Bulletin of Mathematical Biology, Vol. 59, No. 5, pp. 881–896, 1997
Elsevier Science Inc.

0092-8240/97 $17.00 + 0.00

S0092-8240(97)00034-7

VIRUS LOAD AND ANTIGENIC DIVERSITY

■ BARBARA BITTNER,* SEBASTIAN BONHOEFFER and MARTIN A. NOWAK†
Department of Zoology,
University of Oxford,
Oxford OX1 3PS, England

(*E.mail: martin.nowak@zoo.ox.ac.uk*)

In this paper, we analyse mathematical models for the interaction between virus replication and immune responses. We show that the immune system can provide selection pressure for or against viral diversity. The paper provides new insights into the relationship between virus load (= the abundance of virus in an infected individual) and antigenic diversity. Antigenic variation can increase virus load during infections, but the correlation between load and diversity in comparisons among different infected individuals can be positive or negative, depending on whether individuals differ in their cross-reactive or strain-specific immune responses. We derive two models: our first model applies to any replicating parasite that can escape from immune responses; our second model includes immune function impairment, and specifically describes infections with the human immunodeficiency virus (HIV). © 1997 Society for Mathematical Biology

1. Introduction. Virus load is an important determinant of the outcome of infection with many viruses. In HIV infection, for example, asymptomatic patients tend to have low virus load, while patients with AIDS usually have high virus load. Mellors *et al.* (1996) showed that the virus load in the first six months after infection is a strong predictor of the rate of progression to AIDS. Patients with a low virus load remain asymptomatic for a long period, while patients with high virus load progress rapidly to disease and death. In infections with the human T cell leukemia virus (HTLV-1), a large virus load is associated with chronic inflammatory condition, while asymptomatic patients usually have a 10- to 100-fold lower virus load (Bangham, 1993).

Many viruses reproduce rapidly in infected individuals and continuously generate new mutants—so-called quasi-species (Eigen and Schuster, 1977). The antiviral immune response provides selection pressure that favours virus mutants which can escape from immune-mediated destruction. This "antigenic variation" has been observed in several virus infections, including HIV-1, HIV-2, HTLV-1, influenza virus, equine infectious anemia virus,

*Present address: Institut für Algebra und Diskrete Mathematik, Technische Universität Wien, A-1040 Vienna, Austria.
†Author to whom correspondence should be addressed.

and visna medi virus (Balfe *et al.*, 1990; Holmes *et al.*, 1992; Schulz *et al.*, 1990; Burns and Desrosiers, 1991; Overbaugh *et al.*, 1991; Baier *et al.*, 1980; Clements *et al.*, 1980; Salinovich *et al.*, 1986; Ellis *et al.*, 1987). Variation is not only observed in epitopes recognised by antibody responses, but also in epitopes recognised by cytotoxic T cell responses (CTL). There is evidence for escape from CTL responses in infections with HIV-1, HTLV-1, hepatitis B virus, LCMV and mouse retrovirus (Phillips *et al.*, 1991; Niewiesk *et al.*, 1995; Pircher *et al.*, 1990; Moskophidis and Zinkernagel, 1995; Meier *et al.*, 1995; Ferrari *et al.*, 1996; Nowak *et al.*, 1995; Price *et al.*, 1996; Goulder *et al.*, 1996; Borrow *et al.*, 1996).

Virus evolution during individual infections and antigenic variation has been proposed as a potential mechanism for disease progression in HIV infections (Nowak *et al.*, 1990; Nowak *et al.*, 1991; Nowak *et al.*, 1995). This theory of HIV infection is based on the assumptions that: 1) virus load causes disease, 2) immune responses reduce virus load, and 3) virus evolution during single infections increases virus load. The key result of this theory is a dynamic threshold condition, which can be breached by virus evolution leading to increasing antigenic diversity. This shaped the name "antigenic diversity threshold," but variation in other parameters such as development of faster replicating virus strains can also overcome the threshold (Nowak and May, 1992; de Boer and Boerlijst, 1994). It has also been proposed that increasing immune activation leads to increasing virus load and disease progression in HIV infection (McLean and Nowak, 1992a).

Mathematical models of HIV infection have also been developed to describe the dynamics of antiviral drug treatment (McLean and Nowak, 1992b; Frost and McLean, 1994; Wei *et al.*, 1995; Ho *et al.*, 1995; de Boer and Boucher, 1996; Herz *et al.*, 1996; Perelson *et al.*, 1996; Nowak *et al.*, 1997). This has led to estimates of turnover rates of HIV-infected cells and free virus in infected individuals.

In this paper, we provide new insights into the relation between virus load and antigenic diversity. We will first analyse a model without immune function impairment. Such a model does not have a diversity threshold, but antigenic diversity increases virus load during individual infections. This first model applies to all viruses (or more generally, parasites) that cause persistent infections and can generate antigenic variation. The second model includes immune function impairment, and applies specifically to HIV infection.

2. A Mathematical Model. The model describes a replicating viral (or other) pathogen which is opposed by strain-specific and cross-reactive immune responses. Let us consider the following system of ordinary differ-

ential equations:

$$\begin{aligned} dv_i/dt &= v_i(r_i - p_i x_i - q_i z) \qquad i = 1, \ldots, n \\ dx_i/dt &= c_i v_i - b_i x_i \qquad i = 1, \ldots, n \\ dz/dt &= \sum_{j=1}^{n} k_j v_j - bz. \end{aligned} \tag{1}$$

Here, v_i denotes the concentration of virus strain (or mutant) i, x_i is the magnitude of the specific immune response against strain i, and z is the magnitude of the cross-reactive immune response directed at all strains. There are n strains; each strain is characterised by a set of parameters: r_i denotes the replication rate of strain i, p_i is the rate at which the strain is killed by specific immune responses, q_i is the rate at which the strain is killed by cross-reactive immune responses, and c_i and k_i are the rates at which strain i induces specific and cross-reactive immune responses. Specific immune responses x_i decline at rate b_i; the cross-reactive immune response declines at rate b.

The system has 2^n possible equilibria given by the equations

$$v_i = 0 \qquad \text{or} \qquad r_i = \frac{c_i p_i}{b_i} v_i + \frac{q_i}{b} \sum_{j=1}^{n} k_j v_j \tag{2}$$

$$x_i = \frac{c_i}{b_i} v_i \tag{3}$$

$$z = \frac{1}{b} \sum_{j=1}^{n} k_j v_j. \tag{4}$$

It is convenient to rescale the virus concentrations $v_i' = k_i v_i$, which leads to

$$r_i = \frac{c_i p_i}{b_i k_i} v_i' + \frac{q_i}{b} v'. \tag{2a}$$

Here, $v' = \sum_{j=1}^{n} v_j'$. We will now show that among the 2^n possible equilibria, there is only one stable equilibrium. For the stable equilibrium, we must have that all strains with $v_i' = 0$ must be unable to invade, i.e. their transversal eigenvalue $\partial \dot{v}_i'/\partial v_i' = r_i - p_i x_i - q_i z$ at this equilibrium must be negative. (The transversal eigenvalue denotes the per-capita growth rate of virus strain i at the relevant boundary equilibrium where $v_i' = 0$. "Transversal eigenvalue" is an expression coined by Hofbauer and Sigmund (1988).) We assume that there is no specific immune response at the beginning, and

therefore $x_i = 0$. This leads to the condition $r_i - q_i z < 0$ or $r_i/q_i < v'/b$. It follows that we can rank the virus strains according to their ratio r_i/q_i. Without loss of generality, we label the strains such that

$$r_1/q_1 > r_2/q_2 > \cdots > r_n/q_n. \tag{5}$$

This rank order implies that if a certain equilibrium cannot be invaded by strain i, it also cannot be invaded by any strain with an index greater than i. Thus, it is sufficient to consider equilibria of the form $v_i' > 0$ for $i = 1, \ldots, m$ and $v_i' = 0$ for $i = m+1, \ldots, n$. Let us call such an equilibrium E_m, and let us denote the total (rescaled) virus population size at this equilibrium by V_m. By summation of (2a) over $i = 1, \ldots, m$, we obtain

$$V_m = \sum_{i=1}^{m} k_i \alpha_i \Big/ \left(1 + \sum_{i=1}^{m} k_i \beta_i\right) \tag{6}$$

with $\alpha_i = b_i r_i/(c_i p_i)$ and $\beta_i = b_i q_i/(b c_i p_i)$.

Equilibrium E_m is only stable if two conditions are fulfilled. First, strain m must be able to invade equilibrium E_{m-1}. This is the case if

$$r_m/q_m > \frac{1}{b} V_{m-1}, \tag{7}$$

which is equivalent to $V_m > V_{m-1}$. Second, strain $m+1$ must be unable to invade equilibrium E_m. The condition that strain $m+1$ cannot invade is

$$r_{m+1}/q_{m+1} < \frac{1}{b} V_m. \tag{8}$$

It is straighforward to show that this condition is equivalent to V_m being larger than V_{m+1}.

Therefore, the unique stable equilibrium of system (1) is characterised by the highest rescaled virus load among all equilibria. If all strains stimulate the cross-reactive response at the same rate, $k_i = k$, then the unique stable equilibrium is characterised by the highest viral load among all possible equilibria. In this sense, selection maximises viral load. If there are differences in the stimulation of the cross-reactive response, then there is no longer selection for maximum virus load, and a strain which induces a lower viral load can outcompete a strain which induces a higher viral load (Bonhoeffer and Nowak, 1994; Bonhoeffer and Nowak, 1995).

Thus, the way to define the stable equilibrium is the following: order all strains according to their ratios r_i/q_i. Then determine the index m which maximises V_m, the rescaled virus population size. If we go back to our original variables v_i, we can write the equilibrium load v by using (2a) and

(6). With m defined in this way, the stable equilibrium is given by

$$v_i = \alpha_i - \beta_i V_m \tag{9}$$

and the total viral load by

$$v = \sum_{i=1}^{m} v_i = \sum_{i=1}^{m} \alpha_i - \sum_{i=1}^{m} \beta_i V_m. \tag{10}$$

We have now characterised the dynamical behaviour of system (1), and have calculated the unique globally stable equilibrium. In order to gain further analytical insight, particularly into the relation between virus load and diversity, we consider the special case where $p_i = p$, $q_i = q$, $b_i = b$, and $k_i = k$ for all strains. Thus, we analyse the simplified system

$$\begin{aligned} dv_i/dt &= v_i(r_i - px_i - qz) \qquad i = 1, \ldots, n \\ dx_i/dt &= c_i v_i - bx_i \qquad i = 1, \ldots, n \\ dz/dt &= kv - bz. \end{aligned} \tag{11}$$

We rank the strains according to $r_1 > r_2 > \ldots > r_n$. At equilibrium, the individual virus strains have the abundances

$$v_i = \frac{br_i - kqv}{c_i p} \qquad i = 1, \ldots, m \tag{12}$$

and $v_i = 0$ for $i = m+1, \ldots, n$. The total virus population size is

$$v = b \sum_{i=1}^{m} \frac{r_i}{c_i} \Bigg/ \left(p + kq \sum_{i=1}^{m} \frac{1}{c_i} \right), \tag{13}$$

and m is the largest integer fulfilling

$$r_m > kq \sum_{i=1}^{m-1} \frac{r_i}{c_i} \Bigg/ \left(p + kq \sum_{i=1}^{m-1} \frac{1}{c_i} \right). \tag{14}$$

Looking at (12), it is clear that $br_i - kqv$ is a declining sequence with i, and that $br_i - kqv > 0$ if $i \leqslant m$ and $br_i - kqv < 0$ if $i > m$. For a large number of strains m, a good approximation is $br_m - kqv = 0$. Hence, for the relation

between virus load v and diversity m, we derive the result

$$v = \frac{b}{kq} r_m. \tag{15}$$

Figure 1 shows a computer simulation of the model described by (11). New virus strains are added over time, which increases virus diversity and virus load. The figure shows infection dynamics in two patients. One patient has weak strain-specific immune responses, while the other patient has strong strain-specific immune responses. In the weak responder, virus load soon grows to high levels, but there is little selection pressure for antigenic variation. In the strong responder, virus load is downregulated to low levels, but the immune system provides strong selection for variation.

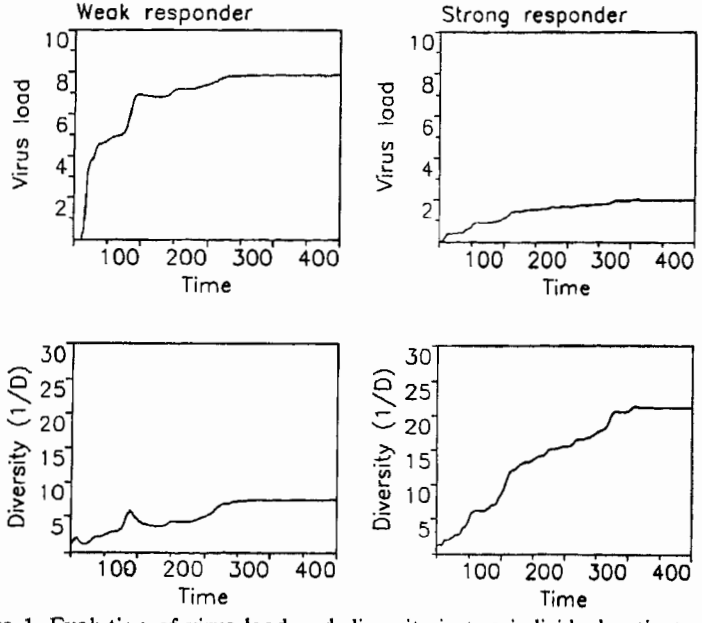

Figure 1. Evolution of virus load and diversity in two individual patients. One patient has a weak strain-specific immune response, while the other patient has a strong strain-specific immune response. In the weak responder, virus load raises quickly to high levels, but the antigenic diversity remains low. In the strong immune responder, virus load is maintained at low levels, while the immune response selects for high diversity. In both cases, antigenic diversity increases virus load. For the computer simulation, we use system (11) with r_i randomly chosen from a uniform distribution between 0 and 1, $k = 0.1$, $p = 1$, $q = 1$, $b = 1$, and $c_i = 0.1$ (for all strains i) for the weak responder and $c_i = 5$ for the strong responder. There is a constant probability over time to produce new antigenic variants, which increases the dimension n of the system. Virus diversity is given by the inverse of the Simpson index D which is defined as $D = \Sigma(v_i/v)^2$.

The result can be understood in terms of competitive exclusion. In the absence of strain-specific immunity, only the strain with the highest replication rate would win. Strain-specific immunity downregulates the abundance of the fastest replicating strains, and therefore allows other strains to persist as well.

In a cross-sectional comparison among different patients, do we expect a positive or a negative correlation between virus load and diversity? If individual patients differ in the strength of their specific immune response (parameters c_i and p) and in the intrinsic viral replication rates r_i, but have the same cross-reactive immune response (parameters k and q), then (14) and (15) predict a negative correlation between virus load and diversity. Patients with stronger strain-specific immune response (high c_i and p) select for more diversity (higher m, lower r_m) and lower virus load v (Fig. 2). Equation (15) also gives an inverse correlation between load and diversity if the only difference among patients is the replication rate r_i of individual virus strains.

If, on the other hand, k and q vary among patients while c_i and p are constant, then we must calculate the relation between kq and m before

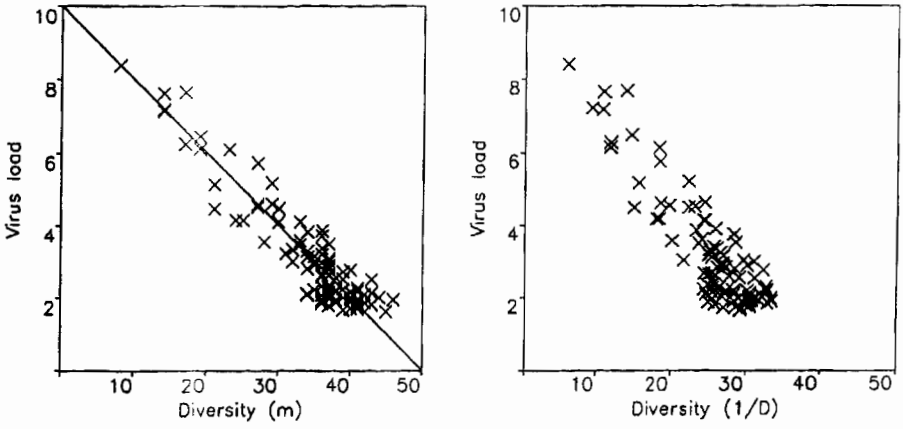

Figure 2. A cross-sectional comparison among different patients gives an inverse correlation between viral load and diversity if the patients differ in their strain-specific immune response. The figure shows equilibrium viral loads of system (11) as given by (13). Individual virus strains have replication rates r_i which are taken from a uniform random distribution on the interval (0, 1). All patients have the same cross-reactive response $k = 0.1$, but for different strain-specific responses c ranging from 0 to 1. (We assume that $c_i = c$ for all strains.) The other parameters are $p = q = b = 1$. There are $n = 50$ strains, and the figure shows 100 patients, each characterised by an x. The left side of the figure shows virus load v versus virus diversity in terms of numbers of strains m. The continuous line represents the approximation given by (15), with a continuous approximation for $r_m = \rho(m) = 1 - m/n$. The right-hand side shows virus load versus the inverse of the Simpson index as a measure for diversity. The Simpson index is defined as $D = \Sigma(v_i/v)^2$.

(15) can provide an answer. We get an elegant analytical approximation if we assume $c_i = c$ for all strains. (This simplification essentially means that all immune responses are induced at the same rate.) Then the index m is defined as the largest integer which fulfills

$$r_m > \sum_{i=1}^{m-1} r_i/(\phi + m) \tag{16}$$

where $\phi = cp/(kq)$ is essentially the ratio of strain-specific over cross-reactive immune responses. (We have also made the approximation that $m - 1 \approx m$.) If there is a very large number of strains, we can make a continuous approximation, $r_i = \rho(i)$, and the index m can be defined as the solution of the equation

$$\rho(m) = \frac{1}{\phi + m} \bar{\rho}(m) m \tag{17}$$

where $\bar{\rho}(m)$ is the average replication rate for all strains with index up to m. If we consider for the r_i a uniform random distribution in the interval $(0,R)$, then we obtain $\rho(m) = R(1 - m/n)$ and $\bar{\rho}(m) = R[1 - m/(2n)]$. Equation (15) leads to

$$m^2 + 2\phi m - 2\phi n = 0. \tag{18}$$

For a given ϕ and n, the number of strains at equilibrium is

$$m = -\phi + \sqrt{\phi^2 + 2n\phi}\,. \tag{19}$$

From (18), we also get $kq = 2cp(n - m)/m^2$, which can be combined with (15) to give

$$v = \frac{bR}{2cpn} m^2. \tag{20}$$

Hence, with the assumption of r_i following a uniform distribution, we obtain a direct correlation between the virus load and the square of virus diversity if individual patients differ in their cross-reactive responses against the virus (Fig. 3).

3. An HIV Model. For HIV, we expand the previous model by including a term for virus-induced impairment of the immune system. The basic model

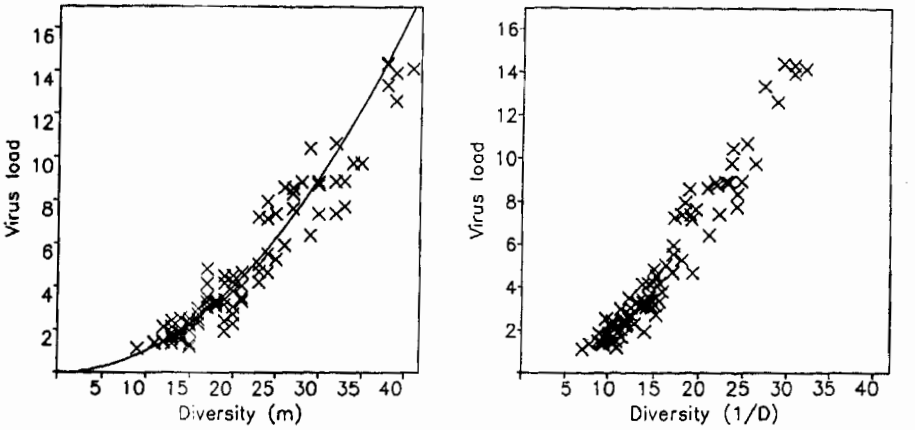

Figure 3. If patients differ in their cross-reactive response against the virus, the model predicts a positive correlation between virus load and diversity. Parameter values: $p = q = b = 1$, $n = 50$, 100 patients, r_i from a uniform distribution on $(0, 1)$, $c = 1$, and k from an exponential distribution with mean 0.2 (simply to get an equal density of points over the interval). The continuous line indicates the analytical approximation given by (20).

then becomes

$$\begin{aligned} dv_i/dt &= v_i(r_i - p_i x_i - q_i z) \qquad i = 1, \ldots, n \\ dx_i/dt &= c_i v_i - \left(b + \sum_{j=1}^{n} u_j v_j\right) x_i \qquad i = 1, \ldots, n \\ dz/dt &= \sum_{j=1}^{n} k_j v_j - \left(b + \sum_{j=1}^{n} u_j v_j\right) z. \end{aligned} \tag{21}$$

In this model, x_i and z denote immune responses which require CD4 cell help. HIV impairs such immune responses by reducing the CD4 cell number and function. This is described by the terms which contain $\Sigma u_j v_j$. Again, without loss of generality, we rank the strains such that $r_1/q_1 > r_2/q_2 > \cdots > r_n/q_n$. There are two possibilities. Either the total virus population size $v = \Sigma_i v_i$ grows to infinity or there is a unique stable equilibrium given by

$$v_i = \frac{b}{c_i p_i \Lambda}\left[r_i\left(1 + \sum_{j=1}^{m} \frac{k_j q_j}{c_j p_j}\right) - q_i \sum_{j=1}^{m} \frac{k_j r_j}{c_j p_j}\right] \tag{22}$$

$$x_i = \frac{1}{p_i}\left[r_i - q_i \sum_{j=1}^{m} \frac{k_j r_j}{c_j p_j} \Bigg/ \left(1 + \sum_{j=1}^{m} \frac{k_j q_j}{c_j p_j}\right)\right] \tag{23}$$

$$z = \sum_{j=1}^{m} \frac{k_j r_j}{c_j p_j} \Bigg/ \left(1 + \sum_{j=1}^{m} \frac{k_j q_j}{c_j p_j}\right) \tag{24}$$

with

$$\Lambda = \left(1 + \sum_{j=1}^{m} \frac{k_j q_j}{c_j p_j}\right)\left(1 - \sum_{j=1}^{m} \frac{r_j u_j}{c_j p_j}\right) + \sum_{j=1}^{m} \frac{k_j r_j}{c_j p_j} \sum_{j=1}^{m} \frac{q_j u_j}{c_j p_j}. \tag{25}$$

The number m is given by the largest index which fulfills

$$r_m/q_m > \sum_{j=1}^{m-1} \frac{k_j r_j}{c_j p_j} \Bigg/ \left(1 + \sum_{j=1}^{m-1} \frac{k_j q_j}{c_j p_j}\right). \tag{26}$$

For the total virus load at equilibrium, we obtain

$$v = \frac{b}{\Lambda}\left[\sum_{j=1}^{m} \frac{r_j}{c_j p_j}\left(1 + \sum_{j=1}^{m} \frac{k_j q_j}{c_j p_j}\right) - \sum_{j=1}^{m} \frac{q_j}{c_j p_j} \sum_{j=1}^{m} \frac{k_j r_j}{c_j p_j}\right]. \tag{27}$$

Whether this equilibrium exists or the virus population grows uncontrolled can be determined in the following way: check inequality (26) for increasing values of m. If the inequality holds for a particular value of m, then check if the corresponding virus load v given by (27) is positive. If it is positive, augment m by 1 and check (26) again. The algorithm ends either if an m is found such that $v < 0$, or if an m is found such that inequality (26) is violated. In the first case, the virus population grows to infinity; in the second case, there is a stable equilibrium given by (22)–(24). Note that v in (27) can only be negative if Λ is negative; hence, $\Lambda < 0$ is essentially a "diversity threshold" condition.

As a special case of (22)–(25), we get the complete equilibrium solution to the basic model of the previous section (1) for $u_i = 0$ for all i. Interestingly, this only affects Λ and v_i, but x_i, z, and v_i/v are not affected by immune function impairment. Note also that $\Lambda > 0$ if $u_i = 0$ for all i. Hence, the diversity threshold is a feature unique to viruses that impair the immune response. (But for other viruses, there is still a "diversity advantage" in the sense that increasing antigenic diversity increases virus load.)

As in Section 2, we will now consider a simplified model to gain some analytical insight into the relation between virus load and diversity in HIV

infections (see also Nowak and Bangham, 1996). Let us consider the special case, where all strains have the same parameters p, q, k, b and u, and only differ in their replication rates r_i and their rate of stimulating the strain-specific response c_i. This leads to

$$\begin{aligned} dv_i/dt &= v_i(r_i - px_i - qz) \qquad i = 1,\ldots,n \\ dx_i/dt &= c_i v_i - (b + uv)x_i \qquad i = 1,\ldots,n \\ dz/dt &= kv - (b + uv)z. \end{aligned} \tag{28}$$

The individual viral strains have the equilibrium frequencies

$$v_i = \frac{1}{c_i p}\left[r_i(b + uv) - kqv\right] \qquad i = 1,\ldots,m \tag{29}$$

and $v_i = 0$ for $i = m+1,\ldots,n$. The total virus load at equilibrium is given by

$$v = b\sum_{i=1}^{m}\frac{r_i}{c_i}\Bigg/\left(p + kq\sum_{i=1}^{m}\frac{1}{c_i} - u\sum_{i=1}^{m}\frac{r_i}{c_i}\right). \tag{30}$$

The number of strains m is defined as the largest integer which fulfills

$$r_m > kq\sum_{i=1}^{m-1}\frac{r_i}{c_i}\Bigg/\left(p + kq\sum_{i=1}^{m-1}\frac{1}{c_i}\right). \tag{31}$$

The denominator of (30) being positive is the diversity threshold condition. If

$$p + kq\sum_{i=1}^{m}\frac{1}{c_i} > u\sum_{i=1}^{m}\frac{r_i}{c_i}, \tag{32}$$

then the virus population converges to the finite equilibrium given above. If the reverse holds, then the immune responses cannot control the virus population. In mathematical terms, the virus population grows to infinity. The cross-reactive response converges to the value k/u, and each strain which fulfills $r_i > kq/u$ will grow.

From (29), we see that $c_i v_i$ declines with increasing i. For a large number m of strains, we can make the approximation $c_m v_m = 0$, which leads to

$$v = \frac{br_m}{kq - ur_m}. \tag{33}$$

This equation describes the relation between virus load v and antigenic diversity m in patients below the diversity threshold.

Let us first assume that patients differ in their strain-specific responses (parameters p and c), but not in their cross-reactive responses (parameters q and k). If the patients are able to control the virus (inequality (32) holds), then (33) describes an inverse correlation between viral load and diversity (Fig. 4). If the patients are unable to control the virus (inequality (32) does not hold), then the virus population tends to very high values, and the number of strains i which will be present is given by the inequality $r_i > kq/u$. But in terms of relative frequency, some of these strains may converge to zero. The minimum number of strains which are necessary to overcome the diversity threshold is given by the smallest interger m such that inequality (32) is violated. Exactly the same situation applies when the only difference among patients is the rate at which individual virus strains replicate.

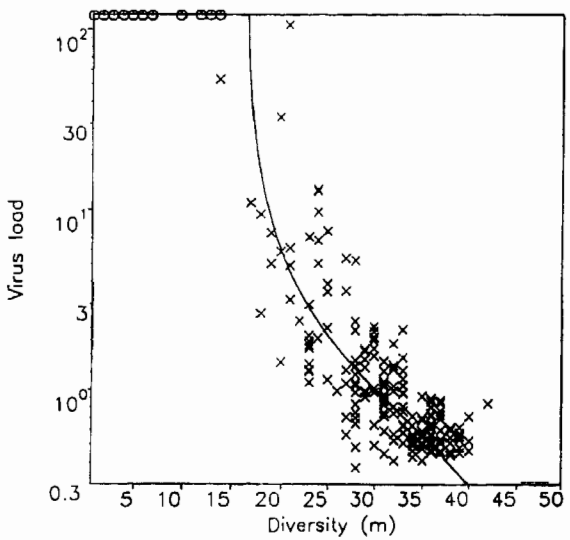

Figure 4. An inverse correlation between viral load and diversity for the HIV model as given by system (28). The patients differ in their ability to mount strain-specific immune responses against the individual virus strains. The individual r_i are taken from a uniform random distribution on the interval (0, 1). There are $n = 50$ strains of virus. Each patient has the same cross-reactive response $k = 1$, but the strain-specific responses vary from $c = 0$ to $c = 50$. Other parameters are $p = q = b = 1$ and $u = 1.5$. The figure shows 200 patients. The crosses indicate patients who can control the virus to a finite equilibrium load; the circles indicate patients who cannot control the virus (because their diversity threshold is overcome). The figure shows virus load versus diversity, as defined by the number m of strains present at equilibrium. For the patients who cannot control the virus, the figure gives the minimum number of strains which can overcome the threshold. The continuous line indicates the analytical approximation given by (33).

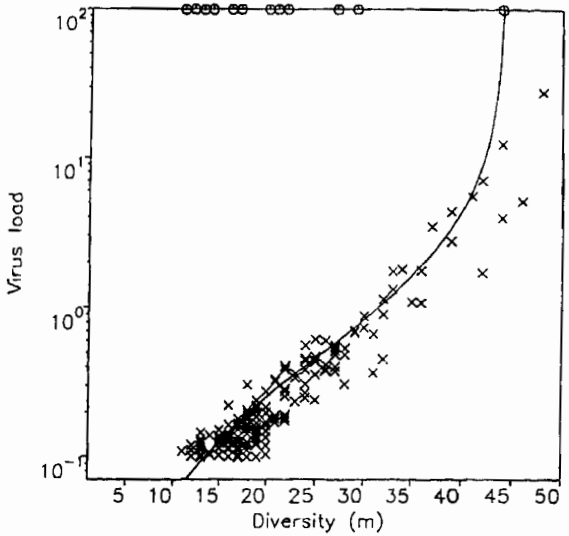

Figure 5. A more complex pattern of viral load and diversity emerges for the HIV model (28) if patients differ in their ability to mount cross-reactive immune responses against the virus. Among those patients who can control the virus, there is a direct correlation between load and diversity. But patients who cannot control the virus are usually already unable to control a very small number of strains. Parameters as Fig. 5, except $u = 1.02$, $n = 100$, $c = 10$, and k varies from 0 to 10 in different patients. The continuous line indicates the analytical approximation given by (34).

If, on the other hand, patients differ in their cross-reactive immune response against the virus (parameters q and k), but not in their strain-specific immune response, we find a positive correlation between virus load and diversity (Fig. 5). For a large number of strains, we can again make a continuous approximation for the replication rates $r_m = \rho(m)$. If the individual r_i are uniformly distributed on the interval $(0, R)$, then $\rho(m) = R(1 - m/n)$. As before, we find that the number of coexisting strains at equilibrium m is the root of the quadratic equation $m^2 + 2\phi m - 2\phi n = 0$ with $\phi = cp/(kq)$ and $c = c_i$ for all strains. Using this equation to eliminate kq in (33), we get

$$v = \frac{bRm^2}{2cpn - uRm^2}. \tag{34}$$

This describes a positive correlation between viral load and diversity.

4. Conclusions

1. In an individual infection, increasing the number m of viral variants (by increasing the total number of variants n) leads to increasing virus load. Antigenic diversity increases virus load.

2. The immune system can select for or against antigenic diversity. The amount of antigenic diversity which is selected in a given patient depends on the ratio of strain-specific to cross-reactive immune responses: the stronger the strain-specific component of the immune system, the more diversity; the stronger the cross-reactive component, the less diversity (19).
3. In cross-sectional comparisons among different patients, the correlation between virus load and antigenic diversity can be positive or negative. If patients differ in their strain-specific responses, then a weak responder allows a high virus load, but also provides little selection for variation, while a strong responder reduces virus load to low levels, but selects for high diversity. The resulting relation between load and diversity is negative. If patients differ in their cross-reactive responses against the virus, then the models predict a positive correlation between load and diversity: weak responders allow high virus load, and also provide low selection for antigenic variation.
4. Recent studies of HIV-1 infection have shown that rapid disease progression is often associated with weak immune responses, high virus load and low genetic diversity (Delwart *et al.*, 1994; Wolinsky *et al.*, 1996). Such a pattern is in agreement with the "antigenic diversity threshold" theory if patients differ mostly in their strain-specific immune responses (Nowak *et al.*, 1996).

Support from the Austrian "Forschungsförderungsfond," SERC and the Wellcome Trust is gratefully acknowledged.

REFERENCES

Baier, M., M. T. Dittmar, K. Cichutek and R. Kurth. 1980. Development in vivo of genetic variability of SIV. *Proc. Natl. Acad. Sci. USA* **88**, 8126–8130.

Balfe, P., P. Simmonds, C. A. Ludlam, J. O. Bishop and A. J. Leigh Brown. 1990. Concurrent evolution of human immunodeficiency virus type 1 in patients infected from the same source: rate of sequence change and low frequency of inactivating mutations. *J. Virol.* **64**, 6221–6233.

Bangham, C. R. M. 1993. Human T-cell leukaemia virus type 1 and neurological disease. *Curr. Opin. Neurobiol.* **3**, 773–778.

Bonhoeffer, S. and M. A. Nowak. 1994. Intra-host versus inter-host selection: viral strategies of immune function impairment. *Proc. Natl. Acad. Sci. USA* **91**, 8062–8066.

Bonhoeffer, S. and M. A. Nowak. 1995. Can live attenuated virus work as post-exposure treatment? *Immunol. Today* **3**, 131–135.

Borrow, P., H. Lewicki, X. Wei, M. S. Horwitz, N. Pfeffer, H. Meyers, J. A. Nelson, J. E. Gairin, B. H. Hahn, M. B. A. Oldstone and G. M. Shaw. 1997. Antiviral pressure exerted by HIV-1 specific CTLs during primary infection demonstrated by rapid selection of CTL escape virus. *Nature Med.* **3**, 205–211.

Burns, D. P. W. and R. C. Desrosiers. 1991. Selection of genetic variants of SIV in persistently infected rhesus monkeys. *J. Virol.* **65**, 1843–1854.

Clements, J. E., F. S. Pedersen, O. Narayan and W. A. Haseltine. 1980. Genomic changes associated with antigenic variation of visna virus during persistent infection. *Proc. Natl. Acad. Sci. USA* **77**, 4454–4458.

de Boer, R. J. and M. C. Boerlijst. 1994. Diversity and virulence thresholds in AIDS. *Proc. Natl. Acad. Sci. USA* **91**, 544–548.

de Boer, R. J. and C. A. B. Boucher. 1996. Anti-CD4 therapy for AIDS suggested by mathematical models. *Proc. Roy. Soc. London B* **263**, 899–905.

Delwart, E. L., H. W. Sheppard, B. D. Walker, J. Goudsmit and J. I. Mullins. 1994. Human immunodeficiency virus type 1 evolution in vivo tracked by DNA heteroduplex mobility assays. *J. Virol.* 68, 6672–6683.

Eigen, M. and P. Schuster. 1977. The hypercycle. A principle of natural self-organization. Part A: emergence of the hypercycle. *Naturwissenschaften* **64**, 541–565.

Ellis, T. M., G. E. Wilcox and W. F. Robinson. 1987. Antigenic variation of CAEV during persistent infection of goats. *J. Gen. Virol.* **68**, 3145–3152.

Ferrari, C., A. Bertoletti, F. Fiaccadori and F. V. Chisari. 1996. *Semin. Virol.* **7**, 23–30.

Frost, S. D. W. and A. R. McLean. 1994. Quasispecies dynamics and the emergence of drug resistance during zidovudine therapy of HIV infection. *AIDS* **8**, 323–332.

Goulder, P., R. Phillips, S. McAdam, B. Colbert, G. Ogg, P. Giangrande, G. Luzzi, B. Morgan, A. Edwards, A. McMichael and S. Rowland-Jones. 1996. Late escape from an immunodominant cytotoxic T lymphocyte response associated with progression to AIDS. *Nature Med.* **3**, 212–217.

Herz, A. V. M., S. Bonhoeffer, R. M. Anderson, R. M. May and M. A. Nowak. 1996. Viral dynamics in vivo: limitations on estimates of intracellular delay and virus decay. *Proc. Natl. Acad. Sci. USA* **93**, 7247–7251.

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**(6510), 123–126.

Hofbauer, J. and K. Sigmund. 1988. The theory of evolution and dynamical systems. Cambridge: Cambridge University Press.

Holmes, E. C., L. Q. Zhang, P. Simmonds, C. A. Ludlam and A. J. Leigh Brown. 1992. Convergent and divergent sequence evolution in the surface envelope glycoprotein of HIV-1 within a single infected patient. *Proc. Natl. Acad. Sci. USA* **89**, 4835–4839.

McLean, A. R. and M. A. Nowak. 1992a. Interactions between HIV and other pathogens. *J. Theor. Biol.* **155**, 69–86.

McLean, A. R. and M. A. Nowak. 1992b. Competition between zidovudine sensitive and resistance strains of HIV. *AIDS* **6**, 71–79.

Meier, U. C., P. Klenerman, P. Griffin, W. James, B. Koppe, B. Larder, A. McMichael and R. Phillips. 1995. Cytotoxic T lymphocyte lysis inhibited by viable HIV mutants. *Science* **270**, 1360–1362.

Mellors, J. W., C. R. Rinaldo, P. Gupta, R. M. White, J. A. Todd and L. A. Kingsley. 1996. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science* **272**, 1167–1170.

Moskophidis, D. and R. M. Zinkernagel. 1995. Immunobiology of cytotoxic T-cell escape mutants of lymphocytic choriomeningitis virus. *J. Virol.* **69**, 2187–2193.

Niewiesk, S., S. Daenke, C. E. Parker, G. Taylor, J. Weber, S. Nightingale and C. R. Bangham. 1995. Naturally occurring variants of human T-cell leukemia virus type 1 Tax protein impair its recognition by cytotoxic T lymphocytes and the transactivation function of Tax. *J. Virol.* **69**, 2649–2653.

Nowak, M. A., R. M. Anderson, M. C. Boerlijst, S. Bonhoeffer, R. M. May and A. J. McMichael. 1996. HIV-1 evolution and disease progression. *Science* **274**, 1008–1010.

Nowak, M. A., R. M. Anderson, A. R. McLean, T. F. W. Wolfs, J. Goudsmit and R. M. May. 1991. Antigenic diversity threshold and the development of AIDS. *Science* **254**, 963–969.

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

Nowak, M. A., S. Bonhoeffer, G. M. Shaw and R. M. May. 1997. Anti-viral drug treatment: dynamics of resistance in free virus and infected cell populations. *J. Theor. Biol.* **184**, 203–217.

Nowak, M. A. and R. M. May. 1992. Coexistence and competition in HIV infections. *J. Theor. Biol.* **159**, 329–342.

Nowak, M. A., R. M. May and R. M. Anderson. 1990. The evolutionary dynamics of HIV-1 quasispecies and the development of immunodeficiency disease. *AIDS* **4**, 1095–1103.

Nowak, M. A., R. M. May, R. E. Phillips, S. Rowland-Jones, D. G. Lalloo, S. McAdam, P. Klenerman, B. Köppe, K. Sigmund, C. R. M. Bangham and A. J. McMichael. 1995. Antigenic oscillations and shifting immunodominance in HIV-1 infections. *Nature* **375**, 606–611.

Overbaugh, J., L. M. Rudensey, M. D. Papenhausen, R. E. Benveniste and W. R. Morton. 1991. Variation in SIV env is confined to V1 and V4 during progression to simian AIDS. *J. Virol.* **65**, 7025–7031.

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 lifespan, and viral generation time. *Science* **271**, 1582–1585.

Phillips, R. E., S. Rowland-Jones and D. F. Nixon. 1991. HIV genetic variation that can escape cytotoxic T-cell recognition. *Nature* **354**, 453–459.

Pircher, H., D. Moskophidis, U. Rohrer, K. Burki, H. Hengartner and R. M. Zinkernagel. 1990. Viral escape by selection of cytotoxic T cell-resistant virus variants in vivo. *Nature* **346**, 629–633.

Price, D. A., P. J. R. Goulder, P. Klenerman, P. J. Easterbrook, M. Troop, C. R. M. Bangham and R. E. Phillips. 1997. Selection of HIV-1 cytotoxic T-lymphocyte escape variants during primary infection. *Proc. Nat. Acad. Sci. USA*, to appear.

Salinovich, O., S. L. Payne, R. C. Montelaro, K. A. Hussain, C. J. Issel and K. L. Schnorr. 1986. Rapid emergence of novel antigenic and genetic variants of EIAV during persistent infection. *J. Virol.* **57**, 71–80.

Schulz, T. F., D. Whitby, J. G. Hoad, T. Corrah, H. Whittle and R. A. Weiss. 1990. Biological and molecular variability of HIV-2 isolates from The Gambia. *J. Virol.* **64**, 5177–5182.

Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. E. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, M. S. Saag and G. M. Shaw. 1995. Viral dynamics in HIV-1 infection. *Nature* **373**, 117–122.

Wolinsky, S. M., B. T. M. Korber, A. U. Neumann, M. Daniels, K. J. Kunstman, A. J. Whetsell, M. R. Furtado, Y. Cao, D. D. Ho, J. T. Safrit and R. A. Koup. 1996. Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection. *Science* **272**, 537–542.

Received 10 September 1996
Revised version accepted 24 February 1997