Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 11514–11519, September 1998
Applied Mathematics, Medical Sciences

Drug concentration heterogeneity facilitates the evolution of drug resistance

(sanctuary sites / HIV)

THOMAS B. KEPLER* AND ALAN S. PERELSON†

*Biomathematics Program, Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203; and †Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545

Edited by Burton H. Singer, Princeton University, Princeton, NJ, and approved July 9, 1998 (received for review January 15, 1998)

ABSTRACT **Pathogenic microorganisms use Darwinian processes to circumvent attempts at their control through chemotherapy. In the case of HIV-1 infection, in which drug resistance is a continuing problem, we show that in one-compartment systems, there is a relatively narrow window of drug concentrations that allows evolution of resistant variants. When the system is enlarged to two spatially distinct compartments held at different drug concentrations with transport of virus between them, the range of average drug concentrations that allow evolution of resistance is significantly increased. For high average drug concentrations, resistance is very unlikely to arise without spatial heterogeneity. We argue that a quantitative understanding of the role played by heterogeneity in drug levels and pathogen transport is crucial for attempts to control re-emergent infectious disease.**

A tremendous range of pest organisms and parasites including bacteria, protozoans, fungi, macroparasites, insects, and weeds have used Darwinian processes to evade chemical control. The resurgence of infectious disease that has become a focal point of global health efforts is due in large part to the evolution of resistance to antibiotic drugs. Here, we concentrate our attention on the important example of drug resistance in HIV type 1 (HIV-1) infection.

Some progress in the quantitative understanding of the evolution of drug resistance by HIV-1 has been made by using relatively simple mathematical models, which consider the body as a single compartment (1–8). Using a protypic model, we show that the range of drug concentrations to which resistance is predicted to arise is unrealistically narrow. For mutants to arise, a parental drug-sensitive virus must replicate at some non-negligible rate. This occurs only at drug concentrations below some threshold level. On the other hand, the mutant must have a distinct advantage to overtake the parental strain. We will see that this occurs only for drug concentrations above some second threshold. The evolution of resistance takes place only between these two thresholds, which in a single compartment system is a narrow window of drug concentrations wherein both production of resistant mutants from their nonresistant precursors, and their subsequent selection can occur.

We suggest that, in nature, the window of opportunity for the generation of resistance is widened by having one compartment in which mutants are generated, such as a "sanctuary" or region of low drug concentration that HIV-1 can enter, and a second compartment in which the drug concentration is high enough to give mutants a selective advantage. For HIV-1, sanctuaries may be physiologically distinguished sites, such as the brain (9, 10) or testes or cell populations susceptible to infection in which intracellular bioavailability of active drug is poor. The process of generating mutants in one compartment and selecting them in another may be repeated several times in the step-wise accumulation of resistance mutations.

The notion that heterogeneity influences drug resistance is not new. Examination of the role of spatial heterogeneity in the spread of insecticide resistance (11, 12) has shown that, under appropriate conditons, spatial transport can give rise to an enhanced rate of spread of the resistant alleles. In the case of antibiotic resistance to tuberculosis, the effects of both temporal and spatial heterogeneity have been modeled (13, 14). Those models suggest that noncompliance to antibiotic regimes, but not spatial heterogeneity, is an important cause of treatment failure. These results cannot simply be applied to the case of HIV-1 infection in which the host–pathogen interactions are different and are described by different models.

Resistance to the protease inhibitor indinavir provides an instructive example because measurable resistance is found only in virus that has acquired at least three amino acid substitutions in HIV-1 protease (15). Granted that one- or even two-base substitutions might be found in the pretreatment viral quasispecies (16, 17), these intermediate strains do not have the ability to grow in the presence of drug at therapeutic concentrations. Therefore, for uniform concentrations of drug in the therapeutic range, these strains cannot produce the three-base mutants that finally do show resistance. But as we show, sanctuaries make it possible.

Approach. Our model will assume a parental virus population, at equilibrium, from which mutants arise. We examine the process by which a strain, j mutations from wild type, produces a more resistant strain with m additional substitutions, where all m are required to produce decreased drug susceptibility. For illustrative purposes, we use the example of indinavir resistance and take $j = 1$ and $m = 2$. This choice is not critical because the number of substitutions required will simply scale the overall time to appearance of the resistant strain. We use $j = 1$ because given the rapid replication rate of HIV-1, one-base change mutants almost certainly preexist (16, 17).

Here, we consider the simplest nontrivial case: two compartments, differing only in size and drug concentration, with movement of virus but not target cells between them. If the sanctuary is a subpopulation of target cells, then movement of target cells is not a concern. If the sanctuary is the central nervous system, then movement is possible but highly restricted due to the blood–brain barrier. Assuming no movement of target cells, this model is sufficiently simple that we can obtain analytical results on the rate at which viable mutants take hold and show that the range of average drug concentrations that allow for the evolution of drug resistance is significantly widened when sanctuaries exist.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

0027-8424/98/9511514-6$0.00/0
PNAS is available online at http://www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office.
Abbreviation: HIV-1, HIV type 1.

MODELS AND RESULTS

The model we use has been adapted from a simple model of HIV-1 dynamics (6, 18) but is quite general and applicable to a variety of other systems (cf. 19, 20).

Our aim is to compute the waiting time for the arrival of the first phenotypically resistant virus that survives and founds the resistant population. We do not model the growth of the resistant population once it has been established; by definition, it will grow to take over. By focusing on the production and initial establishment of the resistant variant, we need only consider the dynamics of the parental strain and the very early stages of the growth of the mutant population. At these very early times, the resistant population does not affect the parental population.

Dynamics of Virus Production. As in ref. 18, let T, T^* and V designate the population densities of uninfected target cells (i.e., cells susceptible to HIV-1 infection), productively infected cells, and free virions, respectively. We assume that target cells are generated by density-dependent proliferation as suggested by recent studies (21, 22) and ignore the possible input of cells from a thymic source. Then this system may be described by the equations

$$\frac{dT}{dt} = rT\left(1 - \frac{T}{T_0}\right) - kVT, \qquad [1]$$

$$\frac{dT^*}{dt} = kVT - \delta T^*, \text{ and} \qquad [2]$$

$$\frac{dV}{dt} = N\delta T^* - cV - kVT, \qquad [3]$$

where T_0 is the equilibrium density of T cells in the absence of virus, r is the maximum rate of T cell population growth, k is the viral infectivity, δ is the per cell rate of productively infected cell death, and c is the rate constant for virus clearance. Death of uninfected target cells is incorporated into the logistic growth term so that r presents the maximum proliferation rate minus the death rate.

Critical Infectivity. Eqs. **1–3** have the property that there is a critical infectivity, k_c, given by $k_c(N-1) = c/T_0$, such that for $k < k_c$, the only stable equilibrium is the noninfected state, $\bar{V} = \bar{T}^* = 0$, $\bar{T} = T_0$, where an overbar represents an equilibrium value. (The condition $k < k_c$ is equivalent to requiring that the basic reproductive number R_0 be < 1 (4).) For $k > k_c$, the stable equilibrium is an infected state with a T cell density less than T_0, i.e., $\bar{V} = r(1 - \bar{T}/T_0)/k$, $\bar{T} = c/(N-1)k$, and $\bar{T}^* = k\bar{V}\bar{T}/\delta$. In this model, if therapy reduces k below k_c, the virus will be eradicated, whereas less effective therapy will simply establish a new (lower) steady–state viral load.

We can compute k_c as a fraction of the infectivity k of the parental virus from measurement of the equilibrium T cell level and knowledge of its virus-free equilibrium value. This gives§

$$k_c = \frac{\bar{T}}{T_0}k. \qquad [4]$$

During the asymptomatic phase of infection, quasi-steady–states are established in which the $CD4^+$ T cell count may typically vary between 200 and 500 cells/mm^3 (23) giving ratios $\bar{T}/T_0$ between 0.2 and 0.5 under the assumption that the normal CD4 count $T_0 = 1{,}000/\text{mm}^3$ and that the T cell count measured in blood is reflective of the T cell levels in tissue. Thus, we would expect—if the assumptions going into the formulation of Eqs. **1–3**, including that of spatial homogeneity, are reasonable—that cutting the infectivity k by 50% to 80%, using antiretroviral drug therapy, should be sufficient to completely eliminate the virus. These predictions are not borne out by clinical practice. Monotherapy with zidovudine, which should have the required efficacy, does not eliminate the virus, even though resistance develops gradually by the stepwise accrual of mutations (24, 25). Similarly, monotherapy with the much more potent protease inhibitors (26) and zidovudine–lamivudine combination therapy (27), cases in which preexisting high level drug resistance is unlikely, do not agree with the prediction of viral elimination.

Thus, even before we consider the evolution of resistance, some assumptions of the model appear to need reconsideration. Several possibilities exist including the role of long-lived productively infected cells (28) and latently infected cells (29, 30, 31), but we will here focus on the assumption of spatial homogeneity.

Production and Selection of Mutants. The parental populations will be considered at equilibrium. The rate at which mutant viral strains are produced from the parental strain is then $\Omega\mu k\bar{V}\bar{T}$, i.e., the product of the specific mutation rate, μ, the rate of infection, $k\bar{V}\bar{T}$, and, since the populations are given as densities, a volume factor Ω to give absolute numbers.

Here, we consider mutations that occur because of errors in reverse transcription. Thus, after a virus infects a cell, the DNA copy of the viral genome that is made, the provirus, may carry a drug resistant mutation. Not only must the resistant provirus be produced, however, it must also "take root." Resistant proviruses may be produced but fail to produce progeny for purely stochastic reasons (32–34). The probability, p, that a provirus will propagate is related to the probability, q, that a free virion will propagate by

$$p = 1 - (1-q)^n, \quad q = \kappa p, \qquad [\mathbf{5a, b}]$$

where κ is the probability that a given resistant virion productively infects a cell and n is the number of progeny produced in the lifetime of an infected cell. The term $(1-q)^n$ in Eq. **5a** is the probability that all n of the progeny of the founder provirus fail to propagate, so $1-(1-q)^n$ is the probability that at least one of the progeny propagates. Eq. **5b** says that a virion propagates if and only if it infects (probability κ) and the provirus propagates (probability p).¶

If n is considered random rather than fixed, we take the expectation of Eq. **5a** over n. For n, a Poisson random variable with mean N and using Eq. **5b**, this becomes

$$p = 1 - \exp(-N\kappa p). \qquad [6]$$

Alternatively, if n is fixed at N, with N large, then Eq. **6** is an excellent approximation to Eq. **5**.

The infection probability κ is related to the kinetic parameters through

$$\kappa = k_r\bar{T}/(c + k_r\bar{T}), \qquad [7]$$

where k_r is the infectivity of the *resistant* virion. This equation comes about from considering the two possible fates of a virion: clearance and infection, and comparing their rates c and k_rT, respectively.

If $N\kappa < 1$, on average less than one cell is infected by the progeny of a productively infected cell, and $p = 0$ is the only non-negative solution of Eq. **6**.‖ By similar logic, one can

§In some models (5, 6, 19, 20), Eq. **1** is replaced by $dT/dt = \lambda - dT - kVT$, and $T_0 \equiv \lambda/d$ is the equilibrium level of target cells in the absence of virus. This model, as well as more general ones in which $dT/dt = f(T) - kVT$, with $f(T_0) = 0$ and $f'(T_0) < 0$, leads to an identical equation for k_c.

¶Here, p and q are the probabilities of a nonterminating line of descent. However, if we instead require that the resistance mutant only propagate for a large number of generations, p and q will be approximated by the solutions of Eq. **5a–5b**.

‖One can deduce this result by considering the graphs of the functions $y = p$ and $y = 1 - \exp(-N\kappa p)$ and observing that if $N\kappa < 1$, the latter function is below the line $y = p$ except when $p = 0$.

deduce that before drug is given, wild-type virus must have $N\kappa = 1$ to establish a quasi-steady–state level, where κ here refers to the infection by wild-type virus. Lastly, for $N\kappa > 1$ there is a positive solution, i.e., the resistant strains have a positive probability of surviving and of replacing the wild-type virus. (Interestingly, the condition $N\kappa = 1$ for wild-type virus is equivalent to the condition $k = c/\bar{T}(N-1)$, i.e., that k is equal to its critical value).

An immediate result is that, before the administration of drug, the drug-resistant mutants that preexist in the parental quasi-species are not self-sustaining. Assuming that there is a cost of resistance, either the infectivity or rate of replication of mutants in the absence of drug will be smaller than that of the wild-type, i.e., $N\kappa$ for the mutant must be less than $N\kappa = 1$ for the wild-type, and so the probability of propagation for these mutants is zero; they are only maintained by continual production from the wild-type.

The mean time τ to appearance of propagating mutants, or founders, is the inverse of their production rate:

$$\tau = (p\Omega\mu k\bar{V}\bar{T})^{-1}. \quad [8]$$

Effect of Drug. For simplicity, we assume the drug is a reverse transcriptase inhibitor and affects the infectivity of the virus. If the drug is a protease inhibitor, noninfectious virions are produced, which to a good approximation can be modeled by a change in infectivity (35). Let ϵ be the plasma (effective) drug concentration. Then, we may write

$$k(z) = k_0/(1 + \epsilon/\mathrm{IC}_{50}) = k_0/(1+z), \quad [9]$$

where k_0 is the viral infectivity in the absence of drug, z is a scaled drug concentration, $z = \epsilon/\mathrm{IC}_{50}$, and IC_{50} is the plasma drug concentration at which k is reduced to 50% of its drug-free value.

For the resistant strain, we use the form

$$k_r(z) = \rho k_0/(1+\beta z), \quad [10]$$

where $\rho \leq 1$ is a factor by which the resistant infectivity is decreased relative to the wild-type in the absence of drug, i.e., a cost of resistance, and $\beta \leq 1$ is a factor by which the drug concentration is effectively reduced for the resistant strain relative to the wild-type. The other viral parameters c and N are assumed unchanged in the presence of drug. In other viral diseases (20), it might be more appropriate to fix k and allow N to change due to drug, or to allow both to change; the results obtained in this alternative approach are qualitatively similar.

Window of Opportunity. One can now see why there is a problem in the generation of propagating resistant mutants. The production rate of resistant viruses, $p(\Omega\mu k\bar{V}\bar{T})$, is a product of two functions: the first, p, the probability of propagation, is an increasing function of z that is zero below some positive value of z, which we call z_L (where $N\kappa(z_L) = 1$), while the second, the overall mutant production rate, is a decreasing function of the drug concentration z that becomes zero at some finite value of z, which we call z_U (where $k(z_U) = k_c$ and hence $V = 0$). As shown in Fig. 1, the product of these functions generates a curve that has a single maximum and which goes to zero at finite values of z on either side of the hump. Richman (36) has presented a strikingly similar figure for the production of drug resistant mutants at different drug concentrations based on qualitative arguments. We now have explicitly calculated the shape of this curve for a simple one-compartment model. What is interesting about our theoretical result is that it shows that, only within a window in z, between z_L and z_U, can mutants be generated. This window is surprisingly narrow with z_U corresponding to a concentration of the order of the IC_{50} of the drug. By going to a two-compartment model with drug concentration differences between the two, the viral window of opportunity is widened considerably, as we now show.

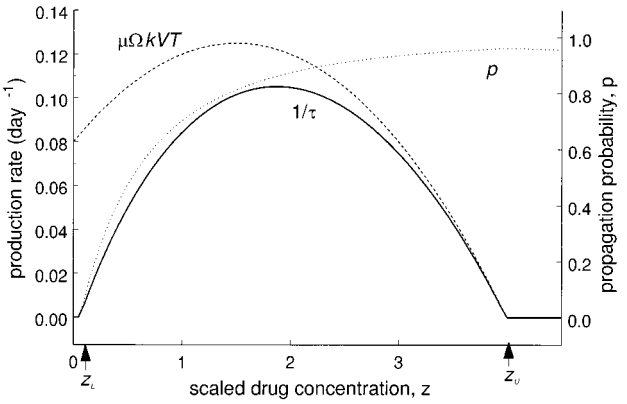

FIG. 1. In a one-compartment model the mean production rate of resistant virions (solid line), equal to the inverse of the mean time to the arrival of the founding resistant virus and labeled $1/\tau$, is nonzero only over a finite window of drug concentrations, z, from z_L to z_U. According to our theory, the production rate is the product of two functions (dashed lines). The curves were computed assuming Eqs. **1–3** were at steady-state and that $k_0 = 1.5 \times 10^{-5}$ mm^3 day^{-1}, $c = 3$ day^{-1}, $T_0 = 1000$/mm^3, $r = 0.01$ day^{-1}, $N = 1000$, $\beta = 0.1$, $\rho = 0.95$. With these values, the equilibrium values of T and V are 200/mm^3 and 5.3×10^5/ml, respectively. Note that above $z = 4$, $k < k_c$ and the production rate falls to zero. The scaling factors for the production rate, $\mu\Omega$ are $\mu = 2 \times 10^{-10}$, corresponding to the acquisition of two independent mutations, and $\Omega = 2.5 \times 10^8$ mm^3, assuming a 5 liter blood capacity and that the T cell count in the blood needs to be multiplied by 50 to account for the 98% of CD4$^+$ T lymphocytes that are in tissues. Note that the production rate curves scale simply with the mutation rate μ, so that if only one mutation were required the waiting time would be orders of magnitude shorter.

Two Compartment Model. Now consider virus occupying two compartments and moving between them by passive transport. The first compartment will be considered the bulk compartment with larger volume and larger drug concentration. The second will represent the sanctuary, with lower drug concentrations and smaller volume. Let subscripts 1 and 2 designate T cells or virus particles per unit volume specific to each of the two compartments, with $u_1 = 1 - u$ and $u_2 = u$, being the relative volumes of the two compartments.

The generalization of Eqs. **1–3** for this case is then

$$\frac{dT_i}{dt} = rT_i\left(1 - \frac{T_i}{T_0}\right) - k_i V_i T_i, \quad [11]$$

$$\frac{dT_i^*}{dt} = k_i V_i T_i - \delta T_i^*, \quad [12]$$

$$\frac{dV_i}{dt} = N\delta T_i^* - cV_i - k_i V_i T_i + D_i(V_{\hat{i}} - V_i), \quad [13]$$

where the circumflexed subscript means "the other one"; e.g., $\hat{2} = 1$.

In this model, we neglect, for simplicity, the passage of target cells between compartments but allow passive transport of virus between the two compartments characterized by the transport coefficients D_i. This form of the transport coefficient arises from requiring that there is no net flow of virus between compartments when the respective concentrations are equal. Requiring that transport itself does not lead to an increase or decrease in the total amount of virus yields

$$0 = \frac{d}{dt}(u_1V_1 + u_2V_2) = u_1D_1(V_2 - V_1) + u_2D_2(V_1 - V_2), \quad [14]$$

where only the transport terms have been considered. This is satisfied for all V_1, V_2 only when $D_1 = u_2D_2/u_1$. Therefore, we assume this condition for our model and write $D_2 = D$ and $D_1 = uD/(1-u)$. We expect that D_i changes with the volume of the compartments in a way that is dependent on the

specific transport mechanisms involved and on geometric details. For example, if transport is by diffusion, the surface area of the interface between the compartments (possibly $u^{2/3}$) enters. Alternatively, if transport is by convection, D_i in simple models depends on the volume flow rate divided by the compartment volume. Here D is a parameter, but we caution that its value can depend on u.

Although we have very little knowledge of the "true" value of D, we find that the system functions as if it has a "sanctuary" only if neither D_1 or D_2 is significantly larger than the viral clearance rate. Thus, we will take D_2, which is the larger of the two (because we are taking compartment 2 as the smaller compartment), to be of the order of c.

It may eventually be possible to make inferences about the magnitude of viral transport from studies of genetically distinguishable viruses isolated in different compartments, e.g., brain isolates vs. spleen isolates (10). However, such data is confounded by differential selection in the two areas, a factor about which very little is known. Here, we focus on the effect of spatial heterogeneity *per se* and neglect differential selection between compartments. Thus, the infectivities k_i are given as in the single compartment model, but now the drug concentrations in the two compartments differ, so we have z_i rather than z.

Production and Selection of Mutants. A virus in compartment i can do one of three things: (*i*) productively infect, thereby producing an average of N progeny, (*ii*) perish without leaving any progeny, or (*iii*) move to the other compartment. Let the probability that the virus infects productively in compartment i be denoted κ_i and the probability that it moves to the other compartment be denoted m_i. If p_i and q_i are the probabilities that a provirus and virus in compartment i propagate, respectively, they will be described by the equations

$$p_1 = 1 - \exp(-q_1 N), \quad q_1 = m_1 q_2 + \kappa_1 p_1, \qquad \textbf{[15a, b]}$$

$$p_2 = 1 - \exp(-q_2 N), \quad q_2 = m_2 q_1 + \kappa_2 p_2. \qquad \textbf{[15c, d]}$$

Eqs. **15a** and **15c** have the same interpretation as Eq. **6**, i.e., in order for a provirus in compartment i to propagate (probability p_i), at least one of its progeny virions must propagate (again, we are assuming a large fixed number or a Poisson distribution of offspring numbers). In order for a virion in compartment 1 to propagate (Eq. **15b**), it must either infect a cell and propagate as a provirus (probability $\kappa_1 p_1$) or move to compartment 2 and propagate (probability $m_1 q_2$).

Eqs. **15a–d** are always satisfied by the trivial solution $(p_1, q_1, p_2, q_2) = (0, 0, 0, 0)$. For any fixed value of the product $m_1 m_2$, we can define a critical curve in the κ_1, κ_2 plane, on one side of which, in addition to the trivial solution, a nontrivial solution is produced:** This curve is the lower branch of the hyperbola given by $(N\kappa_1 - 1)(N\kappa_2 - 1) = m_1 m_2$. For (κ_1, κ_2) values to the right of this curve, a newly produced virion has a positive probability of propagating. Note that when either m_1 or m_2 vanishes, this condition becomes $N\kappa_i > 1$ for either i.

The stochastic parameters are related as before to the dynamic parameters through

$$m_i = \frac{D_i}{D_i + c + k_{ri}\bar{T}_i} \quad \text{and} \quad \kappa_i = \frac{k_{ri}\bar{T}_i}{D_i + c + k_{ri}\bar{T}_i}. \qquad \textbf{[16]}$$

For this system, the waiting time for production of propagating mutant virus is the reciprocal of the sum of the net production rates in the two compartments,

$$\tau = (\Omega\mu[(1-u)k_1\bar{T}_1\bar{V}_1 p_1 + uk_2\bar{T}_2\bar{V}_2 p_2])^{-1}. \qquad \textbf{[17]}$$

Below we provide numerical solutions to the model for a particular choice of parameter values, described in the caption to Fig. 1, which may be viewed as characteristic of a mid-stage AIDS patient. The figures showing these solutions are meant to be illustrative of various general principles and trends that occur as parameters characterizing the sanctuary and drug regime are varied. While, for simplicity, we call τ the mean time to resistance, it is only the time for the initial production of a propagating mutant, not the time until phenotypic resistance would be observed in a patient.

A Widened Window of Opportunity. In a sanctuary, where the drug penetrance is small, partially resistant strains can proliferate and produce more resistant mutants, which can then leave the sanctuary and proliferate in the bulk compartment. Thus, in the presence of a sanctuary, the window of opportunity is dramatically widened. For example, as shown in Fig. 2, when the relative sanctuary volume $u_2 = 0.001$, the window is widened ≈10-fold from an upper threshold of $z_1 = 4$ for the one-compartment model to approximately $z_1 = 40$ for the two-compartment model. Above this new upper threshold, resistance is unlikely. This result is more in line with experience in which drug concentrations of 50- to 100-fold greater than the IC_{50} have therapeutic value.

There are now two distinct regimes to the window. In the first, corresponding to the original window where the infectivity $k_r(z)$ is above its critical value, the step-wise acquisition of mutants is very rapid. In the second regime, where the bulk infectivity $k_r(z)$ is below the critical value, the waiting time between subsequent stages in the acquisition of resistance may be quite long but in the presence of the sanctuary, finite.

In Fig. 2, the sanctuary was completely drug-free. Fig. 3 shows the waiting time between mutations as the drug penetrance, z_2/z_1, is varied. For very high bulk concentrations, z_1, the effect of heterogeneity is essentially lost. The drug concentration in the bulk is too high even for the resistant strain. The sanctuary now acts as a single compartment in which both production and selection must occur. For very small values of z_2, the drug concentration in the sanctuary, we again have increasingly large waiting times (see the curves labeled $z_1 = 16$, 32, or 64 in Fig. 3). This effect, in which evolution of drug resistance is prevented by lowering the concentration of drug in the sanctuary (i.e., lowering the selective advantage of a resistant mutant), is likely particular to the

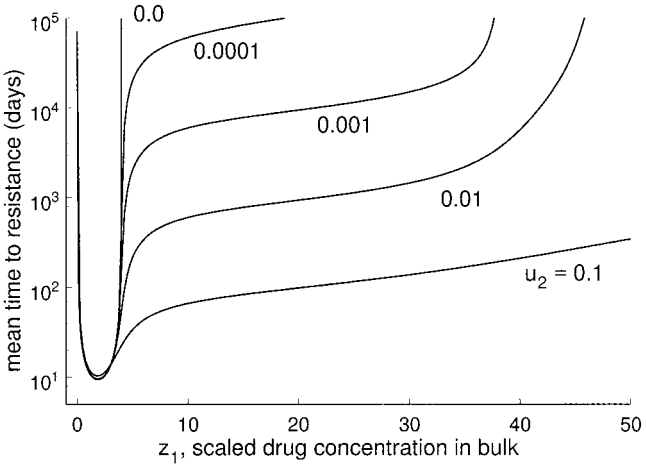

FIG. 2. Mean time to resistance, τ, vs. z_1, the drug concentration in the bulk compartment for various relative volumes, u_2, of the sanctuary, assumed to be drug-free, i.e., $z_2 = 0$. The curve labeled 0.0 has no sanctuary and represents the reciprocal of the quantity shown in the solid line in Fig. 1. Parameters are as in Fig. 1, and $D = 5$ day^{-1}.

Near the critical curve, p_i and q_i are small. Thus, Eqs. **15a–d can be linearized about the origin and written, to first order, in the form $A(p_1, q_1, p_2, q_2)^T = 0$, where A is a matrix. This equation has a nontrivial solution if and only if the determinant of A vanishes. The equation for the hyperbola results from setting the determinant of A to zero.

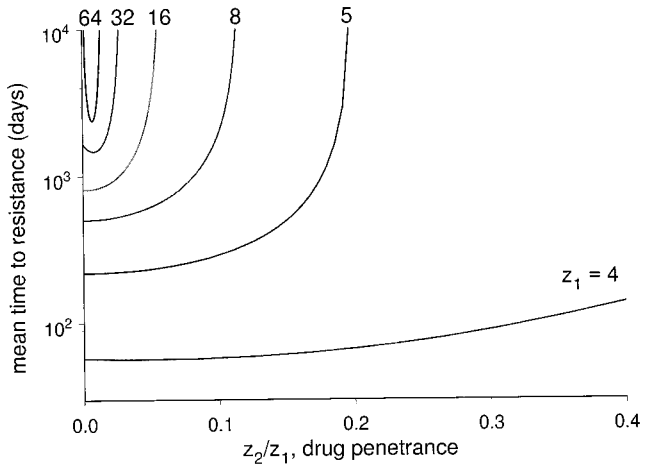

FIG. 3. Mean time to resistance, τ, vs. the drug penetrance z_2/z_1. z_1 is held constant at the indicated value for each of the curves shown and parameters are as in Fig. 2. The sanctuary has relative volume $u = 0.01$. For average concentrations above $z_1 = 4$, a homogeneous system has infinite waiting time (Fig. 1). As seen here, heterogeneous systems are able to produce resistant strains even when $z_1 > 4$.

two-compartment model. In a model with many compartments or a continuum within which gradients of drug concentration can be found, facilitation of the evolution of resistance would be found even at small sanctuary drug concentrations (see below).

The rate at which virus is transported between compartments also affects the mean time needed for drug resistance to arise (Fig. 4). There is an "optimal" transport coefficient at which the heterogeneity between compartments is maximal and for which the mean time to resistance is at a minimum. This optimum occurs for D of order 5–10 days^{-1} for the drug concentrations shown in Fig. 4. For $z_1 = 4$, the mean waiting time is short and virtually independent of the transport coefficient. For higher drug concentrations, the waiting time rises quite sharply as the larger transport coefficient D becomes larger than the clearance rate c and acts like an enhancement of this clearance. In the other limit, in which D is very small, the compartments are approaching isolation, and the resistant mutants produced in the sanctuary cannot easily move to the high-drug compartment where they have an advantage. This effect is not as dramatic when the sanctuary has a drug concentration sufficiently high that the resistant mutants can still propagate without moving to the other compartment, though moving would increase their chances.

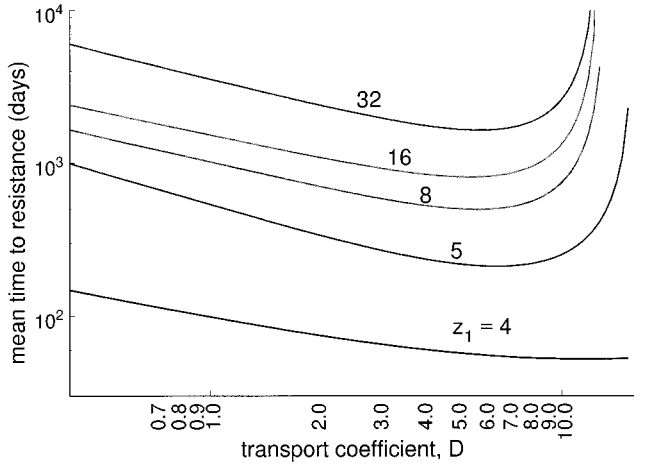

FIG. 4. The effect of increasing transport coefficient D on the mean waiting time until resistance, τ. Compartment 2 is drug-free ($z_2 = 0$) and of relative size $u = 0.01$. Other parameters are as in Fig. 1. The drug concentration z_1 in the larger compartment labels the appropriate curves.

DISCUSSION

The evolution and spread of drug-resistant pathogens is known to occur with great robustness in a wide variety of situations. We have shown that simple models that fail to consider heterogeneity of drug concentration underestimate the range of mean drug concentrations that support the establishment of drug resistance (Fig. 1). The existence of even quite small sanctuaries, places where the drug concentration is much smaller than in the bulk compartment, can greatly enhance the probability of generating resistant mutants. The sanctuaries provide a place where ongoing replication of the parental strain continues and therefore allows mutants to be produced. Once produced, they may then migrate to the bulk areas where the drug concentration is higher and where they can exercise their advantage and replicate.

Our analysis is based upon a relatively simple two-compartment model with fixed drug concentrations in either compartment. Further improvements to this approach can be anticipated. First, a spatially continuous model is likely to reveal further effects of concentration gradients. Our preliminary analyses of models based on partial-differential equations in one spatial dimension (T.K., Babai, and A.S.P., unpublished results) show that the step-wise accumulation of resistance mutations is facilitated by the presence of continuous gradients of drug concentration. Within these gradients, there are locations within which the conditions for growth of any given level of drug-resistance are ideal. As new strains are produced, they migrate and preferentially replicate at their ideal locations, producing the next level of resistant mutants. The process continues with subsequent levels of resistant strains climbing the gradient of drug concentration. Here, we have viewed compartments as being spatially distinct. Another possibility is that drug penetrance differs within different cell populations and thus different cell populations may comprise different compartments. If there were many such populations, it would be analogous to having a gradient in drug concentration.

Another improvement would be to include temporal fluctuations. Drugs are administered at discrete times, so that the concentration of drug fluctuates temporally. We have found a facilitation of drug resistance evolution due to spatial inhomogeneities, and we expect that there may be a similar enhancement due to temporal inhomogeneities.

The role of sanctuaries in the evolution of drug resistance is likely to be even more central for multi-drug therapies. When three or more drugs are used, many mutations are typically required to confer resistance to the therapy as a whole. But even long times on these therapies with no detectable serum virus should be greeted with caution. Waiting times for fully resistant strains in the presence of sanctuaries can be quite long. However, in some cases it is infinite. Thus, emergence of drug resistance is not inevitable; the conditions under which we expect emergence are given by Eqs. **15–17**.

An effective strategy for reducing human sickness and mortality caused by infectious microorganisms will necessitate a far more complete understanding of the large-scale patterns of drug resistance evolution than is presently available. More sophisticated mathematical models that account for spatial and temporal structure, in conjunction with improved experimental measurments of drug and virus levels in multiple compartments, will likely be a tool of great importance in this ongoing endeavor.

We thank J. Mittler, S. Blower, and L. Segel for useful comments and D. Babai for programming help. Portions of this work were performed under the auspices of the U.S. Department of Energy and supported by National Science Foundation Grant MCB–9357637 (to

T.B.K.), National Institutes of Health Grants RR06555 and AI40387 (to A.S.P.), the Santa Fe Institute, and the Joseph P. and Jeanne M. Sullivan Foundation.

1. McLean, A. R. & Nowak, M. A. (1992) *AIDS* **6,** 71–79.
2. Frost, S. D. W. & McLean, A. R. (1994) *AIDS* **8,** 323–332.
3. Stilianakis, N. I., Boucher, C. A., De Jong M. D., Van Leeuwen, R., Schuurman, R. & De Boer, R. J. (1997) *J. Virol.* **71,** 161–168.
4. Nowak, M. A., Bonhoeffer, S., Shaw, G. M. & May, R. M. (1997) *J. Theor. Biol.* **184,** 203–217.
5. Bonhoeffer, S. & Nowak, M. A. (1997) *Proc. R. Soc. London Ser. B* **264,** 631–637.
6. Bonhoeffer, S., May, R. M., Shaw, G. M. & Nowak, M. A. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 6971–6976.
7. Ribeiro, R. M., Bonhoeffer, S. & Nowak, M. A. (1998) *AIDS* **12,** 461–465.
8. Wein, L. M., D'Amato, R. M. & Perelson, A. S. (1998) *J. Theor. Biol.* **192,** 81–98.
9. Pialoux, G., Fournier, S., Moulignier, A., Poveda, J.-D., Clavel, F. & Dupont, B. (1997) *AIDS* **11,** 1302–1303.
10. Wong, J. K., Ignacio, C. C., Torriani, F., Havlir, D., Fitch, N. J. S. & Richman, D. D. (1997) *J. Virol.* **71,** 2059–2071.
11. Spieth, P. T. (1974) *Genetics* **78,** 961–965.
12. Steiner, W. W. M. (1993) in *Evolution of Insect Pests: Patterns of Variation*, eds. Kim, K. C. & Mc Pheron, B. A., (Wiley, New York).
13. Lipsitch, M. & Levin, B. R. (1997) *Antimicrob. Agents Chemother.* **41,** 363–373.
14. Lipsitch, M. & Levin, B. R. (1998) *Int. J. Tuberc. Lung Dis.* **2,** 187–199.
15. Condra, J. H., Schleif, W. A., Blahy, O. M., Gabryelski, L. J., Graham, D. J., Quintero, J. C., Rhodes, A., Robbins, H. L., Roth, E., Shivaprakash, M., *et al.* (1995) *Nature (London)* **374,** 569–571.
16. Coffin, J. M. (1995) *Science* **267,** 483–489.
17. Perelson, A. S., Essunger, P. & Ho, D. D. (1997) *AIDS* **11,** Suppl. A, S17–S24.
18. Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M. & Ho, D. D. (1996) *Science* **271,** 1582–1586.
19. Nowak, M. A., Bonhoeffer, S., Hill. A. M., Boehme, R., Thomas, H. C. & McDade, H. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 4398–4402.
20. Lam, N. P., Neumann, A. U., Gretch, D. R., Wiley, T. E., Perelson, A. S. & Layden, T. J. (1997) *Hepatology* **26,** 226–231.
21. Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M. & Markowitz, M. (1995) *Nature (London)* **373,** 123–126.
22. Sachsenberg, N., Perelson, A. S., Yerly, S., Schockmel, G. A., Leduc, D., Hirschel, B. & Perrin, L. (1998) *J. Exp. Med.* **187,** 1295–1303.
23. Farizo, K. M., Buehler, J. W., Chamberland, M. E., Whyte, B. M., Froelicher, E. S., Hopkins, S. G., Reed, C. M., Mokotoff, E. D., Cohn, D. L., Troxler, S., *et. al.* (1992) *J. Am. Med. Assoc.* **267,** 1798–1805.
24. Larder, B. A. & Kemp, S. D. (1989) *Science* **246,** 1155–1158.
25. Kellam, P., Boucher, C. A. B. & Larder, B. A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 1934–1938.
26. Markowitz, M., Saag, M., Powderly, W. G., Hurley, A. M., Hsu, A., Valdes, J. M., Henry, D., Sattler, F., La Marca, A., Leonard, J. M. & Ho, D. D. (1995) *N. Engl. J. Med.* **333,** 1534–1539.
27. Eron, J., Benoit, S. L., Jemsek, J., MacArthur, R. D., Santana, J., Quinn, J. B., Kuritzkes, D. R., Fallon, M. A. & Rubin, M. (1995) *N. Engl. J. Med.* **333,** 1662–1669.
28. Perelson, A. S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., Markowitz, M. & Ho, D. D. (1997) *Nature (London)* **387,** 188–191.
29. Finzi, D., Hermankova, M., Pierson, T., Carruth, L. M., Buck, C., Chaisson, R. E., Quinn, T. C., Chadwick, K., Margolick, J., Brookmeyer, R., *et al.* (1997) *Science* **278,** 1295–1300.
30. Wong, J. K., Hezareh, M., Gunthard, H. F., Havlir, D. V., Ignacio, C. C., Spina, C. A. & Richman, D. D. (1997) *Science* **278,** 1291–1295.
31. Chun, T. W., Stuyver, L., Mizell, S. B., Ehler, L. A., Mican, J. A., Baseler, M., Lloyd, A. L., Nowak, M. A. & Fauci, A. S. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 13193–13197.
32. Han, S., Hathcock, K., Zeng, B., Kepler, T. B., Hodes, R. & Kelsoe, G. (1995) *J. Immunol.* **155,** 556–567.
33. Radmacher, M. D., Kelsoe, G. & Kepler, T. B. (1998) *Immunol. Cell Biol.*, **76,** 373–381.
34. Leigh-Brown, A. J. & Richman, D. D. (1997) *Nat. Med.* **3,** 268–271.
35. Herz, A. V. M., Bonhoeffer, S., Anderson, R. M., May, R. M. & Nowak, M. A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 7247–7251.
36. Richman, D. (1996) *Antiviral Res.* **29,** 31–33.