certain physiologically distinct sites in the body such as the testes and the brain (Lewis *et al*., 1996; Haworth *et al*., 1998; Schlegel and Chang, 1998). Second, *in vitro* studies have demonstrated that there can be heterogeneities in intracellular drug concentrations. For example, monocyte cell lines are less susceptible to the effects of antiretroviral drugs (Kim *et al*., 1998; Perno *et al*., 1998; Puddu *et al*., 1999). Indeed, these mechanisms have been suggested as possible explanations for the residual replication of HIV-1 that persists in HAART patients (Zhang *et al*., 1999). For the models in Sections 3.1.7 and 3.2.1, near extinction the steady state viral load was extremely sensitive to small changes in drug efficacy. Here we will show that if we model two cocirculating populations of target cells with differing drug efficacies, or two physiologically distinct compartments, one of which is a drug sanctuary such as the brain or testes, a less sensitive relationship between drug effect and steady state viral load results.

As we showed in Section 2, if we simulate therapy using the basic model, the viral load becomes unreasonably low as it oscillates about the steady state. This problem was alleviated by including a subpopulation of chronically infected cells (see Section 3.1.4), and consequently the following two models include this subpopulation as well.

5.1. ***Drug sanctuary created by a physiological barrier.*** We extend the chronically infected cell model in Section 3.1.4 by considering the infection process to occur in two distinct compartments, with one regarded as a drug sanctuary. These compartments are then coupled by allowing transport of virus between the compartments. The extended model is given by the following equations:

$$\frac{dT_1}{dt} = \lambda - dT_1 - (1-\varepsilon)kV_1T_1 \tag{5.1a}$$

$$\frac{dT_2}{dt} = \lambda - dT_2 - (1-f\varepsilon)kV_2T_2 \tag{5.1b}$$

$$\frac{dT_1^*}{dt} = (1-\alpha)(1-\varepsilon)kV_1T_1 - \delta T_1^* \tag{5.1c}$$

$$\frac{dT_2^*}{dt} = (1-\alpha)(1-f\varepsilon)kV_2T_2 - \delta T_2^* \tag{5.1d}$$

$$\frac{dC_1^*}{dt} = \alpha(1-\varepsilon)kV_1T_1 - \mu C_1^* \tag{5.1e}$$

$$\frac{dC_2^*}{dt} = \alpha(1-f\varepsilon)kV_2T_2 - \mu C_2^* \tag{5.1f}$$

$$\frac{dV_1}{dt} = N_T^*\delta T_1^* + N_C\mu C_1^* - cV_1 + D_1(V_2 - V_1) \tag{5.1g}$$

$$\frac{dV_2}{dt} = N_T^*\delta T_2^* + N_C\mu C_2^* - cV_2 + D_2(V_1 - V_2) \tag{5.1h}$$

where T_i, T_i^*, C_i^*, and V_i represent the concentration of HIV-1 target cells, short-lived infected cells, long lived chronically infected cells, and free HIV-1 RNA, respectively, where $i = 1$ in the main compartment and $i = 2$ in the drug sanctuary. We have not explicitly modeled transport of the drug from the main compartment to the sanctuary, rather, we define the parameter f as the factor by which the drug efficacy is reduced in the sanctuary. The transport of virus between the main compartment and the sanctuary is governed by the rate constants D_1 and D_2 and the difference in virus concentration between the two compartments. Definitions for all other parameters follow from the previous models. Notice that apart from the different transport rates and responsiveness to drug therapy, we have assumed that the two compartments are identical by using the same parameters in each. This allows us to test the effect of compartmentalization alone without added complexities. The model has been previously used to study the role drug concentration heterogeneity plays on the generation of drug resistant mutants (Kepler and Perelson, 1998).

There is a single nontrivial equilibrium, and though it can be found in closed form, the expression is prohibitively long to be displayed here or, for that matter, analyzed for its dependence on ε. Instead, we have plotted the steady state viral load as a function of drug efficacy, taking care to choose the parameters realistically (see Appendix). Figure 9 compares the steady state viral load vs main compartment efficacy curve in the spatial two compartment model with the $\overline{V}$–ε curve for the basic one compartment model, equation (3.6), where the same parameters are used in the main compartment. Due to transport from the drug sanctuary to the main compartment, the virus in the main compartment is not eradicated, despite 100% efficacy there. In fact, the curve is concave up near the point of critical efficacy in the one compartment model ($\varepsilon_0 = 0.8$), indicating that the steady state viral load will not be sensitive to small changes in drug efficacy.

As discussed in Section 2.2, in addition to the steady state the system ultimately reaches, it is important to consider the dynamics that lead to it. Figure 10 shows the approach to steady state following initiation of therapy for several different parameter choices. Figure 10(a) compares the trajectories for the one compartment model and the drug sanctuary model; in both cases the drug efficacy was chosen such that the steady state would be 50 copies ml^{-1}. Many years are required to reach the steady state with the one compartment model, and hence only part of

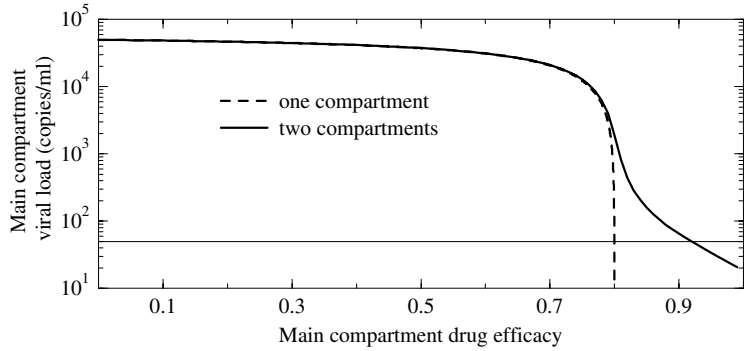

Figure 9. Relationship between steady state virus load and main compartment drug efficacy for one compartment model in equation (2.1) (dashed curve) and the two compartment model in equation (5.1) (solid line). In the drug sanctuary the drug efficacy is reduced by a factor, $f = 0.45$; the remainder of the parameters are defined in the Appendix.

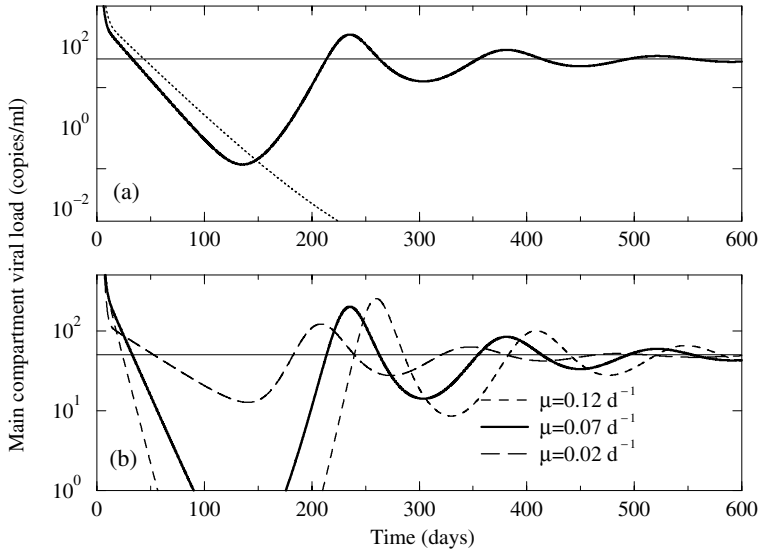

Figure 10. Simulation of therapy with equation (5.1), beginning at $t = 0$ days. The dotted curve delineates 50 copies ml^{-1}, the limit of detection for standard HIV-1 assays. (a) Comparison between dynamics in the one compartment model (dashed curve) and the sanctuary model (solid curve). Drug efficacy was chosen in each case so that the steady state viral load was 50 copies ml^{-1} ($\varepsilon = 0.79984$ in the one compartment model; $\varepsilon = 1.0$ and $f = 0.34$ in the drug sanctuary model). The viral load in the one-compartment model ultimately recovers and reaches 50 copies ml^{-1} but on a time scale longer than that illustrated. (b) Effect of drug therapy in the sanctuary model for three choices of the chronically infected cell decay rate, μ. Parameters are defined in the Appendix, with $\varepsilon = 1.0$ and $f = 0.34$. In this case the model predicts that slower decay rates correlate with smaller though more frequent blips.

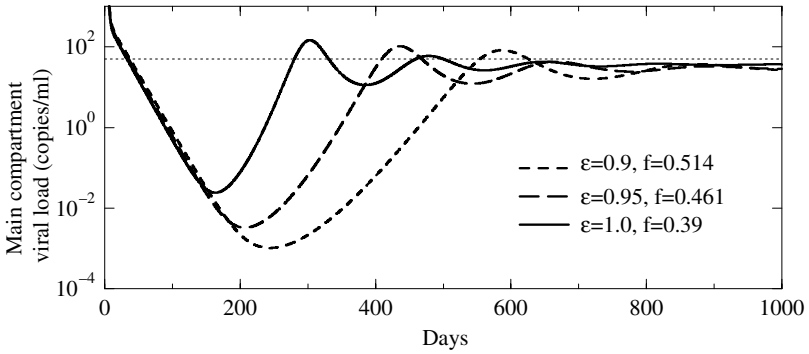

Figure 11. Reducing main compartment drug efficacy while fixing steady state viral load reduces the blip size predicted by equation (5.1). In this case, f was chosen so that the steady state viral load would be 30 copies ml^{-1}. The remaining parameters are defined in the Appendix.

the curve is shown. On the other hand, if the source of virus during therapy is the drug sanctuary, we find that the viral load settles to steady state more rapidly, and oscillates in the process. Since for the majority of the time the viral load remains below the standard threshold of detection, the brief moments while the viral load is above the threshold could serve as a possible explanation for intermittent episodes of detectable viremia in patients whose viral load is otherwise well suppressed (Callaway and Perelson, 2002). These 'blips' in viral load have previously been attributed to factors such as noncompliance on the part of the patient to the drug regimen and activation of latently infected cells. While the results here do not rule out these possibilities, they provide an additional mechanism to account for the observation. Figure 10(b) demonstrates that this oscillatory behavior remains for different parameter choices, indicating that the model behavior is robust to changes in parameters. Furthermore the blip size is typically less than 200 copies ml^{-1}, consistent with findings in an analysis of 124 patient records from several studies (DiMascio, Louie, Ho and Perelson, unpublished observation).

A few interesting predictions follow from the dynamics in Fig. 10 (Callaway and Perelson, 2002). First, since the oscillations decay in time, blips are more likely to occur early after treatment starts. Second, as the initial viral decay rate increases, the peaks of the oscillations are larger; hence blips are more likely to be observed in patients with faster initial viral decay rates.

Figure 11 demonstrates the effect of reducing the main compartment efficacy and simultaneously changing f, the parameter governing drug penetrance in the sanctuary, such that the steady state viral load was 30 copies ml^{-1}. Not surprisingly, blips are much more prominent when the sanctuary penetrance is low. The same conclusion holds if we increase f without simultaneously reducing ε.

The transport rate constants D_1 and D_2 can be thought of as measures of the relative size of the two compartments. In particular, if we define u as the fraction

of the sum of the volumes of two compartments which is occupied by the drug sanctuary, then we have (Kepler and Perelson, 1998)

$$D_1 = \frac{u D_2}{1-u}, \tag{5.2}$$

which means, for the parameters in the Appendix, $u = 0.0053$, i.e., the volume of the drug sanctuary is approximately 0.5% of the total volume.

5.2. ***Differential efficacy in cocirculating target cells.*** Drug efficacy may vary by cell type, not just physiological location. Hence another model which accounts for heterogeneous drug responsiveness is one in which two types of target cells cocirculate in a single compartment, where in one population ($i = 1$) the drug efficacy is ε, while in the second ($i = 2$) the drug efficacy $f\varepsilon$ is reduced by a factor $f < 1$. In this case, the equations are similar to equation (5.1), with the exception that we need only consider one virus population since there is only one spatial compartment:

$$\frac{dT_1}{dt} = \lambda_1 - d_1 T_1 - (1-\varepsilon) k_1 V T_1 \tag{5.3a}$$

$$\frac{dT_2}{dt} = \lambda_2 - d_2 T_2 - (1 - f\varepsilon) k_2 V T_2 \tag{5.3b}$$

$$\frac{dT_1^*}{dt} = (1-\alpha)(1-\varepsilon) k_1 V T_1 - \delta T_1^* \tag{5.3c}$$

$$\frac{dT_2^*}{dt} = (1-\alpha)(1 - f\varepsilon) k_2 V T_2 - \delta T_2^* \tag{5.3d}$$

$$\frac{dC_1^*}{dt} = \alpha(1-\varepsilon) k_1 V T_1 - \mu C_1^* \tag{5.3e}$$

$$\frac{dC_2^*}{dt} = \alpha(1 - f\varepsilon) k_2 V T_2 - \mu C_2^* \tag{5.3f}$$

$$\frac{dV}{dt} = N_T \delta (T_1^* + T_2^*) + N_C \mu (C_1^* + C_2^*) - cV. \tag{5.3g}$$

The definitions of all the variables and parameters are identical to those in equation (5.1). We have, however, used subscripts on several target cell dependent parameters, in particular the infection rate, k_i, target cell death rate constant, d_i, and target cell production rate λ_i. In principle, the infected cell death rate constants, δ and μ might also depend on the target cell type, but we have not introduced that generalization here. As in the previous model, there is a single nontrivial equilibrium, yet the expression for viral load is too cumbersome to display; instead we

show the relationship between drug efficacy and steady state viral load in Fig. 12. The main population parameters are defined in the Appendix, and we chose λ_2, k_2 and d_2 subject to the following constraints: (1) when $f = 0$ and $\varepsilon = 1$ (drug efficacy = 0 in population 2 and efficacy = 1 in population 1), $\overline{V} = 100$ copies ml^{-1}, (2) when $f = 0.5$ and $\varepsilon = 1$, $\overline{V} = 0$ copies ml^{-1}, where the steady state viral load when $\varepsilon = 1$ is given by

$$\overline{V}(\varepsilon = 1) = \frac{(1-\alpha)\lambda_2 N_T + \alpha\lambda_2 N_C}{c} - \frac{d_2}{(1-f)k_2}. \tag{5.4}$$

This is the same constraint used in the Appendix, and is motivated by needing the second population to contribute a small amount to the steady state viral load and do so over a range of drug efficacies. Using these constraints, we find that $\lambda_2 = 31.98$ cells ml^{-1} day^{-1}, and $d_2/k_2 = 100$ copies ml^{-1}. In the main population we have $d_1/k_1 = 12\,500$ copies ml^{-1}, indicating that we must choose d_2 and k_2 significantly different from the parameters for the main population. If we take $d_2 = d_1 = 0.01$ day^{-1}, then $k_2 = 10^{-4}$ ml copies^{-1} day^{-1}, more than two orders of magnitude different from $k_1 = 8 \times 10^{-7}$ ml copies^{-1} day^{-1}. Why the infection rate would be so much higher in the second population could have to do with activation requirements. Most CD4$^+$ T cells are not permissive to infection unless they are in an activated state (Stevenson, 1996). However, for some cell types, such as macrophages, it is believed that there is no activation requirement (Stevenson and Gendelman, 1994), and hence the infection rate constant for such cells would not need to account for the small fraction of cells that are activated and permissive to infection (see Section 3.1.3). Since approximately 1–3% of CD4$^+$ T cells are activated enough to be in cell cycle (Sachsenberg *et al*., 1998), the infection rate of macrophages could be as much as 100-fold higher if their were no activation requirement. Protease inhibitors have lower efficacy in chronically infected macrophages than in chronically infected lymphocytes (Kim *et al*., 1998; Perno *et al*., 1998; Puddu *et al*., 1999), and hence these cells are a possible candidate for the second population of target cells. Recent experiments in the macaque in fact suggest that macrophages may be an important source of virus after CD4$^+$ T cells are depleted (Igarashi *et al*., 2001).

As indicated in Fig. 12, the relationship between steady state viral load and drug efficacy for this system is similar to the one in equation (5.1), providing additional support to the notion that low steady state viral loads are supported only by a small subpopulation of cells in which for some reason antiretroviral drug efficacy is reduced. Furthermore, as we show in Fig. 13, there is a similar oscillatory behavior, providing additional evidence to suggest that blips in viral load could be a natural consequence of population dynamics. Because k_2 is of questionable size, we examined the effect of reducing k_2 on the system's dynamics. Just as increasing f in the model defined by equation (5.1) reduces the blip size, so too does reducing k_2 in this model (results not shown).

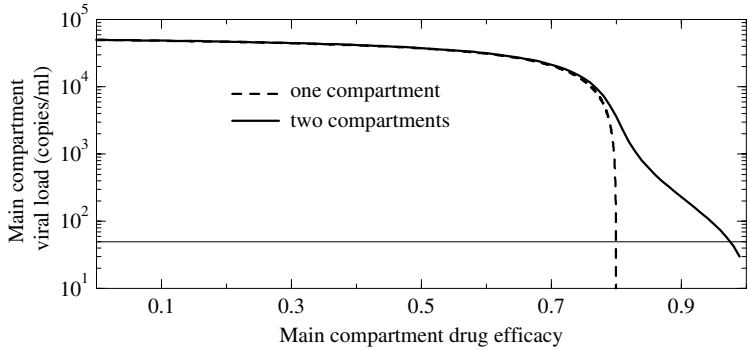

Figure 12. Relationship between steady state virus load and main compartment drug efficacy for equation (5.3). All parameters except those with a subscript 2 are defined in the Appendix, and $\lambda_2 = 31.98$, $d_2 = 0.01$, $k_2 = 10^{-4}$, and $f = 0.34$.

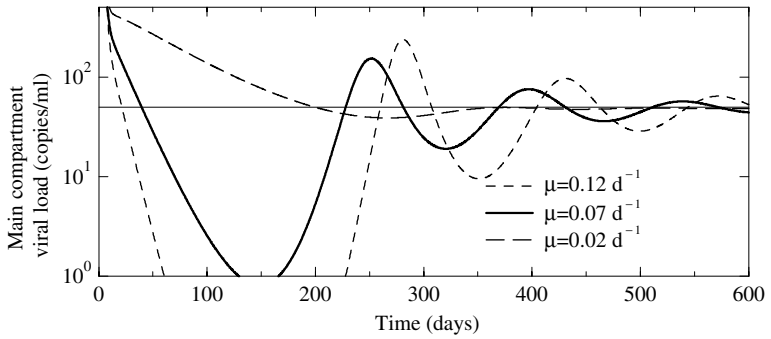

Figure 13. Simulation of therapy, beginning at $t = 0$ days, for equation (5.3). The dotted curve delineates 50 copies ml^{-1}, the limit of detection for standard HIV-1 assays. We have varied the chronically infected cell decay rate, μ. Parameters are defined in the Appendix, with $\varepsilon = 1.0$ and $f = 0.34$.

6. SUMMARY AND CONCLUSIONS

In order to faithfully describe the effects of antiretroviral therapy, models of HIV-1 dynamics should be capable of simulating persistent, low level replication, and extinction of the virus should be unlikely. We have explored the steady state behavior of several models of HIV-1 infection to determine which are most consistent with these features of drug therapy. For most models, the presence of low level replication is extremely sensitive to small changes in drug efficacy. Consequently, for very low viral loads, the probability of extinction is high if the efficacy were to increase by a very small amount. Models which fail in this manner to robustly describe drug therapy include not only the basic model in equation (2.1), but also variants that account for cell mediated killing of infected cells, virus induced killing

of uninfected cells, latently infected cells, quiescent target cells, the specific mechanisms of antiretroviral drug efficacy, and models in which we varied the form of the terms for growth and infection of target cells. This suggests that, although these mechanisms may be important for other aspects of HIV-1 replication, they are not crucial in the maintenance of low viral loads.

Two classes of models do not exhibit extreme sensitivity of the steady state viral load to changes in drug efficacy. In the first of these, the death rate of infected cells is a function of their density. This form is motivated by the notion that infected cells could be cleared by the immune response at a rate proportional to the frequency of infected cells. Although we found that ω, the parameter that governs the size of this effect on the death rate, must be much lower than experimentally estimated values in order to generate this robust relationship, we did find that the dynamics of therapy were much more realistic than in the other variants of the basic model that we surveyed. This surprisingly suggests the immune system may have a role in the persistence of HIV replication.

Models in which target cells respond differentially to antiretroviral drugs are also capable of modeling persistent low steady state viral loads robustly. This points in the direction of improving drug efficacy in all cell types in order to improve the chances of ultimately eradicating HIV-1 from the human body. This could mean developing drugs that cross physiological barriers better, drugs that inhibit P-glycoprotein pumps, or drugs which, unlike protease inhibitors, are not easily removed from the cell interior by efflux pumps (Kim *et al*., 1998; Perno *et al*., 1998; Puddu *et al*., 1999).

Examining the source of virus during blips could provide a means to support the suggestion made here of differing target cell populations having different drug susceptibilities. If a distinct cell type, with distinct cell surface markers, is the source of residual replication and intermittent viremia, those surface markers will be present on the viral membrane, since it is derived from the membrane of the virus producing cell.

Two compartment models hold promise not only for identifying the population of drug resistant cells, but also for timing optimal shifts in therapeutic regimens. If the 'blips' observed in patient data and those in the dynamics of this model are in fact one and the same, the model could be used to predict the timing of blips. This is significant since blips indicate high levels of virus replication, and it is at these times that drug resistant mutants are most likely to appear. This could guide a clinician's decision to change the drugs used in HAART for a particular patient.

7. APPENDIX

Several parameters were taken directly from the literature. In particular, the death rate of infected $CD4^+$ cells, $\delta = 0.7\ \text{day}^{-1}$ (Perelson *et al*., 1996), the clearance rate of free virus particles, $c = 13\ \text{day}^{-1}$ (Ferguson *et al*., 1999; Mittler *et al*., 1999;

Ramratnam *et al*., 1999), and the death rate of chronically infected cells, when included in a model, was $\mu = 0.07$ day^{-1} (Perelson *et al*., 1997). The death rate of CD4$^+$ in humans is not well characterized, and we took $d = 0.01$ day^{-1}, a value derived from BrdU labeling in macaque monkeys (Mohri *et al*., 1998). Quiescent cells are known to have an extremely long half-life (McLean and Michie, 1995), and we used $d_Q = 0.001$ day^{-1}, corresponding to a half-life of approximately 2 years. The average number of virus particles produced during an infected cell life span was taken as $N_T = 100$ (Haase *et al*., 1996).

In order to determine the remainder of the parameters, we used steady state expressions from the two compartment drug sanctuary model in equation (5.1) and the following constraints: (1) there are 1000 target cells per microliter in both compartments in the absence of infection, (2) the steady state viral load in both compartments in the absence of therapy is 5×10^4 copies ml^{-1}, (3) the density of target cells in either compartment during infection in the absence of therapy is 2×10^5 cells ml^{-1}, (4) chronically infected cells contribute to at least 1% of the total steady state viral load (Perelson *et al*., 1997), (5) in the absence of replication in the main compartment ($\varepsilon = 1$), if no drug enters the second compartment ($f = 0$), the main compartment steady state viral load is 100 copies ml^{-1}, and (6) if the drug is 50% effective in the second compartment and 100% effective in the main compartment ($f = 0.5$ and $\varepsilon = 1$), the viral load is zero everywhere. The purpose of constraints (5) and (6) is to ensure that the steady state viral load persists over a broad range of second compartment efficacies, despite complete drug efficacy in the main compartment. Using these constraints and expressions for the steady state viral load, we found that $\lambda = 10^4$ cells ml^{-1} day^{-1}, $k = 8 \times 10^{-7}$ ml copy^{-1} day^{-1}, $\alpha = 0.195$, $N_C = 4.11$, $D_1 = 0.1048$ day^{-1}, and $D_2 = 19.66$ day^{-1}.

Without intercompartment transport, for these parameters the critical efficacy in each compartment would be $\varepsilon = 0.8$. However, as we have chosen the parameters, the transport term in the second compartment acts to increase the viral clearance term from cV_2 to $(c + D_2)V_2$ (when $V_1 \ll V_2$), meaning that extinction can occur in the second compartment even when $f\varepsilon < 0.8$. This is why $\varepsilon = 1$ and $f = 0.5$ is sufficient to extinguish the virus in both compartments. In general, extinction will be determined by a 'critical curve' in the ε–f plane, connecting the points $(\varepsilon, f) = (1.0, 0.5)$ and $(0.8, 1.0)$. We have been unable to find this curve analytically.

For consistency we chose to use the same parameters with other models, where applicable. This resulted in small changes in some of the steady state values. For example, in the basic one compartment model in equation (2.1), the drug free steady state viral load is $\overline{V} = 6.44 \times 10^4$, rather than 5×10^4 as in models with chronically infected cells.

For the latently infected cell model, we chose α, the fraction of infection events which result in a latently infected cell, by using the constraint that there are approximately five latently infected cells per million resting CD4$^+$ cells (Chun *et al*., 1997) at steady state. Using the parameters above with $\delta_L = 0.004$ day^{-1} (Finzi *et al*., 1999) and $a = 0.01$ day^{-1}, this gives $\alpha = 1.5 \times 10^{-6}$.

ACKNOWLEDGEMENTS

Portions of this work were done under the auspices of the U.S. Department of Energy. The work was supported by NIH grant RR06555 and a National Science Foundation Graduate Research Fellowship (DGE96176048).

REFERENCES

Anderson, R. M. and R. M. May (1991). *Infectious Diseases of Humans*, Oxford: Oxford University Press.

Bonhoeffer, S., J. M. Coffin and M. A. Nowak (1997). Human immunodeficiency virus drug therapy and virus load. *J. Virol.* **71**, 3275–3278.

Callaway, D. S. and A. S. Perelson (2002). Intermittent viremia in HIV-1 infected patients may occur in patients fully adherent to potent antiretroviral drug therapy due to drug concentration heterogeneity (in review).

Callaway, D. S., R. M. Ribeiro and M. A. Nowak (1999). Phenotype switching and disease progression in HIV-1 infection. *Proc. R. Soc. Lond. B*, **266**, 2523–2530.

Cavert, W. *et al*. (1997). Kinetics of response in lymphoid tissues to antiretroviral therapy of HIV-1 infection. *Science* **276**, 960–964.

Chun, T. W., R. T. Davey Jr, D. Engel, H. C. Lane and A. S. Fauci (1999). Re-emergence of HIV after stopping therapy. *Nature* **401**, 874–875.

Chun, T. W., L. Stuyver, S. B. Mizell, L. A. Ehler, J. A. Mican, M. Baseler, A. L. Lloyd, M. A. Nowak and A. S. Fauci (1997). Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc. Natl. Acad. Sci. USA* **94**, 13193–13197.

De Boer, R. J. and A. S. Perelson (1995). Towards a general function describing T cell proliferation. *J. Theor. Biol.* **175**, 567–576.

De Boer, R. J. and A. S. Perelson (1998). Target cell limited and immune control models of HIV infection: a comparison. *J. Theor. Biol.* **190**, 201–214.

Dornadula, G. *et al*. (1999). Residual HIV-1 RNA in blood plasma of patients taking suppressive highly active antiretroviral therapy. *JAMA* **282**, 1627–1632.

Ferguson, N. M. *et al*. (1999). Antigen-driven CD4+ T cell and HIV-1 dynamics: residual viral replication under highly active antiretroviral therapy. *Proc. Natl. Acad. Sci. USA* **96**, 15167–15172.

Finzi, D. *et al*. (1999). Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat. Med.* **5**, 512–517.

Finzi, D. *et al*. (1997). Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science* **278**, 1295–1300.

Furtado, M. R., D. S. Callaway, J. P. Phair, K. J. Kunstman, J. L. Stanton, C. A. Macken, A. S. Perelson and S. M. Wolinsky (1999). Persistence of HIV-1 transcription in peripheral-blood mononuclear cells in patients receiving potent antiretroviral therapy. *N. Engl. J. Med.* **340**, 1614–1622.

Gartner, S., P. Markovits, D. M. Markovitz, M. H. Kaplan, R. C. Gallo and M. Popovic (1986). The role of mononuclear phagocytes in HTLV-III/LAV infection. *Science* **233**, 215–219.

Haase, A. T. *et al*. (1996). Quantitative image analysis of HIV-1 infection in lymphoid tissue. *Science* **274**, 985–989.

Haworth, S. J., B. Christofalo, R. D. Anderson and L. M. Dunkle (1998). A single-dose study to assess the penetration of stavudine into human cerebrospinal fluid in adults. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* **17**, 235–238.

Hlavacek, W. S., N. I. Stilianakis, D. W. Notermans, S. A. Danner and A. S. Perelson (2000a). Influence of follicular dendritic cells on decay of HIV during antiretroviral therapy. *Proc. Natl. Acad. Sci. USA* **97**, 10966–10971.

Hlavacek, W. S., N. I. Stilianakis and A. S. Perelson (2000b). Influence of follicular dendritic cells on HIV dynamics. *Phil. Trans. R. Soc. Lond. B* **355**, 1051–1058.

Hlavacek, W. S., C. Wofsy and A. S. Perelson (1999). Dissociation of HIV-1 from follicular dendritic cells during HAART: mathematical analysis. *Proc. Natl. Acad. Sci. USA* **96**, 14681–14686.

Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard and M. Markowitz (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**, 123–126.

Holte, S., A. Melvin, J. Mullins and L. Frenkel (2001). Density-dependent decay in HIV dynamics after HAART (abstr 394), in *8th Conference on Retroviruses and Opportunistic Infections*, Alexandria, VA, USA: Foundation for Retrovirology and Human Health.

Igarashi, T., C. R. Brown, Y. Endo, A. Buckler-White, R. Plishka, N. Bischofberger, V. Hirsch and M. A. Martin (2001). Macrophage are the principal reservoir and sustain high virus loads in rhesus macaques after the depletion of CD4+ T cells by a highly pathogenic simian immunodeficiency virus/HIV type 1 chimera (SHIV): implications for HIV-1 infections of humans. *Proc. Natl. Acad. Sci. USA* **98**, 658–663.

Kepler, T. B. and A. S. Perelson (1998). Drug concentration heterogeneity facilitates the evolution of drug resistance. *Proc. Natl. Acad. Sci. USA* **95**, 11514–11519.

Kim, R. B., M. F. Fromm, C. Wandel, B. Leake, A. J. Wood, D. M. Roden and G. R. Wilkinson (1998). The drug transporter P-glycoprotein limits oral absorption and brain entry of HIV-1 protease inhibitors. *J. Clin. Invest.* **101**, 289–294.

Koenig, S., H. E. Gendelman, J. M. Orenstein, M. C. Dal Canto, G. H. Pezeshkpour, M. Yungbluth, F. Janotta, A. Aksamit, M. A. Martin and A. S. Fauci (1986). Detection of AIDS virus in macrophages in brain tissue from AIDS patients with encephalopathy. *Science* **233**, 1089–1093.

Lewin, S. R., M. Vesanen, L. Kostrikis, A. Hurley, M. Duran, L. Zhang, D. D. Ho and M. Markowitz (1999). Use of real-time PCR and molecular beacons to detect virus replication in human immunodeficiency virus type 1-infected individuals on prolonged effective antiretroviral therapy. *J. Virol.* **73**, 6099–6103.

Lewis, L. L. *et al*. (1996). Lamivudine in children with human immunodeficiency virus infection: a phase I/II study. The National Cancer Institute Pediatric Branch-Human Immunodeficiency Virus Working Group. *J. Infect. Dis.* **174**, 16–25.

Lifson, J. D., G. R. Reyes, M. S. McGrath, B. S. Stein and E. G. Engleman (1986). AIDS retrovirus induced cytopathology: giant cell formation and involvement of CD4 antigen. *Science* **232**, 1123–1127.

Louie, M., B. Ramratnam, R. Kost, A. Hurley, L. Zhang, E. Sun, S. Brun, I. Mcgowan, N. Ruiz, D. D. Ho and M. Markowitz (2001). Using viral dynamics to document the greater antiviral potency of a regimen containing Lopinavir/Ritonavir, Efavirenz, Tenofovir, and Lamivudine relative to standard HAART (abstr 383), in *8th Conference on Retroviruses and Opportunistic Infections*, Alexandria, VA, USA: Foundation for Retrovirology and Human Health.

McLean, A. R. and C. A. Michie (1995). In vivo estimates of division and death rates of human T lymphocytes. *Proc. Natl. Acad. Sci. USA* **92**, 3707–3711.

Mittler, J. E., M. Markowitz, D. D. Ho and A. S. Perelson (1999). Improved estimates for HIV-1 clearance rate and intracellular delay. *AIDS* **13**, 1415–1417.

Mohri, H., S. Bonhoeffer, S. Monard, A. S. Perelson and D. D. Ho (1998). Rapid turnover of T lymphocytes in SIV-infected rhesus macaques. *Science* **279**, 1223–1227.

Natarajan, V., M. Bosche, J. A. Metcalf, D. J. Ward, H. C. Lane and J. A. Kovacs (1999). HIV-1 replication in patients with undetectable plasma virus receiving HAART. *Lancet* **353**, 119–120.

Nowak, M. A. and C. R. Bangham (1996). Population dynamics of immune responses to persistent viruses. *Science* **272**, 74–79.

Pantaleo, G., C. Graziosi, J. F. Demarest, L. Butini, M. Montroni, C. H. Fox, J. M. Orenstein, D. P. Kotler and A. S. Fauci (1993). HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease. *Nature* **362**, 355–358.

Pantaleo, G., C. Graziosi, J. F. Demarest, O. J. Cohen, M. Vaccarezza, K. Gantt, C. Muro-Cacho and A. S. Fauci (1994). Role of lymphoid organs in the pathogenesis of human immunodeficiency virus (HIV) infection. *Immunol. Rev.* **140**, 105–130.

Perelson, A. S., P. Essunger, Y. Cao, M. Vesanen, A. Hurley, K. Saksela, M. Markowitz and D. D. Ho (1997). Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature* **387**, 188–191.

Perelson, A. S., A. U. Neumann, M. Markowitz, J. M. Leonard and D. D. Ho (1996). HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* **271**, 1582–1586.

Perno, C. F., F. M. Newcomb, D. A. Davis, S. Aquaro, R. W. Humphrey, R. Calio and R. Yarchoan (1998). Relative potency of protease inhibitors in monocytes/macrophages acutely and chronically infected with human immunodeficiency virus. *J. Infect. Dis.* **178**, 413–422.

Puddu, P., S. Fais, F. Luciani, G. Gherardi, M. L. Dupuis, G. Romagnoli, C. Ramoni, M. Cianfriglia and S. Gessani (1999). Interferon-gamma up-regulates expression and activity of P-glycoprotein in human peripheral blood monocyte-derived macrophages. *Lab. Invest.* **79**, 1299–1309.

Ramratnam, B., S. Bonhoeffer, J. Binley, A. Hurley, L. Zhang, J. E. Mittler, M. Markowitz, J. P. Moore, A. S. Perelson and D. D. Ho (1999). Rapid production and clearance of HIV-1 and hepatitis C virus assessed by large volume plasma apheresis. *Lancet* **354**, 1782–1785.

Ramratnam, B., J. E. Mittler, L. Zhang, D. Boden, A. Hurley, F. Fang, C. A. Macken, A. S. Perelson, M. Markowitz and D. D. Ho (2000). The decay of the latent reservoir of replication-competent HIV-1 is inversely correlated with the extent of residual viral replication during prolonged anti-retroviral therapy. *Nat. Med.* **6**, 82–85.

Sachsenberg, N., A. S. Perelson, S. Yerly, G. A. Schockmel, D. Leduc, B. Hirschel and L. Perrin (1998). Turnover of CD4+ and CD8+ T lymphocytes in HIV-1 infection as measured by Ki-67 antigen. *J. Exp. Med.* **187**, 1295–1303.

Schlegel, P. N. and S. K. Chang (1998). Physiology of male reproduction: the testes, epididymis, and ductus deferens, in *Campbell's Urology*, Vol. 7, P. C. Walsh, A. B. Retik, E. D. Vaughan and A. J. Wein (Eds), Philadelphia: WB Saunders Co., pp. 1254–1286.

Sharkey, M. E. *et al*. (2000). Persistence of episomal HIV-1 infection intermediates in patients on highly active anti-retroviral therapy. *Nat. Med.* **6**, 76–81.

Siliciano, J. D. and R. F. Siliciano (2000). Latency and viral persistence in HIV-1 infection. *J. Clin. Invest.* **106**, 823–825.

Smith, B. A. *et al*. (2001). Persistence of infectious HIV on follicular dendritic cells. *J. Immunol.* **166**, 690–696.

Sodroski, J., W. C. Goh, C. Rosen, K. Campbell and W. A. Haseltine (1986). Role of the HTLV-III/LAV envelope in syncytium formation and cytopathicity. *Nature* **322**, 470–474.

Stevenson, M. (1996). Portals of entry—uncovering HIV-1 nuclear transport pathways. *Trends Cell Biol.* **6**, 9–15.

Stevenson, M. and H. E. Gendelman (1994). Cellular and viral determinants that regulate HIV-1 infection in macrophages. *J. Leukoc. Biol.* **56**, 278–288.

Wei, X. *et al*. (1995). Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**, 117–122.

Wong, J. K., M. Hezareh, H. F. Gunthard, D. V. Havlir, C. C. Ignacio, C. A. Spina and D. D. Richman (1997). Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science* **278**, 1291–1295.

Yerly, S., T. V. Perneger, S. Vora, B. Hirschel and L. Perrin (2000). Decay of cell-associated HIV-1 DNA correlates with residual replication in patients treated during acute HIV-1 infection. *AIDS* **14**, 2805–2812.

Yoffe, B., D. E. Lewis, B. L. Petrie, C. A. Noonan, J. L. Melnick and F. B. Hollinger (1987). Fusion as a mediator of cytolysis in mixtures of uninfected CD4+ lymphocytes and cells infected by human immunodeficiency virus. *Proc. Natl. Acad. Sci. USA* **84**, 1429–1433.

Zhang, L. *et al*. (1999). Quantifying residual HIV-1 replication in patients receiving combination antiretroviral therapy. *N. Engl. J. Med.* **340**, 1605–1613.

Received 27 May 2001 and accepted 5 September 2001