The evolutionary dynamics of HIV-1 quasispecies and the development of immunodeficiency disease

Martin A. Nowak, Robert M. May and Roy M. Anderson*

This paper presents a theory to explain the development of immunodeficiency disease after a long and variable incubation period of infection with HIV-1. Two assumptions are central to the theory: (1) mutation via reverse transcription during viral replication can generate viral strains resistant to neutralization by antibodies specific to earlier mutants in a particular host; (2) the virus can kill the CD4-positive lymphocytes that play a role in mounting an immunological attack directed at the virus. The theory is examined via the development of a mathematical model which reveals that an increasing number of antigenically distinct viral strains may overwhelm the immune system of the host. As the viral diversity increases beyond a certain level the immune system is unable to suppress the population growth of all the strains simultaneously. The intuitive explanation of this pattern of model behaviour lies in the assumption that each virus can kill CD4-positive lymphocytes that are specific to any of the viral strains, but each lymphocyte only directs immunological attack against a single viral strain. The model captures several observed features of the interaction between HIV-1 and the human immune system: (1) an early peak in viraemia (primary HIV-1 infection) following infection; (2) a long and variable incubation period with low viral abundance for much of the period; (3) an increase of viral density in the final phase of infection as the failing immune system fails to control viral population growth (the appearance of the disease AIDS); (4) coevolution and coexistence of many viral mutants in one infected person, and (5) a positive correlation between the presence of high replicative viral strains and the rate of progression to disease (AIDS).

AIDS 1990, 4:1095–1103

Keywords: HIV, antigenic drift, immunodominant loop, quasispecies, diversity threshold.

Introduction

A characteristic of the immunodeficiency associated with the disease AIDS is the selective depletion of CD4-positive T-helper/inducer lymphocytes in the human immune system [1]. The manifestation is thought to be a direct result of infection by the aetiological agent of AIDS, HIV-1. The steady depletion of CD4 cells during the long and variable incubation period of AIDS (an average of 8–10 years in adults [2]) eventually results in an increased susceptibility to opportunistic infections and malignancies.

Despite intensive study, it is not yet clear how the virus induces this steady depletion of CD4-positive cells. The most obvious explanation is that of the direct cytopathic effects of the virus, but other factors have been suggested. These include syncytia formation, where uninfected cells fuse with infected cells via the interaction of the CD4 receptor and the envelope protein (gp120) expressed on the surface of infected cells, viral induction of autoimmune reactions where the binding of gp120 to uninfected cells generates a target such that autoimmune reactions induce cell death, and the binding of the viral envelope to CD4 molecules on the cell surface acting to disrupt the normal function of CD4-positive T-helper cells [3].

Primary HIV-1 infection is typically associated with viraemia, and, for a short but variable period (weeks to months) after infection, virus replication can be detected

From the Department of Zoology, University of Oxford, Oxford and the *Department of Pure and Applied Biology, Imperial College, London, UK.

Sponsorship: This work was supported by a research grant from the MRC (RMA and RMM), by an Austrian 'Erwin Schrödinger' scholarship (MAN), and by the Royal Society (RMM).

Requests for reprints to: Dr Martin Nowak, Department of Zoology, University of Oxford, Oxford OX1 3PS, UK.

Date of receipt: 21 May 1990; revised: 17 August 1990.

by virus isolation or by viral antigens in blood, or by both. Seroconversion follows, but thereafter virus isolation becomes difficult and viral antigens are often undetectable during the asymptomatic phase between primary HIV-1 infection and the occurrence of AIDS-related complex (ARC) or persistent generalized lymphadenopathy (PGL). As symptoms of disease develop, the ease of virus isolation increases and the fraction of infected cells in peripheral blood (estimated by isolation or the polymerase chain reaction) appears to be 100–1000-fold higher in AIDS patients than in asymptomatic individuals [4].

A striking feature of infection and the development of disease is the high genetic variability in virus isolates obtained either sequentially from the same infected patient or from different patients [5]. During the genetic metamorphosis of the RNA genome of HIV-1 into a DNA provirus upon cell infection, reverse transcription errors produced during DNA strand synthesis become fixed and new 'quasispecies' (populations of closely related but distinct viral genomes) of virus are created. The rate of nucleotide misincorporation is of the order of $> 10^{-4}$ per base per cycle, and for a genome of 10^4 bases this implies a reverse transcription error rate of more than 1 base per genome per metamorphic cycle [6].

Quasispecies are very sensitive to selection pressure such as that exerted by the human immune system (or by *in vitro* culture conditions in the laboratory). The existence of neutralizing antibodies specific to particular HIV-1 antigens has been well documented. The immunodominant loop of the virus, a region of about 30 amino acids within the envelope protein (gp120), appears to trigger neutralization phenomena in infected humans or chimpanzees [7,8]. However, this part of the envelope protein appears to be a highly variable region. For example, viral isolates derived from infected chimps soon after infection appear to be resistant to sera which were able to neutralize the viral isolate used to first infect the host [8]. The change of a single amino acid in gp120 can apparently account for such clonal restriction of neutralizing activity [9]. Infected individuals appear to harbour a quasispecies of the virus, with a broad distribution over the sequence space. Within this quasispecies many immunologically different mutants are found, so that sera from such patients generally neutralize a broad range of isolates. In general, longitudinal observations on patients from the point of infection suggest specific responses early on and a gradual broadening of the immune response during the long incubation period of AIDS.

The term 'quasispecies' was invented by Eigen in 1971 to indicate that virus populations usually never consist of a single wildtype sequence, but rather a whole distribution of different mutants (due to replication errors) [10–13].

A growing body of evidence reveals that the biological characteristics of the various quasispecies differ in such attributes as cell tropism and cytopathic properties [14]. Interestingly, viral replication properties appear to be associated with cytopathic effects (both with respect to direct killing and the ability to induce syncytia formation). Cells infected with slowly replicating viruses contain low levels of viral RNA and cells infected with highly replicating viruses contain high levels of viral RNA (suggesting that control of the rate of viral replication may lie at the transcription level [15]). More importantly, recent work suggests an association between the presence of quasispecies with low replicative ability and no or mild disease and, conversely, the presence of quasispecies with high replicative ability and advanced disease [16]. Furthermore, it has been demonstrated that a change from low to high replicative ability occurs in viral isolates obtained from patients during progression from an asymptomatic state to a state of immunodeficiency and disease [17]. Such changes have been interpreted as signs of increased viral virulence in the host (high replication being associated with severe cytopathicity) but it is unclear whether the described changes are a cause or a consequence of the deterioration of the immune system [18]. Some support for the notion that damage to the immune system (decreased number of CD4-positive cells) facilitates emergence of viral quasispecies with high rates of reproduction has been provided by Yoshiyama *et al.* [19]. To complicate the picture further, an increasing body of evidence points to the coexistence of both high and low replicating viral quasispecies [3]. More broadly, however, the mechanism of apparent change from low to high replicative ability during the progression from the asymptomatic to the symptomatic state during the incubation period of AIDS is unknown at present. In addition, the replicative ability of quasispecies isolated during primary HIV-1 infection, by comparison with those isolated in AIDS patients, is also unclear at present.

In this paper we describe the development and analysis of a simple mathematical model of the interaction of HIV-1 with the human immune system. Our aim is to assess whether interactions between the immune system and viral quasispecies can explain observed patterns of change in viral abundance and diversity, and thence to explain the observed clinical picture in relation to the development of immunodeficiency and disease in infected patients during the long and variable incubation period of AIDS.

The model is based on five key assumptions. These are as follows:

(1) the virus can kill CD4-positive T-helper cells (either directly or via the formation of giant syncytia);
(2) the continual evolution of new resistant viral mutants (via errors in viral replication) enables the total viral population (formed from the summation of all the quasispecies) to evade elimination by the immune system (selection by the immune system giving rise to new variants in the presence of neutralizing antibodies has been observed *in vitro* [20]);
(3) subpopulations of CD4-positive T-helper cells specific to a particular viral strain direct immunological attack against that strain;

(4) each mutant can kill all CD4 cells regardless of their specificity to a particular mutant;
(5) immunological responses to the virus are characterized by a specific response to individual strains and a non-specific general response that acts against all strains (by recognition of epitopes that are conserved among different strains).

Two key assumptions concern the ability of each mutant to infect and subsequently kill all CD4 cell types regardless of their specificity, and the specificity of the immunological response associated with a given cell type to a particular mutant. Current evidence, based on *in vitro* studies and the conserved nature of the gene that appears to control binding to and entry of the virus via the CD4 receptor, supports this view [3,4].

A central conclusion of our investigation is that the human immune system is only able to mount an effective immune response against a viral quasispecies whose diversity is below some threshold value. As the total population of viral quasispecies exceeds this 'diversity threshold' the immune system is liable to 'collapse', being unable to regulate viral replication and CD4 cell destruction.

The model

In this section we present a mathematical model that explores the quantitative consequences of the antigenic drift of HIV. The basic set of equations is

$$\dot{v}_i = v_i(r - sz - px_i) \qquad i = 1....n. \qquad (1)$$

Here v_i denotes the population size of virus strain i, and the superscript 'dot' denotes differentiation with respect to time. The replication rate r (including the whole cycle of infection, and assumed the same for all strains in this initial model) might be thought of as the difference between birth and death rates: $r = bQ - d$ where the birth term includes the probability Q that replication is done without error. We assume that the replication rate is independent of the number of host cells (i.e. the number of host cells in constant); d is the natural death rate of the virus (for example, due to shedding of the envelope protein).

The terms sz and px_i represent unspecific (= cross-reactive) and strain-specific immune reactions, respectively. These can be justified as follows. Let us assume that each mutant i induces the production of certain immune agents (CD4 cells), a fraction of which is specifically directed only against that particular mutant strain (for example, via the immunodominant loop), while the other fraction is directed against more conserved sites (for example the *pol* gene products or conserved regions within the *env* protein) and hence able to react with several different mutant strains. z is the number of immune agents directed against conserved regions and x_i is the number of cells directed specifically against a particular strain.

We can now establish the potential existence of a viral diversity threshold, independent of the details of the differential equations governing the detailed dynamics of x_i and z. The immune system will eventually control strain i if asymptotically $\dot{v}_i < 0$, which is to say if

$$r - sz - px_i < 0. \qquad (2)$$

The immune system can thus control each individual strain only if this inequality holds for all i ($i = 1, .., n$), which implies the restriction that

$$n < n_c(x, z) = \frac{px}{r - sz} \qquad (3)$$

Here $x = \Sigma x_i$. Hence there exists an upper limit, n_c of different strains that can be suppressed simultaneously by the immune system. In this very general discussion, n_c depends on the asymptotic sizes of the specialized and unspecialized T-cell population (x and z, respectively); as spelled out explicitly below, the limiting magnitude of these populations depends on detailed assumptions about how the virus and the immune system interact [see, for example, equation (10)].

Notice that the threshold phenomenon does not arise automatically, but rather depends on the relative magnitude of the parameters characterizing viral replication and immune responses. On the one hand, if the unspecific immune response is sufficiently strong in relation to viral replication rates — that is, if $sz > r$ — then this immune response is, by itself, able to suppress the growth of all viral strains. There will be a rise in viral abundance following the initial infection, but once the unspecific immune response has been mounted the initial strain and all subsequently evolved ones will be suppressed by this generalized response. On the other hand, if the initial viral strain has a replication rate so fast that it can overwhelm both the specific and the unspecific immune responses — that is, if $r > sz + px$ — then the immune response will not be able to cope with this initial infection. This situation corresponds, in effect, to the threshold viral diversity being less than one strain, $n_c < 1$. Between these two extremes — that is, when $sz + px > r > sz$ — lies the interesting region of dynamical behaviour, with its viral diversity threshold, that we now explore in more detail. This situation corresponds to individual viral strains having replication rates that can outrun the unspecific immune response, but not the combined effect of unspecific and specific immune responses (at least initially, when the specific immune response is not diluted by having to cope with an above-threshold number of different strains).

It is straightforward to show that, once the viral diversity exceeds the threshold value ($n > n_c$), then the total amount of virus ($v = \Sigma v_i$) will increase on average. This mathematical fact can be established by considering the product $P = v_1 v_2 .. v_n$ and noting that $P > 0$ if and only if $n > n_c$. For a given value of P the total amount of virus v thus has to be larger than $v_{min} = n(P)^{1/n}$. Therefore v

is bounded from below by v_{min} which increases monotonically.

More generally we might assume that n different mutant strains induce k different antibodies and that

$$\dot{v}_i = v_i(r - \sum_{j=1}^{k} p_{ij}x_j), \quad (4)$$

where p_{ij} is the effect of the jth antibody against the ith mutant. The same arguments as above now lead to a diversity threshold given by

$$n_c = \frac{1}{r}\sum_i\sum_j p_{ij}x_j. \quad (5)$$

Dynamics of the immune response: the simplest model

So far we have made no specific assumptions on the dynamics of the immune response. We now do this.

In the first, simplest model, we do not find a breakdown of immune protection (CD4 cell depletion) as viral abundance escapes control, because the model has the unrealistic feature that arbitrarily high viral concentrations can induce arbitrarily high rates of production of CD4 cells. The model is nonetheless an interesting starting point, because it shows how increasing viral density can eventually lead to uncontrolled viral population growth — beginning at some time that can be long after the original infection. In the following subsection we explore a more realistic model incorporating saturation in the rate of CD4 cell production. A central point in these models is that uncontrolled viral population growth causes the breakdown of the immune system, rather than the converse.

Specifically, suppose the immune cells x_i and z are produced at the rates kv_i and $k'v$ proportional to the density of antigens. The killing of immune cells by viral mechanisms is denoted by the terms uvx_i and uvz. We thus obtain

$$\dot{x}_i = kv_i - uvx_i \quad (6)$$

$$\dot{z} = k'v - uvz \quad (7)$$

$$\dot{v}_i = v_i(r - sz - px_i) + M(v) \quad (8)$$

The stochastic mutation term $M(v)$, representing the appearance of new viral strains, is designed in the following way: the probability that a new mutant is created in the time interval $(t,t + h)$ is given by $bQ'v(t)h$ (as $h \to 0$) where Q' is the probability that mutation yields an escape mutant. We neglect the possibility that mutation leads to mutants that are already present in the system, because the number of immunological different mutants appears to be very large [the combinatorial possibilities of the 19 variable amino acids in the immunodominant loop is 19^{20} and furthermore the shape of the loop can also be altered by mutations in other parts of the envelope protein (P. Nara, personal communication, 1989)].

Note that in this simplest model all viral mutants are assumed to have the same replicative capacity r, the same immunological parameters s, p, k, k' and the same cytopathic capacity u. The mutants differ only in their specific (immunodominant) antigen.

Note that, summing eq. (6) over all strains, we get

$$\dot{x} = v(k - ux). \quad (9)$$

Hence the total number of specific and unspecific immune cells converge monotonically towards the equilibrium value $\hat{x} = k/u$ and $\hat{z} = k'/u$, which implies that the diversity threshold is

$$n_c = \frac{pk}{ru - sk'} \quad (10)$$

Figures 1 and 2 show typical computer simulations of these equations. Initially we observe high levels of virus, but the immune response is slowly increasing and finally manages to suppress the most abundant strain. In the meantime, however, new mutants have been created. The mini-outbreaks of higher virus levels correspond to the occurrence of newly arisen neutralization-resistant mutants. During the course of infection the number of mutant strains continually increases and finally exceeds the threshold value n_c which results in the sudden and continuous rise of in viral abundance.

The unspecific immune response in our model is responsible for the fact that the initial strains are growing to higher levels than the following escape mutants. Roughly speaking, the higher the effect of the unspecific response the higher the difference between the initial peak and the mini-outbreaks in the silent phase. A stronger unspecific (cross-reactive) immune response is therefore also correlated with an increased incubation period (compare Figs 1 and 2).

The immune response against HIV (the sum $x + z$) increases steadily during the course of infection. This reveals an interesting feature: it is not the constant depletion of immune cells that finally allows the virus to escape immune control, but rather a highly diverse virus population manages to escape from an activated immune system. Beyond the threshold level of viral density, as the uncontrolled virus replicates to higher and higher levels, our model becomes unrealistic, because it has the feature that an arbitrarily high virus concentrations can induce an arbitrarily high rate of immune cell production. This oversimplification will be corrected in the next subsection.

If we include a population of immune cells, denoted by y, which is not induced by the virus (because this population, y, is specific against other antigens), but which is killed, we can observe a constant loss of the total number of immune cells $(x + z + y)$ during the course of infection. In mathematical terms we write

$$\dot{y} = -uvy$$

and obtain a monotone decrease in $y(t)$ if the initial density y_0 is larger than the final equilibrium value $\hat{x} = \hat{z}$.

In the final phase, when the threshold is exceeded and the virus 'escapes' control by the immune system, strains that earlier were suppressed can rise again. This happens as follows. After the immune system has suppressed a strain completely (v_i very close to zero) there is no further induction of specific immune cells directed towards that particular strain. But these immune cells are steadily removed by the other virus strains present in the system. After some time the frequency of x_i has dropped below the threshold ($x_i < r - sz/p$) and v_i becomes positive again. This point seems to be important for the accumulation of disease in the final phase of infection.

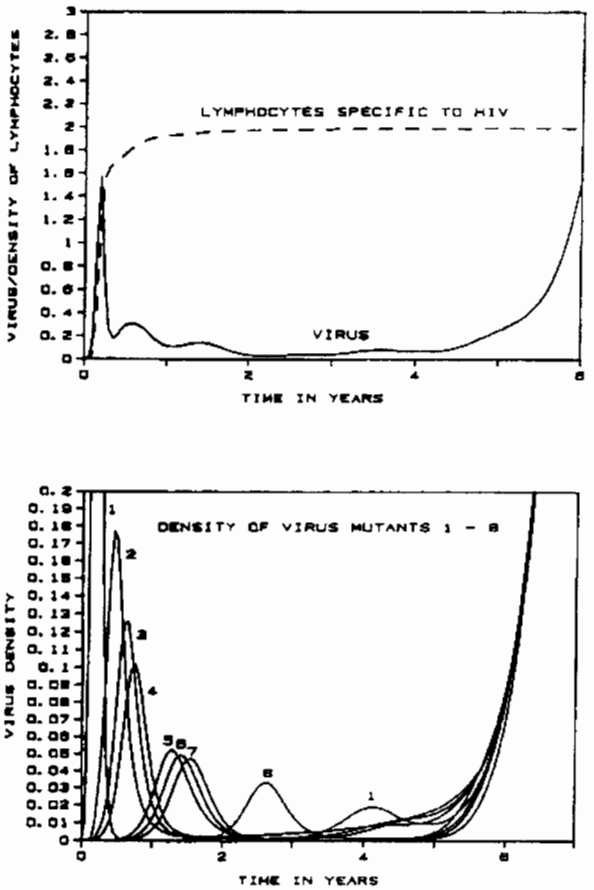

Fig. 1. Numerical simulation of HIV infection, as described by equations (6–8). (a) The solid curve shows the total virus concentration (in arbitrary units), v, and the dashed curve shows the concentration of lymphocytes specific to HIV (the immune response, $x + z$), also in arbitrary units. (b) The sequence of antigenic drift. The individual mutants are down-regulated by the immune response. But new mutants have been generated in the meanwhile. In the final phase we observe a simultaneous rise of all the strains present. In this figure, the parameters have the values: $r = 5$, $s = 4.5$, $p = 5$, $k = k' = u = 1$ and $bQ' = 2$, implying a diversity threshold $n_c = 10$.

Within our model it can also happen that the infection is cleared during the period of low virus concentration following the initial peak. This happens if the mutation rate to produce new resistant mutants is too low. On average we require that each virus strain before being suppressed by the immune system has to produce at least one new escape mutant. Figure 3 shows a simulation where the immune system manages to kill the virus population. This would explain some observations that once-infected patients have lost the virus and recovered completely. This would help explain the observation that some patients appear to be antibody positive but antigen negative (by the polymerase chain reaction technique) whilst others convert from antibody and antigen positive to antibody positive and antigen negative [21].

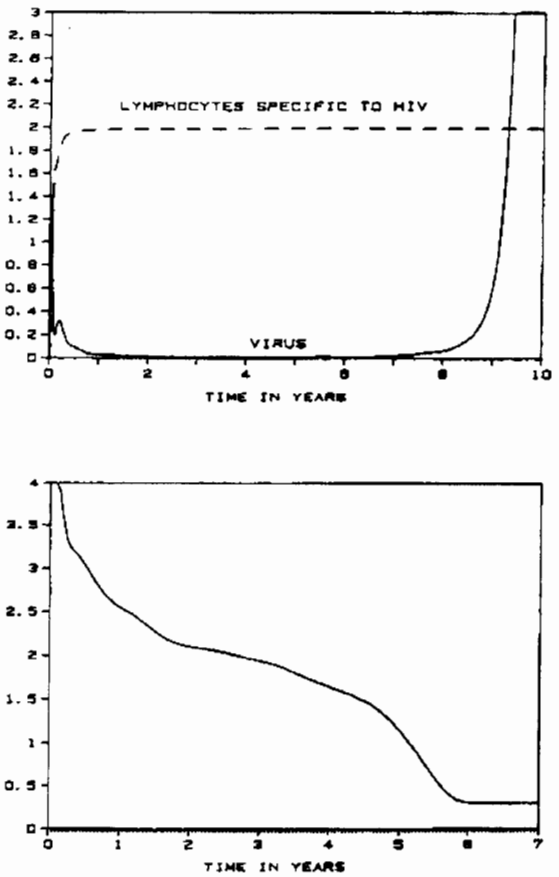

Fig. 2. Another simulation of equations (6–8) displaying an extremely long incubation period with low viral abundance (in comparison with the early peak); the units are as in Figure 1. (a) Viral concentration, v (the solid line) and HIV- specific lymphocyte concentration, $x + z$ (the dashed curve). (b) Total lymphocyte count, $x + y + z$. The parameter values here are: $r = 5$, $s = 4.75$, $p = 5$, $k = k' = u = 1$ and $bQ' = 3$, implying a diversity threshold $n_c = 20$.

Dynamics of the immune response: a more realistic model

To illustrate that the observed dynamics is not a special property of the simple equations for the immune reaction that we have just discussed, we now present a somewhat

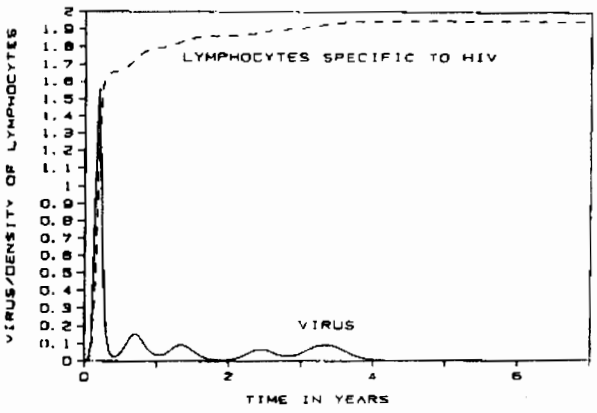

Fig. 3. The virus population is eventually eliminated by the immune response if the production rate of new resistant mutants is too low. In this example the parameters have the values: $r = 5$, $s = 4.5$, $p = 5$, $k = k' = u = 1$ and $bQ' = 1.75$. For this mutation rate the virus population fails to reach the diversity threshold, $n_c = 10$.

more realistic model. The basic structure of the model for a single virus strain is similar to that outlined by Anderson and May [22], but again the framework is broadened to consider the interaction of the immune system with a large number of coexisting viral mutants.

$$\dot{x}_i^0 = K - fx_i^0 - cv_i x_i^0 - uvx_i^0 \qquad (11)$$

$$\dot{x}_i = cv_i x_i^0 + ex_i - uvx_i \qquad (12)$$

Here x_i^0 denotes the density of precursors of immune cells ('inactivated cells') produced at the constant rate K and removed at rate fx_i^0. These precursors are either killed by the virus (rate uvx_i^0) or turned into activated cells by contact with the viral strain i (at rate $cv_i x_i^0$). Activated cells x_i can proliferate (ex_i) or can be killed by the virus (uvx_i).

The steady state for the precursors is given by $x_i^0 = K/(f + cv_i + uv)$. Assuming a steady state for the precursor cells, and including the previous subsection's assumptions about the unspecific cells, we obtain the following set of stochastic differential equations:

$$\dot{v}_i = v_i(r - sz - px_i) + M(v)$$

$$\dot{x}_i = \frac{cKv_i}{f + cv_i + uv} + ex_i - uvx_i$$

$$\dot{z} = \frac{c'Kv}{f + (c' + u)v} + ez - uvz$$

Figure 4 shows a typical simulation. The immune response does not equilibrate, but decreases during the final stage of the disease. The virus population dynamics shows similar behaviour.

In this model the immune cells have the possibility to proliferate indefinitely, if the virus level is low enough ($v < e/u$). For $v > e/u$ the immune cells tend towards the values

$$\hat{x} = \frac{cK}{uv - e} \sum_i \frac{v_i}{f + cv_i + uv}$$

$$\hat{z}\left(\frac{c'K}{uv - e}\right)\left(\frac{v}{d + (c' + u)v}\right)$$

Increasing diversity of the virus results finally in the breakdown of the immune system.

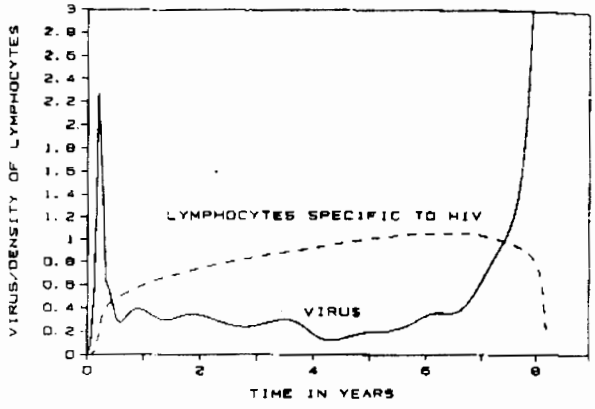

Fig. 4. A simulation of the more realistic model described in subsection 'Dynamics of the immune response: a more realistic model'. The dynamics of the virus population is similar to that in the simpler model, but the immune response breaks down in the final phase of infection.

Selection between strains of different virulence

It has been mentioned that fast-replicating strains are favoured in the final phase of the infection [3]. Within the framework of our model this is obvious.

Let us assume that there are two strains v_1 (fast replicating) and v_2 (slow). We have

$$\dot{v}_1 = v_1(r_1 - s_1 z - p_1 x_1)$$

$$\dot{v}_2 = v_2(r_2 - s_2 z - p_2 x_2)$$

Fast replication (r_1 large) is assumed to be correlated with high immunosuppression (p_1, s_1 large), while the slow-replicating strain is assumed to be a worse target for neutralization (p_2, s_2 low). This assumption is unsupported by empirical evidence at present but motivated by the argument that a slower replicating virus spends a larger fraction of its generation time inside the host cell hidden from neutralizing antibodies directed at the free

virus particle. This area requires further study, perhaps employing models in which replicative ability and antigenicity are uncorrelated.

In the initial phase of infection ($z = 0$, $x_1 = x_2 = 0$) strain 1 will grow faster (because $r_1 > r_2$).

In the long silent phase a well functioning immune system is likely to suppress v_1 more than v_2, such that $r_1 - s_1 z - p_1 x_1 < r_2 - s_2 - p_2 x_2$. During this phase strain 1 might be present in much lower frequency than strain 2. Therefore the immune system selects those strains that cause the long period with low viral activity.

In the final phase of infection, the breakdown of the immune response will again result in strain 1 replicating faster than strain 2 (if $r_1 > r_2$).

According to this argument, the mechanism underlying the observed change from slowly to rapidly replicating strains during the development of AIDS is simply selection. Support for this hypothesis has been reported from a haemophiliac AIDS patient with fast-replicating strains and his sexual asymptomatic partner with slow-replicating strains [19]. Our argument would also imply that in the very beginning of infection, when the immune system has not kicked in, fast replication should be of selective advantage.

The mathematical details of the 'non-neutral' model (different strains having different replication rates) are presented elsewhere (Nowak and May, in preparation).

The effect of chemotherapy or vaccination on the diversity threshold

The relationship

$$n_c = \frac{px}{r - sz}$$

enables us to estimate how changes in the parameter values can affect the maximum number of strains, n_c, that can be simultaneously controlled by the immune system. If a drug, in principle, acts by reducing the rate of viral reproduction, r, then the threshold value, n_c, increases in a hyperbolic manner, reaching infinity if $r = sz$. Since the rate of viral reproduction is itself a combination of variables in our model ($r = bQ - d$), the drug can act either by decreasing the birth rate, b, or the replication accuracy (of the reverse transcriptase) Q, or by increasing the natural death rate, d.

In our model vaccination or any kind of stimulation of the immune response (immunotherapy) has the following effect. Stimulating the specific immune response, by increasing p or x, yields only a linear increase in the diversity threshold, whereas stimulation of the unspecific (cross-reactive) response — by increasing s or z — would yield a fast hyperbolic increase of n_c.

It would therefore be very important to ensure that vaccination and/or immunotherapy acts to stimulate cross-reactive immune responses, rather than strain-specific ones.

Discussion

The models we have described in the preceding sections are obviously very simple caricatures of the true complexity of the interaction between coexisting HIV-1 quasispecies and the human immune system. However, they serve as a starting point for adding further biological realism as knowledge improves, both of the molecular and population genetics of viral replication and persistence, and of the factors that induce a steady decline in CD4 cells during the incubation period of AIDS.

The models make predictions which reflect a number of the observed features of HIV infection and AIDS disease. These are:

(1) a two-peaked pattern of viral abundance during the course of infection, with high total viral abundance in the early stages of infection (the primary HIV-1 infection stage) and a high and increasing viral load during the late stage of infection when the disease AIDS is manifest;
(2) a long and variable asymptomatic phase of infection with low viraemia;
(3) the coexistence of many (immunologically) different mutants throughout the incubation period of AIDS;
(4) a humped pattern in specific immunological responsiveness to viral antigens [specific CD4 cells that target the immunological response, via the stimulation of antibody attack directed against the envelope protein (gp120) of the virus] with low responsiveness during the asymptomatic stage and a decline in measurable responses as ARC and AIDS develop;
(5) the dominance of slowly replicating viruses during the asymptomatic phase of infection;
(6) the dominance of rapidly replicating quasispecies in ARC or AIDS patients.

Equally important, however, is the generation of testable hypotheses. These are as follows. First, the model suggests that the peak in viraemia during primary HIV-1 infection should be characterized by the dominance of quasispecies of the virus with high replicative ability. Second, the diversity within the immunodominant loop (V3) region of the envelope protein (gp120) should increase as a patient moves via the asymptomatic stage of infection to the development of AIDS (concomitant with an increased diversity of quasispecies of the virus). Third, during the asymptomatic phase of infection, small peaks in antigenaemia should be associated with the emergence of a quasispecies of moderate to high replicative capacity. Fourth, and finally, an infective dose containing a large number of quasispecies (perhaps via sexual contact with an AIDS or ARC patient) may be correlated with a shorter

than average incubation period prior to the development of symptoms of immunodeficiency.

The increase of viral diversity should be observed as long as the immune system is strong enough to drive the process of immunological selection. After the virus has overcome the diversity threshold, high virus concentrations kill the immune cells rapidly. In the absence of a relevant immune response in the final phase we need not necessarily observe a high viral diversity, because simply the fastest replicating strains are selected.

Current techniques should, in principle, allow the testing of these hypotheses. For example, Wain-Hobson [23] has reported diversity within the first hypervariable region (V1) of the envelope protein (gp120) on the basis of samples taken from four AIDS patients and two asymptomatic carriers. What is required is more intensive sampling, such that viral isolates can be tested at frequent intervals over the long and variable incubation period of the disease. Similarly, polymerase chain reaction techniques could be employed to assess the diversity of quasispecies both within and between viral isolates taken at sequential time points in the incubation period. It is hoped that the new technologies of molecular biology will encourage researchers to place greater emphasis on examining, in quantitative terms (where possible), the population genetics of HIV-1 populations, both within individual patients and among groups of infected persons (and in different geographical locations).

More generally, the model needs further development as well as testing. The central assumption is that selection by the immune system always acts most severely on the most abundant strain, thereby favouring rare quasispecies or new mutants until they themselves become abundant. The high heterogeneity of the system (many different cell types and many viral quasispecies) permits the coexistence of several competing mutants. Linked to this notion is the further assumption that the coexistence of several different quasispecies cause a dilution effect of the immune response, where each subset of specific CD4-positive T-helper cells is only directed against one quasispecies, but that each quasispecies is cytopathic to all subsets of CD4-positive cells. This assumption creates the 'diversity threshold' where the immune system can control limited diversity of viral types but is unable to constrain viral population growth when many quasispecies are present. This notion is speculative at present but the high mutation rate (created by errors in transcription) of HIV-1 will certainly facilitate the likelihood of ever-increasing genetic diversity in the viral population during the course of the incubation period [24] and the concomitant pressure on the immune system to suppress the growth of each quasispecies.

Finally, with respect to the search for vaccines to prevent infection or immunotherapeutic agents to reduce the likelihood of disease in infected persons, the model provides a template against which to test ideas concerning how to immunize against or treat for infections characterized by high genetic diversity in the infectious agent. Future work in model development will address the question of what is the effect on the course of infection and total viral abundance of immunizing or treating patients with an 'antigen cocktail' consisting of antigens from some subset of those preserved by the different quasispecies of the virus. The question of interest in this context is, given the high mutation rate of HIV-1 during the course of infection, how is diversity within the 'cocktail of antigens' and timing of treatment related to the length of time during which the total viral population can be suppressed to low levels (to avoid the appearance of disease)?

References

1. Garry R: Potential mechanisms of the cytopathic properties of HIV. *AIDS* 1989, 3:683–694.
2. Anderson RM: Mathematical and statistical studies of the epidemiology of HIV. *AIDS* 1988, 3:333–346.
3. Fenyo EM, Albert J, Asjo B: Replicative capacity, cytopathic effect and cell tropism of HIV. *AIDS* 1989, 3:S5–S12.
4. Schnittman SM, Psallidopoulos MC, Clifford-Lane H, *et al.*: The receiver for HIV-1 in human peripheral blood is a T cell that maintains expression of CD4. *Science* 1989, 245:305–308.
5. Saag MS, Hahn BH, Gibbons J, *et al.*: Extensive variation of human immunodeficiency virus type-1 *in vivo*. *Nature* 1988, 334:440–444.
6. Dougherty JP, Trentin HM: Determination of the rate of base-pair substitution and insertion mutations in retrovirus replication. *J Virol* 1988, 62:2817–2822.
7. Goudsmit J: Immunodominant B-cell epitopes of the HIV-1 envelope recognized by the infected and immunized hosts. *AIDS* 1988, 2 (suppl 1):S41–S45.
8. Nara P: AIDS viruses of animals and man. *Los Alamos Science* 1988, 18:55–89.
9. Looney DJ, Fisher AG, Putney SD, *et al.*: Type-restricted neutralization of molecular clones of human immunodeficiency virus. *Science* 1988, 241:357–359.
10. Eigen M: Self organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 1971, 58:465–523.
11. Eigen M, Schuster P: The hypercycle. *Naturwissenschaften* 1977, 64:541–565.
12. Eigen M, McCaskill JS, Schuster P: The molecular quasispecies. *Adv Chem Phys* 1989, 75:149.
13. Nowak M, Schuster P: Error thresholds of replication in finite populations. *J Theoret Biol* 1989, 137:375–395.
14. Saki K, Dewhurst S, Ma X, Volsky DJ: Differences in cytopathogenicity and host cell range among infectious molecular clones in human immunodeficiency virus type 1 simultaneously isolated strains an incidental. *J Virol* 1988, 62:4078–4085.
15. Peterlin BM, Luciw PA: Molecular biology of HIV. *AIDS* 1988, 2 (suppl 1):S28–S40.
16. Tersmette M, Grutess RA, de Wolf F, *et al.*: Evidence for a role of virulent human immunodeficiency virus (HIV) variants in the pathogenesis of acquired immunodeficiency syndrome: studies on sequential HIV isolates. *J Virol* 1989, 63:2118–2125.
17. Cheng-Mayer C, Seto D, Talene M, Levy JA: Biologic features of HIV-1 that correlate with variation in the host. *Science* 1988, 240:80–82.
18. Levy JA, Cheng-Mayer C, Talemo M, Evans LA, Humsy J: Biobreak logic heterogeneity of HIV and its relationship

to pathogenesis. *J Cell Biochem* 1989, suppl 13B [abstract GO1T].

19. YOSHIYAMA H, *ET AL*: Transmission and genetic shift of HIV *in vivo*. *Mol Biol Med* 1987, 4:385–396.
20. CANN AJ, KARN A: Molecular biology of HIV: new insight into the virus life-cycle. *AIDS* 1989, 3 (suppl 1):S19–S34.
21. HEWLETT IK, LAURIAN Y, EPSTEIN J, *ET AL*: Assessment by gene amplification and serological markers of transmission of HIV-1 from haemophiliacs to their sexual partners and secondarily to their children. *J Acquired Immune Deficiency Syndromes* 1990, 3:714–720.
22. ANDERSON RM, MAY RM: Complex dynamical behaviour in the interactions between HIV and the immunosystems. In *Cell to Cell Signalling* edited by Goldbetter A. New York: Academic Press, 1989, pp 335–349.
23. WAIN-HOBSON S: HIV genome variability *in vivo*. *AIDS* 1989, 3 (suppl 1):S13–S18.
24. NOWAK M: HIV mutation rate. *Nature* 1990, 347:522.