W. David Wick · Otto O. Yang

# War in the Body

## The Evolutionary Arms Race Between HIV and the Human Immune System and the Implications for Vaccines

Springer

# War in the Body

W. David Wick • Otto O. Yang

# War in the Body

## The Evolutionary Arms Race Between HIV and the Human Immune System and the Implications for Vaccines

Springer

W. David Wick
Seattle, Washington
USA

Otto O. Yang
Geffen School of Medicine,
Department of Medicine
UCLA Medical Center
Los Angeles, California, USA

ISBN 978-1-4614-7293-3 ISBN 978-1-4614-7294-0 (eBook)
DOI 10.1007/978-1-4614-7294-0
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013935710

© Springer Science+Business Media New York 2013
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

# Preface

It has often been said that nothing in biology is comprehensible except in the light of evolution.[1] HIV/AIDS provides the latest—and most sobering—proof of this adage.

Evolution in the microbial world burst into scientific and public-health consciousness 60 years ago, with the appearance of antibiotic-resistant bacteria. In this well-known episode, the bugs developed resistance to penicillin in a matter of months after it was introduced (around 1947) in the clinic; some bacterial strains are now resistant to every antibiotic we possess. If you become infected by one, you might as well be living in the 19th century. HIV is presumably second only to bacteria in the rate it evolves: roughly a million times faster than mammalian evolution. Only a half-dozen decades after the probable introduction of HIV into the human population (most likely from chimps in Western Africa), variants of HIV have diverged to such an extent that if we were discussing something other than viruses we would call them separate species. But HIV evolution is not just of interest to biologists; it matters profoundly to doctors and their patients. In 1987, the first drug with an impact on HIV was tested in a clinical trial. At first, it appeared that a treatment, if not a cure, was at hand; but, 6 months later, the treated patients were found to be progressing to AIDS as fast as untreated. The virus had acquired mutations, negating the drug's benefit in every subject. It required a combination of three drugs to (partially) overcome the resistance problem.

The immune system fares better than a single drug, suppressing the infection for a decade on average; but the virus eventually learns to evade specific responses and escapes control. The implications of this evolutionary battle for vaccine design cannot be overstated. (After 25 years of research, we still have no licensed vaccine for HIV/AIDS.)

[1] Usually attributed to evolutionary biologist Theodosius Dobzhansky.

The motivation for writing this book derived from a scientific disagreement with our colleagues about why, and how, HIV evolves in a patient after infection. Most published discussions of HIV evolution *in vivo*[2] derive from "population genetics," a field founded in the 1920s by J.B.S. Haldane, Sewall Wright, and especially Ronald A. Fisher[3] (who published the first book about evolution written by a mathematician, in 1930). We will argue that the correct picture derives from a different tradition: that of "ecological genetics." The distinction between these perspectives has to do with enemies. Population genetics postulates that evolution goes on among rival species (or variants of a given species) who compete for niches or resources in a fixed, unreactive environment. The canonical system studied by population geneticists for the last 80 years has been, and remains, fruit flies raised in cages.[4] By contrast, ecological genetics emphasizes that all living organisms have, besides rivals, enemies—i.e., predators and parasites—and escaping them can be the driving force in evolution. In the case of HIV, either the virus or its principal foe—the human immune system—can be regarded as the predator (since in fact HIV targets certain immune-system cells); we prefer to imagine the so-called "killer" T-cells as the predators and HIV as the prey. However you think about it, HIV infecting a human body is not like fruit flies implanted in a jar.

[2] "*In vivo*" is a Latin phrase biologists use to refer to observations or experiments made in living organisms. Here we are contrasting it to "in populations," meaning (for HIV) the study of the epidemic around the globe. Of course, the antithesis of "*in vivo*" is properly "*in vitro*," meaning studies performed in a test tube or glass well.

[3] The same Fisher who introduced the p-value, the exact test, and randomization in clinical trials. Fisher was not satisfied with merely being one of the greatest statisticians who ever lived, but evidently wished to be the Einstein of evolutionary biology as well.

[4] For example, see Nature **467**: 587 (September 2010). The authors report on an evolutionary experiment with *Drosophila*, aka fruit flies, in which deep gene sequencing indicated that "... unconditionally advantageous alleles rarely arise, [or] are associated with small fitness gains ..." the usual conclusion from these types of experiments (see Chapter 1, Section 1.9). We find it remarkable that population geneticists do not entertain the notion that creatures living in cages, isolated from predators, for thousands of generations did not acquire new adaptive mutations in their genomes precisely because little advantage could be thereby obtained. One imagines that evolutionary pressures on free-living fruit flies are quite different.

To state the issue plainly: Fisher made a colossal mistake by leaving enemies out of his thinking and his models.[5] Fisher did have his reasons, of course; indeed, a kind of divorce accompanied the birth of mathematical biology. Ecologists and population geneticists parted ways, each chanting the mantra: "Ecology is short, but evolution is long." That is, changes in an ecosystem (e.g., a new variety of lion moving into your neighborhood, say) are fast (making an impact in your lifetime), but evolutionary changes (e.g., primates developing an upright-walking stance and ability to hurl spears) are slow (taking thousands of generations). Thus, when modeling one kind of change, the thinking went, you could ignore the other. The split has persisted to this day.

But HIV resists absorption into either system. It takes around 3 weeks for the immune system to contain an HIV infection and barely more time for HIV to escape by mutation from a single drug or immune response. The timescales of "ecology" (if we may be permitted to use the term when referring to what goes on in our bodies) and retroviral evolution are identical. Assuming the short-*vs*.-long mantra applies anywhere, it surely does not to HIV *in vivo*.

The historical parallel can be found in the work of field biologists such as E. B. Ford (who studied, among other things, how moths varied their spots to escape predation by birds[6]) later in the 20th century, which developed into the new paradigm Ford dubbed "ecological genetics." In reference to theory, the clash between population genetics and ecological genetics is perhaps most dramatically represented in "Van Valen's law," also known as the the "Red Queen Hypothesis."[7] It was proposed in 1973 by a professor of evolutionary biology at the University of Chicago, Leigh Van Valen.[8] Van Valen's key

[5]In addition, Fisher's "Fundamental Theorem of Natural Selection" was nonsensical, and by his choice to fix demographic population size he missed the possibility of quasi-species. See Chapters 1 and 8.

[6]Ford published the first book on the subject, bestowing its moniker, in 1964 [95].

[7]The reference is to the Alice books. At one point Alice remarks that she is tired of running and the Red Queen replies, "Now, *here*, you see, it takes all the running *you* can do to stay in the same place. If you want to get somewhere else, you must run at least twice as fast as that."

[8]1935–2010; New York Times obit, 10/31/2010. The NYT obituary noted that Van Valen could never get his idea published in an established journal (presumably due to opposition from population geneticists), and had to resort to founding his own and publishing himself. The citation is [299]; a scanned copy can be found on the Van Valen website.

insight was that the Fisherian conception of evolution as a friendly hill-climbing competition was misbegotten. The correct metaphor is military: evolution is an arms race.

Both population genetics and the theory we will expound here are formulated using mathematical models. Unlike population genetics, ecological genetics—because of its more complicated understanding of events—is not generally given abstract formulations. Not surprisingly, the modeling described here is specific to the context of HIV *in vivo* and the immune response. But this is not unusual; mathematicians have been at work modeling biology in particular settings for 9 decades or so. In ecology, models of predator-and-prey appeared in the 1920s, and of pathogen-and-host by the 1950s. Models of HIV infection and the immune response were proposed in the late 1990s. We have used such models in order to, among other things, estimate the number of HIV-infected cells a "killer" T-cell can kill every day, in the body of an infected patient. In 2006, several colleagues and an author used a stochastic version of our model to predict that T-cells alone could abort a retroviral infection, a very controversial conception that was verified a few years later in vaccine experiments in monkeys. Likewise, the motivation for modeling HIV evolution *in vivo* was to generate novel predictions for experiments in animals or interventions in the clinic. The reader will find roughly a dozen such predictions in the pages of this book.[9] Here we mention only the most important one: it may be possible to design a vaccine that blocks HIV from escaping the immune system's clutch.

We wished to write for a heterogeneous audience, including researchers, physicians, patients, teachers, and students; indeed, anyone with an interest in HIV, the immune system, evolution, or modeling in biology or medicine. This desire lead to the inevitable decision to split the book into two parts. Part I contains the biology, description of the models, some easily-grasped formulas, and the conclusions, expressed plainly in the text or in computer-generated figures. There are no prerequisites to read this part. As in any scientific field, technical jargon is unavoidable, so, in addition to defining each term when introduced, we have added a Glossary of frequently used phrases and acronyms after the text. Part II contains the Greek-lettered equations. We have included a brief tutorial about modeling and exercises

[9]See the entries listed under "predictions" in the Index.

in the chapters. We discuss the philosophy of modeling—a topic which can be neglected by the practitioner only at great risk—in the Introduction to Part I.

We wish to thank several colleagues for their help in writing this book: Peter Gilbert, for many discussions about biostatistics, and Fusheng Li, ditto for biological data mining. Finally, W.D.W. thanks the Fred Hutchinson Cancer Research Center and the US National Institutes of Health for supporting in part the research described here (*via* grant 1R01AI05428). The opinions expressed in this book are solely those of the authors and do not represent the views of the National Institute of Allergy and Infectious Diseases or the NIH.

Seattle, WA, USA — W. David Wick
Los Angeles, CA, USA — Otto O. Yang
September 2011[10]

[10]About dates: the bulk of this work was written before 2007. As a result, some of the material, in particular the figures and parameter tables, could be updated to reflect new information from experiments reported after that date. Comments about updating can be found in the Notes to the chapters. However, we are not aware of the demise of any of the principal conclusions of this book on the basis of more recent experiments, although that is a common fate of theories in biology.

# Contents

# List of Figures

# Chapter 1

# Introduction

## 1.1 The HIV Epidemic and Its Origins

In 1984, two groups of investigators—Luc A. Montagnier's at the Pasteur Institute in Paris and Robert C. Gallo's at the National Institutes of Health in Bethesda, Maryland—announced the discovery of the human immunodeficiency virus (HIV in this book, although H.I.V. in the New York Times).[1] The clinical manifestations of a new disease, acquired immunodeficiency syndrome (AIDS), had been observed 3 years earlier, in a population of men who have sex with men, in several cities of the United States.[2] At this writing, HIV has established a global pandemic, among the worst in recorded history; 33 million people are currently infected worldwide, with 2.5 million new infections each year.[3] At least 25 million have died of AIDS. The primary modes

[1]The 2008 Nobel Prize in Medicine was awarded to Montagnier and Françoise Barré-Sinoussi for the discovery of the human immunodeficiency virus and Harald zur Hausen for discover of the human papilloma virus, which causes cervical cancer. The Karolinska Institute's opinion about the discoverers of HIV is controversial, coming after a long and acrimonious dispute between the French and American camps.

[2]A brief account of a cluster of cases in Los Angeles of combined pneumocystis pneumonia and Kaposi's sarcoma appeared in the Morbidity and Mortality Weekly, published by the Centers for Disease Control (CDC), in June, 1981. The combination of a rare form of pneumonia with a rare skin cancer alerted epidemiologists that a new infectious disease may have appeared.

[3]These are revised estimates by UNAIDS, released on November 20, 2007. The current estimate of 33.2 (confidence interval, [30.6,36.1]) million infected replaced an earlier estimate of 39.5 ([34.1,47.1]); the UN agency's revisions reflects lowered estimates primarily of the epidemic in India.

W.D. Wick and O.O. Yang, *War in the Body: The Evolutionary Arms Race Between HIV and the Human Immune System and the Implications for Vaccines*,
DOI 10.1007/978-1-4614-7294-0_1, © Springer Science+Business Media New York 2013

of transmission at this time are unprotected sexual activity and intravenous drug users sharing syringes. Although the government's top-ranking doctor, Edward Brandt Jr., said shortly after the virus was discovered that he was optimistic that a vaccine would be available by 1987, 20 years later no vaccine against HIV/AIDS has been licensed.

HIV is a retrovirus that primarily targets certain cells of the human immune system.[4] The prefix "retro" refers to an aspect of the virus's lifecycle in the body ("*in vivo*" for biologists, who prefer the Latin): in order to replicate, HIV must first integrate its genes, stored on molecules of RNA, into the host's DNA. This process violates the once-canonical doctrine about the flow of biological information in cells (from DNA into RNA and then to proteins); hence the name. The retroviridae are a ubiquitous family of parasites, or at least fellow-travelers, of vertebrate animals; thus monkeys are infected by various strains of simian immunodeficiency virus (SIV); cats, by feline immunodeficiency virus (FIV); mice, by murine acquired immunodeficiency virus (MAIDS); and so forth. In some cases, these viruses in their natural hosts do not cause disease, but when transmitted to a new host can become pathogenic. A well-studied instance is SIV in sooty mangabeys and African green monkeys, which is tolerated for the animal's lifetime, while injecting variants of this virus into Asian macaques has proved useful for studying simian AIDS. Lessons learned, it is hoped, will be relevant also for human disease. SIV and HIV are also known as "primate lentiviruses" ("lenti" is Latin for "slow"; the appellation refers to the long time between infection and disease).

The HIV epidemic in humans is thought to have begun 50–100 years ago, in at least two zoonotic transmission events. The variety that has spread around the world, called HIV-1, probably originated when an SIV-infected chimpanzee[5] was butchered for its meat, which is consumed in parts of forested central Africa. The accused strain of SIV is prevalent in chimps in southern Senegal and does not appear to make them sick—rendering HIV-1 another, and very important, example of a microbe that became pathogenic after jumping to a new host. A different set of strains also circulating in humans (primarily in West Africa, but also in Europe, especially Portugal, and southwestern India), collectively called HIV-2, probably

[4] "Virus" is Latin for poison or slime; these disease-causing agents—originally called "filterable viruses" because they passed through the finest sieves in the laboratory—were discovered around 1915.

[5] Pan troglodytes troglodytes; the "greater" chimpanzee.

derived from a smaller monkey, a sooty mangabey,[6] possibly from seven separate transmission events. HIV-2 also causes AIDS, but disease is frequently less severe and patients live significantly longer. The virus is also more difficult to transmit, perhaps explaining why HIV-2 is endemic in various regions rather than pandemic worldwide. Co-infection is common in areas where both viruses are prevalent. There is another family of clinically-relevant retroviruses, which cause some cancers in humans, called HTLV (which were discovered before HIV). Because the laboratory and clinical research that has informed our modeling was conducted with various strains of HIV-1, we will limit our discussion in this book to these viruses, which we collectively label simply as "HIV" unless particular variants need be described.

The dynamics of HIV transmission in sexual and injection networks has an obvious bearing on the epidemic and has attracted much interest from epidemiologists, statisticians, and mathematical modelers over the years. The focus of this book, however, is on a different "epidemic": the one in the cells of an infected patient's body, and, in particular, the evolution of the virus in that single infection, during the patient's lifetime.

### 1.1.1 Notes

For the discovery of HIV: [104]. For its zoonotic origins: [161, 296]. Steve Self and an author wrote one of the first papers about sexual networks and HIV; they explored the impact of "superspreaders" on the problem of estimating vaccine efficacy in a clinical trial: [308].

## 1.2 HIV *In Vivo*: Part I. Time-Course and Target-Cells

The course of an HIV infection is conventionally divided into three stages: primary viremia, chronic or asymptomatic phase, and AIDS. The first, or primary, phase begins, in the sexually-transmitted case, with a localized infection in a mucous membrane, then spreads to lymph nodes. (Lymph nodes are cell-concentration regions of the lymphatic system, a secondary circulation in the body that organizes and sustains immune responses.) HIV can infect many cell types, but the primary targets of this virus are the

[6] Cercocebus atys.

macrophages and the CD4+ T lymphocytes. Macrophages ("big eaters") are immune-system cells that consume bacteria and matter from dead cells as well as having other immune functions. T lymphocytes, also called thymus-derived or T-cells, are also immune-system cells. The descriptive terms reflect a T-cell's bodily visitations: after generation in bone marrow, as for all blood cells, they transit through the thymus gland and reside for part of their lifetime in lymph nodes. There are several varieties of T-cells. HIV's T-cell targets are dubbed by immunologists "CD4+"; the plus sign means the cells stain positively by a reagent for the surface marker, called CD4. Another kind of immune-system cell, called the dendritic cell (DC), frequents mucous membranes and skin, where it picks up pathogens and transports them to lymph nodes for display to T-cells. HIV may hitch a ride on DCs, but does not reproduce in them. As we write, the question of which cell type is the most-likely portal for HIV's entry into the body—macrophage, CD4+ T-cell, or DC—is still unsettled.

Although macrophages support HIV replication, most virologists believe that HIV is primarily a disease of T cells. We will use the acronym "PIT" in this book for "productively-infected target" cell, by which we shall usually mean a CD4+ T-cell. The adjective "productive" refers to on-going viral replication, and distinguishes PITs from a class of latently-infected cells (discussed in Chapter 2.)

Viruses are obligate parasites, which means that they can reproduce only in their targeted "host" cells. A virus separate from its host cell is a lifeless particle, called a "virion", made up of protein and either DNA or RNA molecules. See Figure 1.1.

The latter, constituting the viral "genome", stores the information necessary to replicate the virus. HIV's genome consists of two strands of RNA. Its life-cycle in CD4+ T cells proceeds as follows. See Figure 1.2. First, a free virion latches on to a CD4 molecule and another molecule called a co-receptor on the surface of the target cell. The virion binds with these cell-membrane receptors and is engulfed into the cell cytoplasm, where it uncoats and releases its genome and certain proteins such as the celebrated "reverse-transcriptase" (RT) enzyme. (RT, discovered in 1970 by Temin and Baltimore, is the molecule that makes retroviruses retro.) Next, RT, in collaboration with host enzymes, transcribes viral genes (from both RNAs; the details are in Chapter 10) into a single strand of DNA. The later penetrates the cell's nucleus, where it is inserted into the host's genome. At this stage, the viral DNA is said to constitute a "provirus".

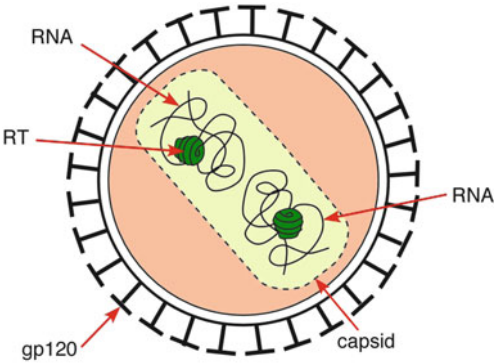


Figure 1.1: Cartoon version of an HIV virion.

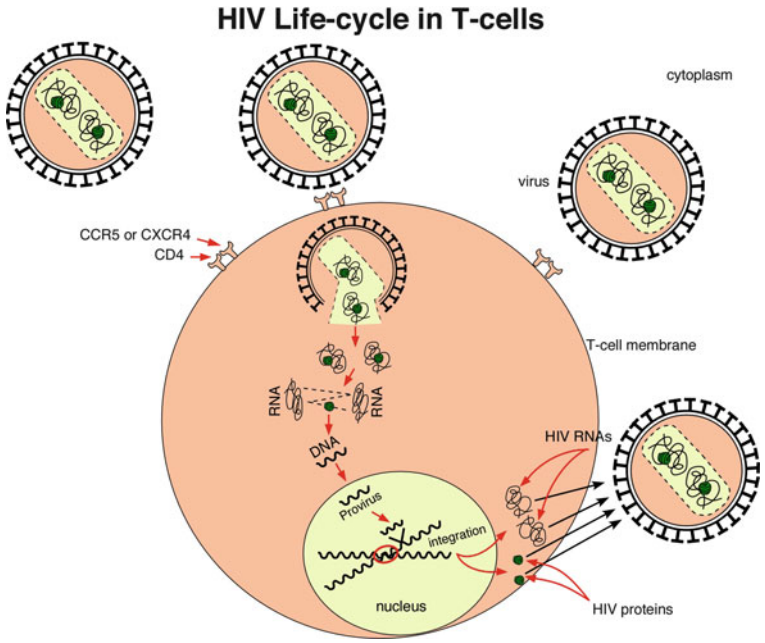


Figure 1.2: Cartoon version of HIV's life-cycle *in vivo*.

The stage is set for the final step in viral replication: synthesizing and releasing new virions. The host's gene-expression system, unable to distinguish self from non-self DNA, transcribes the viral genes into RNA and proteins. These assemble, with some help from host proteins, into new virions. In the last act, either the virions bud directly from the cell membrane, or the cell's excretion machinery releases them, into the extra-cellular medium where they can infect more target cells.[7] The whole process—from initial penetration of the cell membrane to release of the first new infectious virion, often referred to as the "eclipse period"—takes 2–3 days.[8]

Although most of the action occurs where T-cells mostly reside—in lymph nodes and other organs, including the spleen, tonsils, and gut—the amount of virus is usually measured, for practical reasons, in peripheral blood. The level of virus is called "viral load" (VL) or simply "viremia", and is usually expressed as virions per milliliter (ml) of blood. (Sometimes RNA copies is meant instead, which causes confusion because there are two per virion. Perhaps more logical would be total-body burden of virions, but that is not a simple laboratory measurement.) As the infection proceeds, the VL typically reaches ten million or more, peaking in 20–40 days; in this period the patient often reports symptoms similar to that of the flu. This stage is called "primary viremia". At this time the CD4-bearing T-cells in peripheral blood typically drop by around 50 %, from a normal 1,000 per microliter—in immunology, cell concentrations are given "per $\mu l$"—to around 500. Due to the many controversies about HIV and T-cell dynamics, which will be a recurrent theme in later sections, it is not known what this implies about the infection.

After the primary stage, the VL typically falls by a factor of 100–1,000 ("2–3 logs"), for reasons that are also controversial and discussed in Section 1.6. The CD4+ T cell population in peripheral blood usually recovers somewhat, but not to prior levels. The stage that follows is called the "chronic", "asymptomatic", or (by mathematicians, rather imprecisely), the "steady-state" period. "Quasi-steady-state" would be more accurate. Characterized

[7]There is also a reported route for virions to pass directly between cells by hijacking the immunological synapse formed between APCs and T-cells or creating a "virological synapse" between two CD4+ T-cells. Although this mechanism for cell-to-cell spread had been demonstrated *in vitro*, its importance *in vivo* is unknown.

[8]The eclipse period was incorrectly estimated in 1996, from data about declining viremia after HAART, to be 1.1 day [237], a figure which made it into many models including some of my own. For discussion of this issue, see the Notes to Section 1.7.

by almost-stable viral loads and CD4 cell counts and lack of AIDS symptoms, this period lasts a highly-variable time averaging 10 years. The onset of the last phase—AIDS—has since 1993 been defined by the Centers for Disease Control as a CD4 count of 200 or less and at least one AIDS-defining diagnosis of an infection, symptom, or cancer rare among immunocompetent persons. The cause of the transition—chronic phase to AIDS—is also unknown at this time. Some conjectures will be discussed in this book.

### 1.2.1 Notes

Good reviews of the biology of HIV infection include [11, 92, 255, 285]; the last also contains articles on other viruses (such as SIV and LCMV) that establish chronic infections in animals. Direct cell-to-cell spread and the "virological synapse": [152, 244]. CXCR4 expression on naïve T-cells was essential in these experiments and the observation limited to X4 viruses (see Section 1.6), so the importance of this mechanism in primary infection by an R5 virus is unclear.

## 1.3 HIV *In vivo*: Part II. The Mutation Machine

Perhaps the most striking fact about HIV *in vivo* is its extraordinary replication rate. Unlike some viruses (such as chickenpox or herpes), HIV never enters a dormant or "latent" stage, but reproduces continuously in the body over the whole time-course of infection. The number of PITs in the chronic phase is in the range 10–100 million, and the turnover time is 2–4 days. HIV's mutation rate is also remarkable: it is at least five orders-of-magnitude higher than for DNA-bearing, eukaryotic organisms. The rate was measured in the early 1990s in the test tube, with HIV propagating in immortalized T-cell lines (again, biologists prefer the Latin, and refer to the observation as "*in vitro*", literally in glass) and yielded the average figure: about 0.3 changes per genome per replication cycle. The cause of this error-prone replication is sloppy reverse-transcription of viral RNA into DNA. Now the HIV genome is quite small, even among viruses: about 9.6 kilobases. This yields a probability for substituting a particular nucleotide by another of about $3\times10^{-5}$ per cycle. From these figures it easily follows that every possible HIV mutation is made every day in the body of an infected person, although that does not

mean that any newly-created variant is destined to replace the existing, or "wild-type", strain—an issue that will greatly concern us in this book.

The extremely high rate of generating variants sets the stage for evolution to act in an infected person's lifetime. Indeed, after HIV infection the virus evolves rapidly, changing by as much as 1 % per year in part of the envelope protein. When a variant with a mutation enjoys a replicative advantage, geneticists refer to "positive selection"; and it is frequent in HIV evolution. Selection acts on "phenotypes", or behavioral repertories, which derive from amino-acid sequences in proteins. Each amino-acid is coded in the genome by a triple of nucleotides but there is some redundancy. Nucleotide changes that do not alter amino-acids are referred to as "silent" or "synonymous". When a synonymous change occurs, or a substitution of an amino-acid for another that does not change the function of the protein, geneticists invoke the word "neutrality".[9] Much neutral variation is observed in the HIV genome, which is not surprising given the high mutation-rate, and also many positively-selected events which may reflect the virus improving its reproductive facility in its host cells or expanding its range to new cell-types. But many, even the majority, of adaptive amino-acid changes allow the virus to evade immune pressure. This form of *in vivo* evolution—which we will refer to by the generic term, "escape"—is the central focus of this book.

### 1.3.1 Notes

The high mutation-rate of RNA viruses was discovered in the 1970s; for the rate in HIV, see [191]. That HIV might mutate and thus escape a T-cell immune response was proposed in 1991 [240]. HIV makes every possible mutation every day: [63]. For a recent review of HIV genetics and AIDS: [220].

The rapid evolution in HIV's envelope protein (Env) in the first year after infection is thought to represent mainly escape from antibodies [261], discussed in Section 1.5. For the frequency of positively-selected mutations, see [70, 334, 335]. That a large fraction are escapes from CTLs: [3, 44, 189]. In one study of four patients for 5 years post-infection, 53 % of 98 a.a. changes were CTL-related [3]. In another recent multicenter study of 75 primary

[9] It is sometimes asserted that only non-synonymous mutations can confer an advantage, but there are situations—involving RNA secondary structure or codon usage—where synonymous changes also affect fitness. Nevertheless, geneticists often use the ratio of synonymous to non-synonymous variation as a measure of the relative importance of neutral "drift" and selection.

HIV patients, a minimum of 20–40 % of amino-acid changes in the Nef and Pol proteins in the first 2 years after infection were probably CTL escape mutations [44].

## 1.4 The Experience with Drugs

In 1987, only 4 years after the laboratory isolation of HIV, the first drug that had a clinical impact on HIV disease was approved by the FDA. This drug, zidovudine, also called AZT, was but the fifth antiviral drug ever licensed and one of very few to be rationally designed.[10] The drug had been shown, by researchers at the FDA, the National Cancer Institute, and Burroughs Wellcome Co. (the drug's sponsor), in *in vitro* studies to inhibit the reverse-transcriptase (RT) enzyme; only a year later, a clinical trial was halted early after the observation that patients on the drug had lowered viral burdens and better quality of life. At least, that is, for about 6 months. Follow-up revealed that as many patients taking AZT progressed to disease as would be expected in a drug-naïve group. Investigators next tested the virus from progressors and discovered that it had changed—grown less sensitive to the drug. Sequencing revealed that the virus had made three-to-four mutations in the RT region of its genome that had allowed it to escape suppression by the drug. The hope expressed (by, among others, mathematicians) that AZT would prevent AIDS was dashed.

In retrospect, the only surprising part of this episode—besides the unprecedented pace of drug discovery and approval—is the 6 months. Recall that HIV *in vivo* is capable of making every possible mutation every day; hence, one would expect it could evade one drug in a few weeks, if possible at all. Escape from monotherapy with other drugs, often by a single amino-acid-change, is indeed that fast. It turned out that, in the case of AZT, one

[10] A rationally-designed drug is one that was developed on the basis of hypotheses about disease mechanisms at the molecular level, perhaps including detailed X-ray photographs revealing the structure of targeted proteins. By contrast, most drugs in clinical use, even today, were discovered in screening programs and the mechanism of action is often unknown. The earlier anti-viral drugs licensed were acyclovir, amantadine, vidarabine, and ribavirin, active against herpes, flu, and hepatitis. The hurdle in discovering anti-viral drugs is that viruses hijack cellular enzymes to facilitate reproduction, so drugs acting against them are often toxic for the host. Bacteria bring along their own enzymatic machinery and thus are easier to target.

of the amino-acid-replacements the virus made to escape the drug required two nucleotide changes—which is a rarer event.

By 1995, pharmaceutical companies had invented other anti-HIV drugs, especially a new class called protease-inhibitors (which acted against a different HIV enzyme). Two groups enlisted patients in experiments in which they were given three drugs at a time. This regimen—called "triple-combination" therapy, or "HAART" (highly-active antiretroviral therapy)—proved able to reduce viral-loads in the patient's blood streams almost to unmeasurable levels. One of the physician-investigators in these experiments, David Ho, appeared on the cover of Time magazine (as "Man of the Year"), and once again there were predictions that HAART would cure HIV, or at least prevent AIDS.[11] For many patients with access to these drugs, triple-combination therapy has prolonged their lives; but for others resistance eventually deve lops—a scenario that we will encounter again, in another context, in this book.

### 1.4.1 Notes

For the history of antivirals before AZT see [64]. For development and licensing of the first anti-AIDS drug, AZT, see [337]. For the HAART experiments, see [129, 305].

## 1.5 The "Killer" T Cells

T (for thymus) cells are part of the "adaptive immune system". The phrase distinguishes the T-cells and B-cells from other immune-system actors that generate the "innate immune response". As the name implies, the adaptive response improves with exposure to the pathogen, for which the host retains a memory; by contrast, the innate response simply detects certain patterns in pathogen proteins or nucleic acids and makes a fixed response. Considering the breadth of the microbial world, storing all existent pathogen-patterns in the human genome is presumed impossible; thus the common conjecture holds that the adaptive immune system evolved in vertebrate animals to defend against diverse enemies: some ancient and some never before encountered.

[11] The assertion was made at the time that 36 months of drug treatment would amount to a cure. This prediction failure was probably due to neglecting the possibility of latency at the cellular level (see Chapters 2 and 10) rather than escape by mutation.