The state of latency occurs when a microbe's persistence in a host produces host damage without perturbing homeostasis sufficiently to cause clinical symptoms or disease. The mechanisms contributing to latency are diverse and depend on the nature of both the microbe and the host. Latency has advantages for both host and microbe. The host avoids progressive damage caused by interaction with the microbe that may translate into disease, and the microbe secures a stable niche in which to survive. Latency is clinically important because some latent microbes can be transmitted to other hosts, and it is associated with a risk for recrudescent microbial growth and development of disease. In addition, it can predispose the host to other diseases, such as malignancies. Hence, latency is a temporally unstable state with an eventual outcome that mainly depends on host immunity. Latency is an integral part of the pathogenic strategies of microbes that require human (and/or mammalian) hosts, including herpesviruses, retroviruses, Mycobacterium tuberculosis, and Toxoplasma gondii. However, latency is also an outcome of infection with environmental organisms such as Cryptococcus neoformans, which require no host in their replicative cycles. For most microbes that achieve latency, there is a need for a better understanding and more investigation of host and microbial mechanisms that result in this state.

Find the latest version:
https://jci.me/136221/pdf
The state of latency in microbial pathogenesis

Liise-anne Pirofski1 and Arturo Casadevall2

1Division of Infectious Diseases, Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, New York, USA. 2Department of Molecular Microbiology and Immunology, Johns Hopkins School of Public Health, Baltimore, Maryland, USA.

The state of latency occurs when a microbe’s persistence in a host produces host damage without perturbing homeostasis sufficiently to cause clinical symptoms or disease. The mechanisms contributing to latency are diverse and depend on the nature of both the microbe and the host. Latency has advantages for both host and microbe. The host avoids progressive damage caused by interaction with the microbe that may translate into disease, and the microbe secures a stable niche in which to survive. Latency is clinically important because some latent microbes can be transmitted to other hosts, and it is associated with a risk for recrudescence of microbial growth and development of disease. In addition, it can predispose the host to other diseases, such as malignancies. Hence, latency is a temporally unstable state with an eventual outcome that mainly depends on host immunity. Latency is an integral part of the pathogenic strategies of microbes that require human (and/or mammalian) hosts, including herpesviruses, retroviruses, Mycobacterium tuberculosis, and Toxoplasma gondii. However, latency is also an outcome of infection with environmental organisms such as Cryptococcus neoformans, which require no host in their replicative cycles. For most microbes that achieve latency, there is a need for a better understanding and more investigation of host and microbial mechanisms that result in this state.

This article discusses the complex state of latency in microbial pathogenesis using the damage-response framework (DRF) as a conceptual guide. The DRF puts forth an integrated theory of microbial pathogenesis and infectious diseases that incorporates the role of the host, as well as the microbe, to define the outcome of host-microbe relationships. The DRF was first proposed in 1999 as a solution to the late-20th-century problem of explaining the increase in human disease with microbes that were thought to be nonpathogenic or very rare causes of disease in those with intact immunity (1–3). These microbes often resided on the skin and/or mucosal surfaces of normal people, e.g., staphylococci, enterococci, enteric Gram-negatives, and Candida spp., and included those that were previously characterized as “opportunists” because they were nearly always observed as causes of disease in patients with immune impairment, e.g., Pneumocystis, Cryptococcus, Histoplasma, and Toxoplasma spp. During this time, there was also a recrudescence of adult disease caused by viruses with the capacity for latency, e.g., Epstein-Barr virus, cytomegalovirus, and other herpesviruses. There were three main factors driving the emergence of these microbes: (a) breach of barriers as reflected by the increased use of plastic catheters and advances in surgery; (b) exogenous immunosuppression as a consequence of advances in the treatment/management of malignancy, inflammatory diseases, and organ transplantation with cytotoxic and/or immunosuppressive agents; and (c) endogenous immunosuppression brought on by the AIDS pandemic. Collectively, these factors led to an unprecedented number of individuals with acquired immune impairments ranging from loss of barrier immunity to deficits in humoral and/or cellular immunity.

The DRF provided a tractable and integrated theory that clinicians and researchers could use to explain, predict, and propose new approaches for the therapy and study of the pathogenesis of common, emerging, and reemerging microbes (3). The basic tenet of the DRF is that the relevant outcome of host-microbe interaction is host damage, which is a function of the host response and can be mediated by host or microbial factors or both. As illustrated in Figure 1, the relationship between host damage and the host immune response is depicted by a simple, upright parabola, whereby damage is plotted on the y axis and the host response, ranging from weak to strong, is plotted on the x axis. The threshold for disease, a clinical state, is represented by a horizontal line across the y axis. This schema makes it possible to explain several important principles, including that (a) microbial infection may result in damage without causing clinical disease and (b) host damage and disease occur in the setting of either an insufficient or an exuberant immune response to a microbe. While the latter concept was novel when the DRF was first proposed, today it is common to attribute clinical disease to host inflammation, even in the setting of microbial clearance. In fact, we attribute this change in emphasis to the success of the DRF in reframing questions of microbial pathogenesis.

A major contribution of the DRF to clinical practice and research was its ability to account for different outcomes of infection with a single microbe without invoking confusing and inconsistent terminology, e.g., pathogen, nonpathogen, opportunist, primary pathogen, colonizer, commensal (2, 4). To describe the state of host-microbe interaction, the DRF puts forth five states of microbial existence in a host: (a) infection, (b) colonization, (c) commensalism, (d) disease, and (e) latency (also referred to as chronicity and/or persistence) (2). According to the DRF, there are no pathogens, commensals, etc.; there are only microbes and hosts whose interaction results in one of these states. The DRF proposes...
The concept of latency is useful for clinicians and basic scientists alike. For clinicians, understanding the state of a microbe in a host can inform an understanding of disease pathogenesis and prognosis, as well as the use of antimicrobial agents, making this an evidence-based decision, as well as decisions on implementation of infection control measures (5). For researchers, accounting for the state of the microbe in a host makes it possible to devise more precise questions to unravel pathogenesis and virulence.

**What is latency?**

Clinically, at the bench, and in experimental models, the state of latency is one of quiescence, whereby there are no clinical symptoms in the host and there is no evidence of microbial growth, despite the continued presence of the microbe in the host. For microbes, the state of latency may be due to a genetic change that blocks replication entirely, a slow rate of replication that may not be detectable with available tools, or sequestration that escapes detection. None of these strategies produce a degree of damage that translates into clinical disease. Generally, the state of latency is invoked for microbes formerly associated with known acquisition and/or clinical illness. As such, there is no clinical evidence of the microbe, but it remains present in the host and the risk of recrudescence disease looms. Recrudescence disease is often characterized as “reactivation.” Traditional examples of latent microbes include *Mycobacterium tuberculosis* and herpesviruses. The state of latency contrasts with microbial states that feature either acquisition without clinical symptoms or disease, e.g., colonization or commensalism, or acquisition followed by continued microbial growth resulting in symptoms, disease, elimination, or death of the host.

Classical, functional views of latency, whereby there is no clinical disease and no evidence of microbial growth, represent the tip...
of a complex iceberg. *M. tuberculosis* is a microbe that is universally accepted as capable of latency. On the other hand, based on its clinical course, human immunodeficiency virus (HIV) was originally believed to enter a state of latency after an initial illness that was followed by a long asymptomatic state. However, advances in understanding the biology of the virus and its interaction with the host revealed that instead of being quiescent, the virus continues to replicate during the asymptomatic period between infection and AIDS. This period of containment by the host immune system resulted in a clinically latent state that lasted until the immune system had sustained a degree of damage that could not prevent the emergence of clinical disease, often from an intercurrent infection. Another way in which HIV latency is characterized is as the state in which reservoirs of virus may remain after administration of agents to induce a cure (6).

**Latency in the context of the DRF**

The state of latency is referred to as a state of persistence to convey that certain latent microbes continue to replicate in, and/or are recovered as “infectious” particles from, their hosts (7–9). Thus, clinical latency can occur in the setting of either persistence, a state characterized by ongoing microbial (usually viral) replication, albeit at a reduced rate; or latency, a state characterized by microbial presence without ongoing replication. While persistence best describes the state of HIV and other microbes in hosts that exhibit clinical latency despite ongoing viral replication, microbial persistence and latency each result in an outcome of host-microbial interaction that does not feature clinical signs or symptoms. From the viewpoint of the DRF, persistence and latency are comparable states, with the caveat that persistent and latent microbes have different biological strategies for survival and for immune evasion and may evade different host immune responses. As above, latency is not associated with a degree of damage that perturbs homeostasis or translates into clinical disease. While this concept may appear simplistic, it provides the flexibility that is necessary to incorporate and explain emerging evidence that microbes not previously thought to exhibit latency can exhibit periods of clinical quiescence. One example of this phenomenon is Ebola virus, which may recrudesce after recovery from disease (10); another is the growing recognition that adenoviruses may persist and reactivate in immunosuppressed hosts (9).

One variable that affects the state of latency is time. For certain infectious diseases, the term “latency period” is used to describe the time between microbial acquisition and clinically evident disease. Historically, recognition of this period made it possible to identify that the yellow fever virus was transmitted to humans by mosquitoes following an extrinsic latency period (11). This usage of the term “latency period” is often used synonymously with “incubation period.” Although “latency period” does not represent classical latency, it is important to account for this period, because it delineates a time during which a microbe can be transmitted from person to person while the transmitting individual is asymptomatic. This period has been an intense focus of the investigation of transmission of the 2019 novel Coronavirus (12, 13). Challenges to our understanding of this period have increased since the 2010s, when it was discovered that Ebola virus disease could recrudesce (14, 15) and that Zika virus could persist in semen months after initial acquisition (16, 17). A common thread that links Ebola and Zika to prolonged existence in a host is the presence of host reservoirs that contain and sequester these microbes. While these reservoirs remain incompletely understood, from a clinical standpoint, the presence of these viruses in an asymptomatic host blurs the line between microbial persistence, latency, and clinical disease. From the standpoint of microbial pathogenesis, the state of a “quiescent” microbe in a host is crucially important for epidemiology, diagnosis, and prevention strategies, as well as approaches to therapy and vaccine design.

**Why is latency important?**

Latency provides a mechanism for a microbe to establish an ongoing or persistent niche in a host. For the microbe and the host, latency offers a relatively safe reservoir. It allows the microbe to travel with the host with the possibility of reaching new populations. The presence of a latent microbe signals the potential for reactivation and future disease. This may or may not occur. In humans, death of the host with a latent microbe generally signals the death of the microbe. However, in feral hosts, death can enable the microbe to find new hosts if a carcass is eaten or returns to the soil when the host dies and decomposes, provided the microbe remains viable. In this regard, it has been suggested that *Coccidioides immitis* reached South America after being carried by migrating humans (18), possibly in a latent state. For the individual carrying a latent microbe, there is the potential for damage that translates into another disease. For example, lung inflammation from latent *M. tuberculosis* can progress to scar formation and subsequent carcinoma. In experimental *Cryptococcus neoformans* infection in rats, control of pneumonia is associated with persistence of fungal cells in granulomas that alters the pulmonary immune response to promote airway hyperresponsiveness (19). Cryptococcal antigens may also trigger asthma in children (20). These examples show that establishment of the latent state can alter immune and inflammatory responses in the organs involved and predispose the host to very different diseases.

**Why do certain microbes elicit latency?**

Not all microbes exhibit the capacity for latency. In this respect, microbes with prolonged incubation periods between infection and disease (clinical latency) differ from microbes with classically defined persistence or latency. The former microbes are clinically quiescent during the period of clinical latency because the microbial burden and/or immune response does not elicit sufficient damage to affect homeostasis and translate into clinical disease. In contrast, microbes that exhibit “classical” latency or persistence infect the host and enter this state without necessarily causing clinical symptoms or progressing to disease. The microbial mechanisms and immune evasion strategies that lead to the state of latency are complex and varied. Some microbes capable of establishing the classical state of latency, such as *M. tuberculosis* and *C. neoformans*, enter a state characterized by a host response that features persistence of the microbe within tissue granulomas. The initial state of infection with these microbes may not be associated with a state of disease. Other microbes, such as herpesviruses, assume a state whereby their genomes are maintained in host cells and viral progeny are not produced, but possess the potential for reactivated viral
First, there is persistence in granulomas; this is exhibited by the microbe to vary, there are three general microbial persistence strategies. For survival, such as persistence in the environment, which ensure their ability to live with limited resources. microbes that require human hosts to survive employ different mechanisms, similar to those they might employ in animal cells also reflect this strategy. Persistence of microbes that can live outside of a human host, there may be complete dependence on the host for survival, and latency is part of the microbe’s replicative and survival strategies. Third, there is microbial persistence in protected sites whereby immunity does not eradicate microbes; this is exemplified by Zika virus and Ebola virus persistence, although to date, there is no evidence that these microbes exhibit classical latency (see above).

Examples of latency in phylogenetically distant microbes

The series on Latency in Infectious Disease in the *JCI* features Reviews on five microbes capable of establishing latency: human immunodeficiency virus (HIV) (50); herpes simplex virus (51); *Mycobacterium tuberculosis*; *Toxoplasma gondii* (52); and *Cryp
tococcus neoformans* (53). In considering the topic of latency in microbial pathogenesis, one can identify similarities among these very different types of microbes, with the caveat that each host-microbe relationship that results in microbial latency is unique with respect to microbial biology, host immunity, and how latency is established and maintained. These five microbes illustrate these distinctions, as each can persist in a host it has entered without causing clinical symptoms, yet the outcome and details of their pathogenic processes are very different. Although aspects of latency with each of these microbes are briefly mentioned above, this section considers their similarities and differences through the lens of the DRF.

Viruses

HIV and herpes simplex (HSV) are viruses that establish lifelong relationships with their hosts, during which the host can experience extended periods without clinical disease. However, these viruses differ in their lifestyles, host locations in which they reside, and mechanisms of virulence and host damage. Both viruses are acquired by infection of cutaneous or mucocutaneous surfaces, including by sexual transmission, although HIV can be acquired by inoculation into the blood (e.g., by injection, and/or historically by transfusion). HIV initially causes a systemic process characterized by viral replication in immune cells that may be marked by a clinical syndrome with fever, adenopathy, rash, and other manifestations. This initial response is followed by a quiescent clinical period that can last for years, and this period has been referred to as latency. If untreated, HIV continues to replicate in host cells at a rate that depends on host immune mechanisms until it results in progressive damage to the immune system, severe immunosuppression characterized by CD4+ T cell depletion, and a cascade of immunological deficits that also affect antibody and innate immunity. Hence, this long asymptomatic period is not a stable state.
Instead, it inexorably progresses to disease and is thus different from the latency states described for other microbes, for which infection can be stable for the life of the host. The immune impairment stemming from the ability of HIV to persist and continue to replicate leads to recrudescence or reactivation of classically latent microbes, such as C. neoformans and T. gondii, which the host normally controls with immunological mechanisms that are destroyed by HIV-mediated damage. HSV is like HIV in that it may be acquired by sexual transmission and may reactivate in the setting of immune suppression. However, unlike HIV, which resides in CD4+ T (and potentially other immune) cells, latent HSV resides in neurons, commonly affecting the trigeminal (HSV-1) and sacral (HSV-2) ganglia. During latency, HSV persists in these cells by downregulating host immune molecules to avoid triggering a virus-clearing immune response. Hence, in latency, HSV-mediated damage is limited to infected cells, whereas reactivation leads to viral replication and axonal transport to mucocutaneous sites, where a lytic cellular infection is manifested as blisters, erosions, and painful lesions. Both HIV and HSV are transmissible during latent and disease states.

**Bacterium**

*M. tuberculosis* is a bacterium that is acquired by inhalation of aerosolized organisms emitted by a host with disease (tuberculosis). Pulmonary infection, which follows inhalation, is then followed by either clearance, latency, or disease. If the initial host immune response fails to clear the mycobacterial inoculum, it is contained in the lungs by a granulomatous tissue response that can be asymptomatic and lead to a latent state in which the microbe persists in granulomas. The size of the granuloma varies in different individuals and can range from microscopic to visible by chest radiography. The establishment and maintenance of *M. tuberculosis* latency results in cellular, tissue, and organ damage, with the extent of damage varying depending on the individual host and the nature of the host-microbe interaction, including such variables as inoculum, host genetics, host nutrition, and location of infection in the lung. The persistence of viable organisms in granulomatous lesions presents a future danger to the host from reactivation of mycobacteria growth and progression to disease. Potential mechanisms of reactivation include interference with host defense by drugs that affect the immune response, such as steroids and immunomodulatory biologics; coinfection with agents that affect immunity, such as HIV; and aging, as there is a lifetime risk of *M. tuberculosis* reactivation, which increases with time. The persistence of inflammation in the lung can also result in malignant transformation and the development of scar carcinomas. *M. tuberculosis* is not transmissible during latency. Transmission requires a level of bacillary burden that occurs in disease, which is marked by mycobacterial proliferation in lung tissue, tissue destruction, and the formation of infectious aerosols during expectoration and coughing.

**Parasite**

*T. gondii* is an apicomplexan parasite that can infect numerous hosts but requires cats to complete its replicative cycle. Human infection is acquired by ingestion of *T. gondii*-infected animal tissue, which results in gastrointestinal infection followed by widespread dissemination to organs and skeletal muscle. The process of infection and dissemination is often asymptomatic, with the latent state being established as the parasite infects individual cells in a complex cellular biological process that involves the formation of a specific parasitic vacuole in which *T. gondii* persists until the tissues are ingested by another host. If host immunity is impaired, the parasite can reactivate and replicate in tissue, creating a necrotic lesion. Before the current era of antiretroviral therapy, patients with HIV infection and advanced immunosuppression (AIDS) often developed cerebral toxoplasmosis, which appeared as ring-enhancing lesions in CT scans. Hence, the state of latency of *T. gondii* is associated with cellular damage to parasite-infected cells. In addition, there is some evidence that high tissue parasitic loads can impair organ function, as suggested by the association of infection with some psychiatric disorders.

**Fungus**

*C. neoformans* is a soil fungus that is the causative agent of cryptococcosis, a life-threatening disease that usually presents as meningitis or meningoencephalitis. Unlike HIV, HSV, *M. tuberculosis*, and *T. gondii*, *C. neoformans* has no need for an animal host in any part of its replicative cycle. As such, its ability to cause human disease appears to represent an instance of accidental virulence (54), whereby traits honed for environmental survival allow it to survive in humans. Despite their large phylogenetic distance from one another, the pathogenesis of *C. neoformans* has remarkable similarities to that of *M. tuberculosis*. Both begin with an initial pulmonary infection that either is cleared from the lung or becomes latent in a granuloma (for *C. neoformans*, a cryptococcoma; for *M. tuberculosis*, a tuberculoma). Cryptococcomas are sites of granulomatous inflammation that are associated with damage to lung tissue. Like *M. tuberculosis*, *C. neoformans* is a facultative intracellular microbe that replicates in macrophages, in which it produces cellular damage. Fungal cells can remain viable in lung tissue for years and reactivate to cause progressive replication and disease years after initial infection. Most individuals with latent *C. neoformans* infection are entirely asymptomatic. However, as with *M. tuberculosis*, herpesviruses, and *T. gondii*, the latent state of *C. neoformans* can progress to disease when there is a perturbation of host immunity that impairs cellular, humoral, and/or innate immune mechanisms used for microbial containment.

In summary, for each of the microbes discussed, the state of latency is associated with persistent host immune system (cellular and/or tissue) damage that does not affect homeostasis in a way that results in clinical symptoms or disease. Of the five microbial examples, HIV differs from the others in that the quiescent asymptomatic period that is characterized as a period of latency is accompanied by progressive viral replication and inexorable damage to the immune system that inevitably leads to disease and death if untreated. In contrast, the most common outcome of latent HSV, *M. tuberculosis*, *T. gondii*, and *C. neoformans* in immunologically normal individuals is a continued asymptomatic period of clinical latency that endures unless this state of host-microbe interaction is disrupted by factors that impair host immunity or other homeostatic mechanisms. Persons with immune impairment due to HIV or the use of cytotoxic or lymphocytic drugs and/or biologics for malignancy, inflammatory (rheumatological) disease, or solid organ transplantation are at risk for reactivation of *M. tubercu-
losis, T. gondii, and C. neoformans. The common denominator in each of these host-microbe relationships is that the symptomatic period associated with latency is unstable and fraught with potential for future danger. On the other hand, latency may last for the life of the host. Historically the latent state has received relatively little attention compared with the state of disease. This probably reflects the challenges involved in studying the state of latency, which is usually clinically quiescent, stable, and difficult to study without perturbing. Currently there is a paucity of animal models to investigate latency, and in vitro systems are unlikely to fully recapitulate the in vivo state. Future studies should develop such models with the goal of further defining the host and microbial interactions that contribute to persistence, stability with the potential for instability, and, in some cases, instability that characterize latent microbes. This may reveal new strategies to control the progression from latency to disease and develop vaccines for microbes with the potential for latency. However, accomplishing this will probably require the development of new tools and techniques to study the state of latency without perturbing it. Given the importance of latency in microbial pathogenesis, a renewed focus on this state should be a research priority, for it is likely to provide new actionable insights into the host-microbe relationship that could be used to benefit human health.

Address correspondence to: Liise-anne Pirofski, Division of Infectious Diseases, Albert Einstein College of Medicine and Montefiore Medical Center, 1300 Morris Park Avenue, Bronx, New York 10461, USA. Phone: 718.430.2940; Email: lpirofski@einsteinmed.org. Or to: Arturo Casadevall, Department of Molecular Microbiology and Immunology, Johns Hopkins School of Public Health, 615 N. Wolfe Street, Room E5132, Baltimore, Maryland, USA. Phone: 410-955-3457; Email: acasade@jh.edu.


