Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 6314–6319, May 1998
Immunology

Multiple modes of cellular activation and virus transmission in HIV infection: A role for chronically and latently infected cells in sustaining viral replication

ZVI GROSSMAN*†, MARK B. FEINBERG*‡, AND WILLIAM E. PAUL*§¶

*The Office of AIDS Research, National Institutes of Health, Bethesda, MD 20892; †Department of Physiology and Pharmacology, Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel; ‡Emory University School of Medicine, Atlanta, GA 30322; and §Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892

Contributed by William E. Paul, March 30, 1998

ABSTRACT **CD4$^+$ T cell activation, required for virus replication in these cells, occurs in local microenvironmental domains in transient bursts. Thus, although most HIV originates from short-lived virus-producing cells, it is unlikely that chronic infection is generally sustained in rapid continuous cycles of productive infection as has been proposed. Such continuity of productive infection cycles would depend on efficient long-range transmission of HIV from one set of domains to another, in turn requiring the maintenance of sufficiently high concentrations of cell-free virus across lymphoid tissues at all times. By contrast, long-lived cellular sources of HIV maintain the capacity to infect newly activated cells at close range despite the temporal and spatial discontinuities of activation events. Such proximal activation and transmission (PAT) involving chronically and latently infected cells may be responsible for sustained infection, particularly when viral loads are low. Once CD4 cells are productively infected through PAT, they can infect other activated cells in their immediate vicinity. Such events propagate locally but generally do not spread systemically, unlike in the acute phase of the infection, because of the early establishment of protective anergy. Importantly, antiretroviral drug treatment is likely to differentially impact long-range transmission and PAT.**

Antiretroviral drugs presently used to treat HIV-infected people do not block virus particle production in previously infected cells; rather, they act by preventing *de novo* infection. Therefore, when it was demonstrated that such drugs reduce plasma concentrations of HIV RNA by up to two orders of magnitude within 2 weeks, the implication was that the major HIV-producing cells are short-lived (1, 2). This has led to the premise that chronically and latently infected cells do not play a significant role in the dynamics of HIV infection. The events that sustain viral load have been depicted as rapid continuous cycles of productive infection and death of CD4$^+$ T cells.

These considerations focus on the overall turnover of virus but do not take into account microenvironmental aspects of immune activation that impose constraints on the interaction of the virus with its target cells. In physiological contexts, infected cells that themselves do not contribute substantially to the viral load may nevertheless play an essential role in sustaining the infection.

Proximal Activation and Transmission (PAT) Versus Long-Range Transmission (LRT)

Activated CD4$^+$ T cells are susceptible to infection with HIV. One determinant of such infection is the concentration of infectious virus in the locality of a susceptible cell. Such HIV could emerge from two classes of sources.

First, it could be produced at a distance by T cells that themselves had been recently infected. This mode, designated long-range transmission (LRT), may well be dominant when viral burdens are high, as we will discuss later. The emphasis on long-range and high burdens stems from consideration of the spatial and temporal pattern of the immune response. Activation of CD4$^+$ T cells, required for virus replication in these cells, is believed to occur in local microenvironmental domains in the form of transient bursts (3). Thus, the continuity of productive infection cycles involving infected T cells that have short life spans would depend on efficient LRT of HIV from one set of domains to another. This, in turn, would require the maintenance of sufficiently high concentrations of ambient virus across lymphoid tissues at all times. To stably sustain an infection, an ambient concentration of HIV should exist in extracellular fluid sufficient to infect a number of activated CD4$^+$ T cells equal to the number of infected cells that die in the same period of time.‖

Alternatively, the HIV responsible for infecting a CD4$^+$ T cell could have been produced locally, coupled with an immunologic interaction with an antigen-presenting cell (APC) that results in the activation of the T cell. If such immunologic interactions are highly localized and sporadic, the cell initially responsible for local HIV production typically would be long-lived, because such an infected cell is more likely to persist between subsequent activation episodes. It could be a chronically infected APC such as a macrophage or dendritic cell. The APC then would both activate and infect the antigen-specific CD4$^+$ T cell. Alternatively, the T cell may be activated by an uninfected APC in the presence of another cellular source of virus or the activation event could be followed closely by encounter with a source of HIV. The other source of virus could be a chronically infected cell, an HIV-binding follicular dendritic cell, or a latently infected CD4$^+$ T cell, activated specifically or as a bystander (4). Once new CD4$^+$ T cells are locally activated and infected, they can in turn infect other activated cells in their vicinity, amplifying the initial events. Some local infection events would result in new chronically or latently infected cells or in durable binding of HIV to new follicular dendritic cells, providing opportunities for further

Abbreviations: LRT, long-range transmission; PAT, proximal activation and transmission; APC, antigen-presenting cell; CTL, cytotoxic T lymphocyte; HAART, highly active antiretroviral treatment.

¶To whom reprint requests should be addressed at: Building 10, Room 11N311, National Institutes of Health, Bethesda, MD 20892. e-mail: wepaul@nih.gov.

‖This in turn requires that the basic reproduction ratio (BRR) of the infection should be larger than 1. BRR is a measure of infectivity; it is the number of cells each infected cell would infect when it is first introduced into a noninfected host. It is a function of the probability that a susceptible cell becomes infected and of the number of susceptible cells.

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.

cycles. In these cycles the longer-lived cellular sources of infectious virus serve as sparks, igniting transitory flames of HIV replication via a chain reaction of productive infection events.

This form of viral spread, designated proximal activation and transmission (PAT), is intimately tied to the process through which a $CD4^+$ T cell becomes susceptible to infection. The relevance of this mode is underscored by realizing that T cells activated in the presence of a cellular source of HIV are much more likely to become infected than are activated cells simply exposed to the ambient HIV in extracellular fluid (5). This would be particularly true when viral burdens are low. PAT is more effective than LRT because the proximity to the viral source provides higher local concentrations of particles and because the time contingency of activation and virus transmission may render the target cells more susceptible and the virus more infective. Indeed, cell-free HIV is believed to rapidly lose infectivity. Long-lived cellular sources of HIV provide the capacity to infect newly activated cells at close range despite the discontinuity of activation events.

This concept emphasizes the lack of spatial and temporal uniformity of virus production throughout the tissue. The continuity of productive infection is a systemic observation but need not exist at the local level. In the PAT mode, infection episodes are largely isolated from each other.

Wain-Hobson and his colleagues (6) have provided evidence in support of a spatial organization of HIV replication and T cell activation. They analyzed microdissected white pulps obtained from spleens of HIV-infected individuals who had undergone splenectomy. Each infected white pulp sample displayed unique HIV quasispecies, implying relative spatial isolation of infected foci. They concluded that "the distribution of HIV sequences within a white pulp was indicative of founder effects whereby one or a few proviruses became amplified locally." T cells found in individual white pulps from infected and uninfected spleens also manifested exquisite compartmentalization. When T cell receptors expressed by $CD4^+$ T cells were characterized, it was observed that individual white pulps displayed a unique pattern of V_β chain expression. Hence, "the singular compartmentalization of both HIV and T cells is indicative of little or no traffic between adjacent white pulps" (6).

These findings are consistent with PAT. Efficient LRT would wipe out the heterogeneity of HIV quasi-species at the microscopic level. Wain-Hobson and colleagues (6) proposed that the maintenance of virus replication in an individual white pulp depended on continuing recruitment of uninfected $CD4^+$ T lymphoblasts as a result of antigen activation. Our PAT concept agrees with this view of the isolation of individual foci of infection but modifies and extends it, emphasizing (*i*) the temporal discontinuity of local virus replication and (*ii*) cycles of regeneration of chronic and latent infection coupled to bursts of productive infection.

Control of HIV Infection by the Host: The Inadequacy of LRT-Based Models

The steep exponential increase of HIV concentration in blood after initial infection indicates that each infected cell initially can infect several susceptible cells, implying a large reproduction ratio (substantially greater than 1).** The rise in HIV burden during acute infection is followed by a rapid fall in the amount of virus, often of several orders of magnitude, and a near steady state is reached. At steady state, by definition, each infected cell that dies is replaced, on average, by one other infected cell.

Any theory of the dynamics of HIV infection during this chronic phase must take into account the following constraints: (*i*) virus load is typically much lower than in the acute phase, and (*ii*) initially there is only a modest decline in total CD4 cells, which may be largely caused by redistribution (7–9), and there is no substantial general disruption of immune responses to non-HIV antigens.

We discuss models of HIV infection that implicitly assume nominally efficient LRT to be the only relevant mode of transmission. We indicate the limited compliance of these models with the above constraints.

Model 1: Depletion of Susceptible Cells. The striking fall in viral burden at the end of the acute phase could be caused by the very efficiency of the process of cytopathic infection that would largely deplete the pool of infectable cells (10). The ensuing equilibrium would be characterized by a small number†† of virus-producing cells that in turn infect a small number of new target cells as soon as the latter are generated. Accordingly, even at low ambient viral concentrations, infection is highly efficient. If it were not, susceptible cells should accumulate once again.

This model complies with the first constraint (low virus load) but not with the second (undisrupted immunity). Because the thymus appears to contribute little to T cell replenishment in adults (11), most replacement of dying $CD4^+$ T cells must be through division of peripheral T cells. If the preparation for such division renders cells susceptible to infection and if infection is efficient, then replenishment of the peripheral $CD4^+$ T cell population should be strikingly impaired. In reality, however, $CD4^+$ T cell numbers often remain high for several years.

One way out of this contradiction is to assume that the state of cellular activation associated with self-renewal *in vivo* does not render $CD4^+$ T cells susceptible to infection. It has been reported that $CD45RA^+$ cells stimulated to proliferate by a mixture of cytokines preserved their CD45RA status and showed limited capacity for viral replication (12). But even if only immunologically activated cells are highly susceptible, infection and killing of such cells that is efficient enough to prevent their accumulation should severely impair immune function. Although immune functions may be impaired to some degree even at early stages, such impairment is generally not substantial. Indeed, even in patients with moderately advanced infection, CD4 responses to antigens associated with cytomegalovirus, for example, remain intact (13). Furthermore, impairments that are observed also might be explained by cellular anergy.

Model 2: Anergy. The finding that, in chronically infected individuals, the proportion of cells expressing memory/effector markers such as CD45RO or activation markers such as HLA-DR is greatly elevated, rather than reduced, but the number of infected cells is quite low (14), argues strongly against the depletion model.

We have postulated a mechanism that would reduce the susceptibility of the CD4 cell population to HIV infection without the substantial depletion of activated cells; namely, the induction in T cells by chronic stimulation of a reversible resistance to full activation, or anergy (8, 15). This in turn is based on the principle that lymphocytes that are chronically stimulated raise their activation thresholds and thus become insensitive to full activation by their cognate antigens (16–18).

The anergy model holds that many of the activated cells are actually resistant to infection. Indeed, when authentically activated T cells are generated in chronically infected individuals as a result of immunization, rapid increases in viral burden do occur, mimicking, although to a lesser extent, the type of increase seen upon acute HIV infection (19–21).

Our proposition is consistent with the functional defects seen in peripheral T cells from HIV-infected persons (22). Chronic stimulation and the induction of anergy is not restricted to HIV-specific T cells because diminished responsiveness, with a

**More precisely, the reproduction ratio in this case is a measure of the efficiency with which the virus both activates CD4 cells, generating more target cells, and infects them.

††With the simplest model, the steady-state number of susceptible cells would be reduced, compared with their number in the absence of infection, by a factor $1/R$, where R is the basic reproduction ratio.

possible impact also on the numbers and tissue distribution of T cells (15), is broadly manifested, encompassing the responses to recall antigens, alloantigens, and mitogens (22). Cytokines such as tumor necrosis factor, interleukins 1 or 6, or interferon α may chronically stimulate naive cells, produce a change in their surface phenotype, and induce a state of anergy within them. Indeed, this could explain the observation that the depletion in CD4 cells in the course of HIV infection is most notable among naive ($CD45RA^+$, $CD62L^{bright}$) cells despite the insensitivity of such cells to infection by HIV (12, 23–25). Viral products (e.g., gp120, TAT) also have been suggested to induce such changes in CD4 cells.

The systemic level of such anergy would be controlled by the degree of chronic stimulation which, in turn, should depend on the viral load. Moreover, because the level of anergy controls the number of infectable cells, these two parameters are dynamically coupled, and such coupling can readily lead to a steady state with a reproduction ratio of 1.

Thus, the fact that the virus does not expand in the lymphoid tissues of a chronically infected individual, as it does in the acute phase, may indicate that the activation of most lymphocytes is of the kind that renders them resistant, rather than susceptible, to *de novo* infection by HIV and/or to virus replication. By contrast, infection of properly activated cells occurs normally. Nonetheless, it seems unlikely that at steady state the infection of fully activated cells can be highly efficient (as in the acute phase) because if it were, virtually all nonanergic cells that recently have been activated should be destroyed. As we have argued before, this would disrupt the replenishment of the peripheral T cell population and the development of effective immune responses. Therefore, model 2 by itself is no less problematic with respect to the second constraint (undisturbed immunity) than model 1 even if, as we postulate, anergy is a major feature of chronic infection and anergic cells are resistant to viral replication.

Model 3: Conventional Immunological Control. Immune destruction of infected cells before their production of large amounts of virus or antibody-mediated neutralization of HIV also could account for the rapid fall of HIV from its peak in the acute phase. Indeed, this is the conventional explanation. Such a fall implies that the killing of potential HIV producers must be very efficient. If this mechanism is so efficient, why is it incapable of eliminating the virus completely? In contrast to models 1 and 2, where the level of virus and the number of infectable cells are interlocked in a predator–prey-like relationship, the link between the amount of virus and the activity of cytotoxic T lymphocytes (CTLs) is much less direct, allowing for a substantial overshoot in the numbers of CTLs. Thus, CTLs that could lower viral burden by several orders of magnitude should eliminate the virus altogether. Although examples have been reported in which CTLs appear to have eliminated HIVs that express sequences for which the CTLs were specific, this can be accounted for by a modest reduction in the fitness of such HIVs relative to competing species (Angela McLean, personal communication).

Invoking PAT

The concept of nominally efficient infection of activated cells by ambient HIV, stabilized by mechanisms that limit and adjust the levels of susceptible cells and/or infectious virus, is ridden with self-consistency problems. We invoke PAT partly because of these problems and mainly because we find it biologically more plausible, in light of what we know about spatially structured immune responses in the lymphoid tissues and given the direct evidence for highly localized interactions of HIV and T cells in these tissues (6). We propose that at least during part of the chronic stage, (*i*) LRT is genuinely inefficient, so that immune homeostasis and function remain largely intact; (*ii*) the infection is sustained via multiple local cell activation and virus transmission events; and (*iii*) active induction of cellular anergy, especially in HIV-specific T cells, is key to the stability of the level of infection.

The Acute Phase

We envision the events during the acute phase of HIV infection as two coupled chain reactions, activation and infection, driving each other. A small number of infected APCs and $CD4^+$ T cells initiate the process. The exponential increase in the number of activated lymphocytes is driven by increased antigen stimulation caused by the growing amounts of virus, by the development of activated APCs, and by production of viral products and growth stimulatory cytokines. The rapid increase in HIV levels, in turn, is driven by efficient infection of the increasing number of locally activated susceptible cells. Efficiency can be attributed in this case to the proximity of the cellular sources of infection to activated cells and to the proximity of coactivated and sequentially activated cells to each other. Thus, the large burst of virus production after the onset of infection is not necessarily caused by efficient LRT. As the systemic concentration of virus increases, LRT also may gain in efficiency so that at high viral burdens it may become the dominant mode.

Why does the process terminate after a few weeks, with a striking drop in virus levels (26)? Both CTLs (27–29) and the cytopathic effect of HIV (10), especially at the high concentrations reached at the peak of viremia, may contribute to the fall in viral burden, but we postulate that changes in the pattern of cellular activation play a key role in controlling the infection.

Acute immune responses are self-limiting even when the antigenic stimulus persists (30). We have proposed (8, 15) that this is caused by two built-in control processes: (*i*) a shift in balance between proliferation and differentiation toward the latter with subsequent death of the differentiated cells (17, 31); indeed, there is evidence for an early selective depletion of virus-specific $CD4^+$ T cells (32); (*ii*) elevation of the activation thresholds of recurrently stimulated lymphocytes that have failed to differentiate, resulting in a reversible induction of proliferative anergy (16, 17, 33). Consistent with the second mechanism are the substantial numbers of $HLA\text{-}DR^+$ T cells that are uninfected.

The Chronic Phase

If LRT is inefficient, then why does the infection persist once the acute phase has been controlled? We propose that this is because the lymphoid tissues have been seeded during the acute phase with chronically and latently infected cells[§§] that are relatively long-lived and site-specific and can spark new local bursts of infection in proximally activated target cells. Typically, activation is induced by antigens other than HIV; such local activation events occur normally in response to ongoing environmental stimuli. Because specifically and nonspecifically activated cells tend to cluster because of proliferation, recruitment, and bystander effects (34), a self-limiting chain reaction of efficient productive-infection events is envisioned to occur at the site of activation. Several such bursts can concurrently arise in different locations, collectively generating a viral load.

Such bursts would be the miniscale analogs of the acute phase. A large-scale recurrence of the acute-phase burst is hindered because of the chronic activation-associated anergy that already exists in a proportion of the $CD4^+$ lymphocytes

[§§]Here, and in the following, "chronically and latently infected cells" should be understood as including also "chronically infective" cells, even if the latter are not actually infected. We have in mind, in particular, follicular dendritic cells that are believed to be capable of binding infectious HIV for relatively long periods of time.

and possibly also because of the action of anti-HIV antibody and CTLs. These controls and the short life span of productively infected CD4 T cells would ensure that activation–infection events generally result in extinction of this major source of virus at the site of activation.

On the other hand, chronically and latently infected cells maintain the capacity of the microenvironment to infect newly activated CD4 T cells long after a burst of activation and infection has subsided. During each burst, some new macrophages and possibly dendritic cells become infected (35), replacing those that die. Some CD4 T cells may also become latently infected. Therefore, such bursts, when they occur, restore or increase the local levels of chronic and latent infection. These forms of infection provide the sparks that ignite new local bursts of virus replication. This scenario is in line with the observation that there is limited trafficking of infected cells between adjacent white pulps in HIV-infected spleens (6).

HIV-specific T cells maintain their anegry through recurrent interactions with HIV-presenting cells. Transient activation–infection episodes not only generate new virus but also reinforce the anergic state of HIV-specific T cells. Furthermore, anergy is widespread and apparently not limited to HIV-specific T cells. This is probably because cytokines produced by the latter and viral products induce cells of other specificities to enter and remain in an activation-resistant state. Indeed, there is now considerable evidence (36, 37) supporting the capacity of anergized cells to induce unresponsiveness in competent lymphocytes (16, 17). Relatively high levels of nonspecific stimulation (chronic bystander effect) may cause further elevation of the cellular activation thresholds, maintaining this state of anergy.

Though wide-spread, anergy is not global; many naive cells and memory cells that had not been anergized or that have recovered from their anergy would be activated by local triggering events. T cell responses depend broadly on the pattern of stimulation to which cells are exposed over time (16). In the tissues of HIV-infected individuals, anergy and full activation both are induced because the level and pattern of stimulation of individual cells is not uniform.

In the proposed scenario of miniscale local bursts initiated by chronically and latently infected cells, infection would be limited to cells that become activated close to a source of HIV. This explains the restricted impact of HIV infection on immune function. Only a small variable fraction of the activation events would be associated with PAT bursts, depending on the extent of chronic and latent infection.

The Dynamics of PAT Cycles

At steady state, the average duration of complete cycles of HIV replication in the PAT mode is essentially equal to the average turnover time of chronically and latently infected cells, rather than to that of productively infected $CD4^+$ T cells. This is also the turnover time of the sustained fraction of HIV, which is relevant in considering the evolution of new variants; virus that does not find its way into the DNA of chronically or latently infected cells is largely lost. Obviously, this turnover is much slower than if cohorts of productively infected short-lived cells sequentially and continuously infect each other, as in the commonly accepted paradigm. A typical cycle includes a short episode of intensive virus replication in coactivated or serially activated cells and a much longer period in which little or no replication occurs.

Unlike the LRT mode, PAT is not self-limiting (that is, controlled by running out of targets), but is rather rate-limited by an external factor, namely, the initiation of activation episodes. These events are largely external to the infection process, as they probably reflect primarily stimulation by antigens other than HIV. The frequency of PAT cycles, although it corresponds to the turnover rate of chronically and latently infected cells as we have indicated, does not reflect inherent life spans of these cells but rather the average frequency at which local infection episodes are initiated.

If PAT is not self limiting, then how does it reach a quasi steady state? We suggest that the relative stability of the viral load (and of other immunologic parameters) reflects the overall tendency of domains seeded by chronically and latently infected cells to recycle themselves, rather than expand, and the inefficiency of establishing new domains. We suggest that chronically and latently infected cells that have been activated produce large amounts of virus and have short life spans, as do productively infected CD4 cells. When a PAT episode occurs, some preexisting latently infected lymphocytes and chronically or latently (35) infected macrophages and dendritic cells are locally activated, intracellular virus replication is triggered or enhanced (35, 38), and most of these cells then probably die within a short time, to be replaced by new chronically and latently infected cells as the activation wave subsides. Within some radius from the center of a local activation episode, all or most preexisting chronically and latently infected cells would be activated and erased, either as a result of participating in cognate interactions or as affected bystanders. As the activation process subsides, incomplete cellular activation can presumably lead to the formation of some new chronically and latently infected cells at approximately the same site. A stable balance between the local regeneration and elimination rates of such cells in PAT cycles can be reached if the regeneration rate is relatively constant whereas elimination is proportional to the local density of these infected cells.

Although this scenario, which explains the overall relative stability of the viral load, is hypothetical, it is consistent with the observed robustness of founder effects at the microscopic scale in the lymphoid tissue (6).

Progression to AIDS

In addition to reinfection cycles at fixed domains, we envision flare-ups of infection at previously uninfected sites as a result of the migration of infected cells, most likely latently infected CD4 cells in the resting state. We propose that disease progression is associated with a spread of the tissue domain containing chronically and latently infected cells. In turn, such a spread will lead to an increase in overall burst frequency and may be accompanied by more extensive anergy and by qualitative changes of the tissue infrastructure. At a later stage, when the chronic reservoir has expanded beyond a certain limit, the concentration of extracellular virus may exceed the efficiency threshold for the LRT mode to become self-sustained. In this mode, the infection would have a substantial impact on the numbers and function of T cells, as we have discussed. LRT and PAT events always overlap. Because of this overlap and because the viral load is strictly controlled in LRT, to maintain an average reproduction ratio of 1, the increase in viral load is limited (within one or two orders of magnitude over several years).

Systemically, the viral load initially would reflect the amount of chronic and latent infection established in the acute phase and later would change only at a slow rate. Indeed, there is a correlation between the size and duration of the acute phase and the virus load (39).

Why is the spreading process slow? A partial explanation might be that establishing a new chronically infected domain is a rare event. It would depend on igniting bursts of infection in previously uninfected sites by ambient virus or by the traffic of infected cells and their subsequent activation, followed by the generation of a certain number of chronically and/or latently infected cells. LRT is typically inefficient; the concentration of ambient virus is generally too low. As for the migration of latently infected cells, because of their dispersion and the smallness of the pool (14), their activation at new sites

typically would occur as a single cell per site per event. Furthermore, in the context of a local response to a given antigen, such ad hoc activation of a migrating resting cell would typically be a nonspecific event, that is, a bystander effect. Single cells may have a small sparking efficiency, being eliminated in most cases without triggering a regenerative PAT episode. On the other hand, established chronically infected domains containing more than some critical number of chronically infected cells, HIV-binding follicular dendritic cells, and stationary latently infected cells (possibly with a partially activated phenotype) would have a higher sparking probability and could facilitate infected-cell replacement cycles. Once a domain with a sufficient cross-section is established, it will be relatively stable.

An alternative or complementary explanation is that potential founder cells, even when they trigger a local burst, have low probability for establishing chronic infection in new microenvironmental domains. On the other hand, the previous presence of anergic cells that can function also as suppressor cells (36, 37) in established chronically infected domains might increase the chance of arresting cellular activation and HIV transcription in newly infected cells during PAT events and thereby facilitate recycling of latently infected cells in such domains. The determinants of susceptibility to latent and chronic infection are presently unknown.

The postulated inefficiency of trafficking infected cells and ambient virus in establishing new chronically infected domains contrasts with the successful establishment of multiple such domains, in a relatively short time, during or after the acute phase of HIV infection. This difference might be attributed to the larger concentrations of virus achieved during acute infection, to the more extensive activation, or to qualitative differences in the profiles of cellular activation in the acute and chronic phases.

The relative stability of the number of latently infected CD4 cells, which is a correlate of slow progression, also can be considered from a systemic perspective. On the whole, the rate of latent infection of new cells is small. On average, however, latently infected T cells are long-lived and, therefore, *de novo* latent infection need only compensate, roughly, for the small rate of loss of such cells. Typically, in individuals defined as progressors, the rate of *de novo* infection would be larger than the death rate, but the net growth, the difference of two small numbers, would be small.

A number of observations support the hypothesis that progression to AIDS reflects a gradual transition from PAT to LRT. First, a paradoxical observation is that genetic diversity of HIV quasi species evolves faster during slow progression to AIDS, which is associated with lower viral burdens, than during rapid progression, in spite of the higher turnover of the virus in the second case (40). This has been attributed to more effective host-mediated selection pressures on replicating quasispecies in slow progressors (40). We can offer an alternative explanation. According to our hypothesis, low viral loads and slow progression may be associated with PAT as the dominant mode, whereas LRT becomes effective when viral loads are high. In the first case, the virus can evolve independently in multiple sites (6), leading to the accumulation of different quasispecies. In the second case, emerging species compete for predominance everywhere, resulting in more limited diversity.

Compatible with our hypothesis are also observations related to the inflection point that characterizes the long-term kinetics of T cell counts in HIV-infected persons (41). Eighteen to 24 months before the onset of AIDS, an unexplained fall in blood $CD8^+$ lymphocyte counts and in total T cell counts begins, replacing the pattern of earlier years of rising CD8 cell counts that roughly compensate for the fall in CD4 cell counts. In parallel, the rate of decline in $CD4^+$ T cell counts increases. The failure to maintain a relatively constant T cell count appears to be associated with accelerated progression to AIDS and has been termed homeostatic failure (41). Interestingly, viral loads typically maintain high but relatively fixed levels after the inflection point, and further diversification of the HIV quasispecies repertoire does not occur (Joseph Margolick, personal communication).

The inflection point itself and the other characteristics can be rationalized under the assumption of a switch from PAT to LRT. Such a transition would increase the impact on the immune system substantially because the maintenance of a reproduction ratio of 1 now requires a systemic reduction in the numbers of activated T cells, as we have discussed above for LRT-based models. In addition, LRT models, unlike PAT, readily predict a relatively fixed virus concentration even as the efficacy of infection increases, independently of the detail of these models (42). Finally, LRT would be associated with interspecies competition leading to the dominance of a few HIVs and a low rate of HIV-sequence diversification.

Differential Effect of Antiviral Drugs on LRT and on PAT?

Although the numbers of productively infected CD4 cells are dramatically reduced by available drug-combination treatments in patients with high viral loads, it is quite possible that infection continues to be actively sustained by the PAT mode (Z.G., M.B.F., V. Kuznetsov, D. S. Dimitrov & W.E.P., unpublished work). Antiretroviral drugs may reduce the rate of infection of activated CD4 lymphocytes by ambient virus below the self-maintenance threshold but may not reduce the efficiency of the local mode, involving relatively persistent cellular sources of virus and direct transmission from cell to cell, below this threshold. In particular, if the antiretroviral regimen is not sufficient to block all production of infectious HIV, the close-range infection occurring during cell–cell interaction and activation may be less inhibitable than is infection that depends on the ambient concentration of HIV in the extracellular space.

The large reduction of viral loads observed under highly active antiretroviral treatment (HAART) may reflect a transition from LRT to the PAT mode. Alternatively, if PAT was also the relevant mode before the initiation of treatment, the reduction may reflect a failure to efficiently amplify local cell-to-cell infection events.

Thus, the possibility must be considered that in some or all treated patients having undetectable plasma HIV RNA, a low level of infection is actively sustained. Virus production in such individuals may not be attributable solely to activation and virus production in cells that had been latently infected before institution of treatment or to residual active infection in isolated sites that are inaccessible to the drugs. Rather, virus also may be produced by recurrent rounds of productive infection, collectively replenishing, and depending upon, a small but stable pool of chronically and latently infected cells. Indeed, it has been reported that unintegrated viral DNA is present in CD4 cells of treated individuals, indicating ongoing viral replication (43), because such DNA is believed to be unstable. Even if HIV replication is very low in treated individuals, the infrastructure of chronically and latently infected cells may still be largely or partly intact. This may explain why, after cessation of treatment, HIV usually rebounds to pretreatment levels. Moreover, sustained low-level infection, with a reduction in systemic activation, could be associated with a risk of rebound to a level even higher than the baseline, at least transiently. This is because potential target cells may have tuned down their activation thresholds and thus escaped from a state of partial anergy in which they were relatively resistant to infection.

HAART can be extremely effective in reducing HIV burden but this does not imply that it does so by completely blocking viral transmission. We have argued (Z.G., M.B.F., V. Kuznetsov, D. S. Dimitrov & W.E.P., unpublished work) that data showing expo-

nential decay of virus in treated individuals are more consistent with variable degrees of inhibition than with complete blocking, but the death of the infected cells themselves is better represented by a distribution of time delays than by exponential decay kinetics. Furthermore, if as we propose HAART is more effective in blocking LRT of HIV than in inhibiting PAT and long-lived infected cells are responsible for ongoing cycles of infection in treated individuals, viral eradication may not be achieved by this therapeutic strategy alone. Either more effective drug therapy, capable of interrupting PAT, or interventions such as controlled activation of residual infected cells and/or induction of effective anti-HIV immunity under the cover of the reduction in viral replication induced by HAART, may be required if all infected cells are to be eliminated.

Conclusion

HIV may sustain itself in the body at different stages either (*i*) by continuously replicating in the $CD4^+$ T cell population or (*ii*) by maintaining the capacity of tissues to activate and infect $CD4^+$ cells. Chronically and latently infected cells may not contribute significantly to the viral load but may be instrumental in sustaining the infection. We further suggest that although the numbers of productively infected CD4 cells are dramatically reduced by available drug-combination treatments, it is quite possible that the infection continues to be actively sustained in the PAT mode.

Our propositions are based on several lines of evidence and are amenable to testing. Testable are the proposed functional profile of activated cells from infected lymphoid tissue or the frequency of chronically and latently infected cells in lymphoid tissue as an index of progression. The concept of PAT predicts that infected cells found in lymphoid tissue should manifest clustering, at the microscopic scale, into isolated foci or islands. Indeed, such clustering has been observed by *in situ* hybridization detection of infected cells (P. Greenberg, personal communication). The issue of active infection occurring in the presence of drugs is presently being addressed, indirectly, by increasing the sensitivity of viral load assays, by quantifying the persistence of integrated and unintegrated HIV in the lymphoid tissues, and by analyzing the character of virus in latently infected cells (43–45).

The prevailing focus on productive-infection cycles and the view that latently and chronically infected cells are important only as reservoirs represent an overly simplistic approach to HIV dynamics. The linkage between infection and activation at the cellular level should be better understood, taking into account the intricate nature of each of these processes. In particular, the determinants of susceptibility to latent and chronic infection and the homeostatic mechanisms that regulate the turnover of provirus-containing cells need to be defined.

We thank Tom Folks and Sandy Livnat for their critical review of this manuscript and for their useful suggestions.

1. Wei, X., Ghosh, S. K., Taylor, M. E., Johnson, V. A., Emini, E. A., Deutsch, P., Lifson, J. D., Bonhoeffer, S., Nowak, M. A., Hahn, B. H., *et al.* (1995) *Nature (London)* **373,** 117–122.
2. Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M. & Markowitz, M. (1995) *Nature (London)* **373,** 123–126.
3. Gulbranson-Judge, A. & MacLennan, I. (1996) *Eur. J. Immunol.* **26,** 1830–1837.
4. Tough, D. F., Borrow, P. & Sprent, J. (1996) *Science* **272,** 1947–1950.
5. Dimitrov, D. S., Willey, R. L., Sato, H., Chang, L. J., Blumenthal, R. & Martin, M. A. (1993) *J. Virol.* **67,** 2182–2190.
6. Cheynier, R., Henrichwark, S., Hadida, F., Pelletier, E., Oksenhendler, E., Autran, B. & Wain-Hobson, S. (1994) *Cell* **78,** 373–387.
7. Rosenberg, Y. J., Zack, P. M., White, B. D., Papermaster, S. F., Elkins, W. R., Eddy, G. A. & Lewis, M. G. (1993) *AIDS Res. Hum. Retroviruses* **9,** 639–646.
8. Grossman, Z., Bentwich, Z. & Herberman, R. B. (1993) *Clin. Immunol. Immunopathol.* **69,** 123–135.
9. Grossman, Z., Herberman, R. B, Vatnick, N. & Intrator, N. (1998) *J. AIDS Hum. Retrovirus* **17,** 450–457.
10. Phillips, A. N. (1996) *Science* **271,** 497–499.
11. Mackall, C. L., Fleisher, T. A., Brown, M. R., Andrich, M. P., Chen, C. C., Feuerstein, I. M., Horowitz, M. E., Magrath, I. T., Shad, A. T., Steinberg, S. M., *et al.* (1995) *N. Engl. J. Med.* **332,** 143–149.
12. Woods, T. C., Roberts, B. D., Butera, S. T. & Folks, T. M. (1997) *Blood* **89,** 1635–1641.
13. Waldrop, S. L., Pitcher, C. J., Peterson, P. M., Maino, V. & Picker. L. J. (1997) *J. Clin. Invest.* **99,** 1739–1750.
14. Chun, T. W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J. A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T. C., *et al.* (1997) *Nature (London)* **387,** 183–188.
15. Grossman, Z. & Herberman, R. B. (1997) *Nat. Med.* **3,** 486–490.
16. Grossman, Z. & Paul, W. E. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 10365–10369.
17. Grossman, Z. (1993) *Immunol. Rev.* **133,** 45–73.
18. Goodnow, C. C. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 2264–2271.
19. Staprans, S. I., Feinberg, M. B., Grant, R. M., Barbosa, P., Elbeik, T., Follansbee, S. E. & Hamilton, B. L. (1995) *J. Exp. Med.* **182,** 1727–1737.
20. Stanley, S. K., Ostrowski, M. A., Justement, J. S., Gantt, K., Hedayati, S., Mannix, M., Roche, K., Schwartzentruber, D. J., Fox, C. H. & Fauci, A. S. (1996) *N. Engl. J. Med.* **334,** 1222–1230.
21. Brichacek, B., Stevenson, M., Pirruccello, S., Janoff, E. N. & Swindells, S. (1996) *J. Infect. Dis.* **174,** 1191–1199.
22. Shearer, G. M. & Clerici, M. (1991) *AIDS* **5,** 245–251.
23. Zack, J. A., Arrigo, S. J., Weitsman, S. R., Go, A. S. & Chen, I. S. (1990) *Cell* **61,** 213–222.
24. Roederer, M., Raju, P. A., Mitra, D. K., Herzenberg, L. A. & Herzenberg, L. A. (1997) *J. Clin. Invest.* **99,** 1555–1564.
25. Spina, C. A., Prince, H. E. & Richman, D. D. (1997) *J. Clin. Invest.* **99,** 1774–1785.
26. Koup, R. A., Safrit, J. T., Cao, Y., McLeod, G., Borkowsky, W., Farthing, C. & Ho, D. D. (1994) *J. Virol.* **68,** 4650–4655.
27. Safrit, J. T. & Koup. R. A. (1995) *Curr. Opin. Immunol.* **7,** 456–465.
28. Borrow, P., Lewicki, H., Wei, X., Horwitz, M. S., Peffer, N., Meyers, H., Nelson, J. A., Gairin, J. E., Hahn, B. H., Oldstone, M. B. A., *et al.* (1997) *Nat. Med.* **3,** 205–211.
29. Fauci. A. S. (1996) *Nature (London)* **384,** 529–534.
30. Moskophidis, D., Lechner, F., Pircher, H. & Zinkernagel, R. M. (1993) *Nature (London)* **362,** 758–761.
31. Grossman, Z. (1982) *Eur. J. Immunol.* **12,** 747–756.
32. Rosenberg, E. S., Billingsley, J. M., Caliendo, A. M., Boswell, S. L., Sax, P. E., Kalams, S. A. & Walker, B. D. (1997) *Science* **278,** 1447–1450.
33. Grossman, Z. & Singer, A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 14747–14752.
34. Mitchison, N. A. (1995) *Immunologist* **3,** 5–11.
35. Orenstein, J. M., Fox, C. & Wahl, S. M. (1997) *Science* **276,** 1857–1861.
36. Qin, S., Cobbold, S. P., Pope, H., Elliott, J., Kioussis, D., Davies, J. & Waldmann, H. (1993) *Science* **259,** 974–977.
37. *Immunol. Rev.*, Vol. **149**.
38. Folks, T. M., Justement, J., Kinter, A., Dinarello, C. A. & Fauci, A. S. (1987) *Science* **238,** 800–802.
39. Nowak, M. A., Lloyd, A. L., Vasquez, G. M., Witrout, T. A., Wahl, L. M., Bischofberger, N., Williams, J., Kinter, A., Fauci, A. S., Hirsch, V. M., *et al.* (1997) *J. Virol.* **71,** 7518–7525.
40. Delwart, E. L., Pan, H., Sheppard, H. W., Wolpert, D., Neumann, A. U., Korber, B. & Mullins, J. I. (1997) *J. Virol.* **71,** 7498–7508.
41. Margolick, J. B., Munoz, A., Donnenberg, A. D., Park, L. P., Galai, N., Giorgi, J. V., O'Gorman, M. R. & Ferbas, J. (1995) *Nat. Med.* **1**, 674–680.
42. Bonhoeffer, S., Coffin, J. M. & Nowak, M. A. (1997) *J. Virol.* **71,** 3275–3278.
43. 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.
44. 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.
45. 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.