Commentary

Error catastrophe and antiviral strategy

Manfred Eigen*

Max-Planck-Institut für Biophysikalische Chemie, Karl-Friedrich-Bonhoeffer Institut, Am Fassberg 11, D–37077 Goettingen, Germany

The term "error catastrophe," originally introduced in the theory of molecular evolution (1), has become fashionable among virologists. In a recent paper in PNAS (2), it was suggested, on the basis of quantitative sequence studies, that ribavirin, a common antiviral drug, by its mutagenic action drives poliovirus into an error catastrophe of replication, thereby turning a productive infection into an abortive one. Previous studies by Loeb and his group (3, 4) on the AIDS virus (HIV) and by Domingo, Holland, and coworkers (5, 6) on foot-and-mouth disease virus (FMDV) have led to similar conclusions, suggesting a paradigm shift in antiviral strategies (7). A recent issue of PNAS presents a paper by Grande-Pérez *et al.* (8), which deals with the "molecular indetermination in the transition to error catastrophe," shedding light on the complexity of the mechanisms involved in virus infection and stressing the need for a careful molecular analysis of the detail, which may differ greatly from one virus to another. Because of its practical relevance for developing potent antiviral drugs and, beyond that, its general importance for an understanding of molecular evolution, this commentary will highlight the theoretical basis and point out the kind of conclusions that can be drawn in discussing experimental results.

The term error catastrophe is of a descriptive nature and lacks a clear-cut definition. A catastrophe is usually triggered if certain tolerances are exceeded. For replication, there is indeed such a limiting value of error or mutation rate that must not be surpassed if the wild type is to be kept stable. We call this limit the "error threshold." Why is it a sharply defined limit? Why does the efficiency of replication not vary monotonically with the error rate? The information stored in the genomic sequence melts like ice at 0°C. This comparison is indeed a very apt one. The information melts away in a process that has all the physical characteristics of a first-order phase transition requiring cooperative behavior with unlimited coherence lengths, as we encounter in the melting of a solid or the evaporation of a liquid at its boiling point. The error threshold is caused by the inherent autocatalytic nature of replication, which represents not only the transfer of information from one generation to the next, as would be the case for a message sent through a transmission channel. Rather, replication provides an exponential proliferation of the information contained in the sequence as a whole. In the population formed, this results in competition among the various slightly differing sequences, which behave as cooperative units. Natural selection is a direct consequence of this competitive replication. It presupposes differences in efficiency of replication without excluding neutral mutants. Neutral copies, all belonging to the group of best-adapted ones, are selected against the rest, but because of their inherently reproductive behavior, they continue to compete with one another in a stochastic manner. Kimura and Ohta (9) called this nondeterministic fluctuating selection "non-Darwinian," although Darwin himself anticipated it. Kimura and Ohta's stochastically fluctuating selection reminds us of "critical phase transitions," as found in ferro- or antiferromagnetism or liquid-gas transformation near the critical point where, in analogy to neutrality among replicative units, the densities of the liquid and gaseous phases become equal, with the consequence of density fluctuations on all scales of spatial dimensions manifesting themselves in the phenomenon of "critical opalescence."

However, note that these phase transitions associated with natural selection do not take place in the space–time coordinates of our physical space. They refer rather to an abstract "information space" and are therefore not easy to visualize, because they may appear scattered in physical space and over extended periods of time. Information space is a discrete point space with a metric named after Richard Hamming (10). Each of the possible 4^N sequences of length N is assigned to one and only one point, with all neighborhoods among sequences correctly ordered according to their kinship distances. This "spatial" order requires a 2^{2N}-dimensional Hamming space. The dynamical equations of the rise and fall of populations can be written in a fairly general phenomenological form, yielding the quasispecies model (11, 12). A quasispecies is a population structure in information space and is the "condensed" mutant distribution that results from the phase transition representing natural selection. It has been termed "quasispecies" because the whole distribution behaves "quasi" as a single species, because it is determined by one (namely the largest) eigenvalue of its system of dynamical equations. The eigenvalues, being invariants of the equations, are determined as soon as the mutant spectrum is defined, regardless of whether the final stationary population structure is achieved. Rather than elaborating on further details of theory, I shall now discuss the important parameters that determine selection and hence also the behavior of virus populations, as expressed in the work this commentary refers to.

Fig. 1 shows a computer simulation of a model case that is representative of the phenomenon of error catastrophe. Such simulations were first performed by Schuster and Swetina (13). The present example was computed by Tarazona (14). It shows the stationary structure of a population consisting of binary sequences of length $N = 20$, in which all sequences have equal values of all their replication rates except for one sequence, which shows a 10-fold higher rate. The error rate $(1 - q)$, i.e., the relative number of misincorporations per site, has been assumed to be uniform for all sequences in the distribution. Fig. 1 shows a plot of the relative population number of the steady-state population against the error rate $(1 - q)$, the numbers 0, 1, 2, etc., referring to 0 errors (= master sequence) and the sums of all of the 1, 2, 3, error sequences, respectively. The error threshold is seen clearly at $1 - q \approx 0.11$. Although the individual curves vary quite markedly with the error rate, the order of the quasispecies, represented by the consensus sequence, is clearly conserved up to the "melting point," i.e., the error threshold. Above the error threshold, each of the $\approx 10^6$ (i.e., 2^N) possible individual sequences occurs with the same probability of $\approx 10^{-6}$. Because the distribution was centered around the master sequence (0

See companion article on page 12938 in issue 20 of volume 99.

*E-mail: Manfred.Eigen@gwdg.de.

www.pnas.org/cgi/doi/10.1073/pnas.212514799

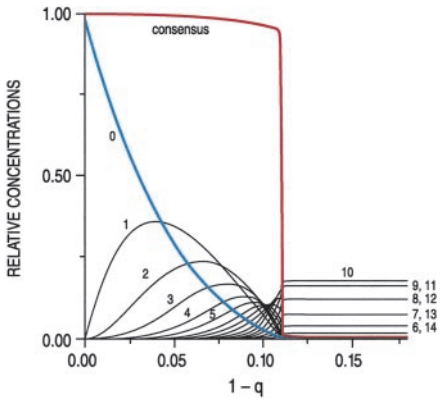


Fig. 1. Relative population numbers of binary sequences S_k (ordinate) as functions of single-digit error rate $(1 - q)$ (abscissa). The length of all binary sequences is $N = 20$. All $2^N \approx 10^6$ sequences are degenerate in their reproductivity except for one ''master'' sequence S_m, which reproduces 10 times more efficiently than the rest. The resulting quasispecies distribution is centered at the master sequence (''0'' errors). The numbers 1, 2, . . . 20 refer to the sum of all 1-, 2-, . . . 20-error mutants. The red curve refers to the consensus sequence, which shows a sharp first-order phase transition at the error threshold.

errors), the sum of all k-error sequences is given by the binomial coefficient $\binom{N}{k}$, which has its maximum at $k = N/2$. That is why the sum curve for 10 errors shows the most frequent representation (corresponding to about 17.6% of the $\approx 10^6$ sequences). Above the threshold, no memory of the former wild type remains.

The model simulated in Fig. 1 suffices to show the clear analogy to a first-order phase transition (14) applying to the information content of the quasispecies as a whole. The transition would be even sharper for (more realistic) longer sequences, reaching a slope of ∞ for $N \to \infty$, whereas smaller selective advantages of the wild type would just cause a shift of the transition point along the abscissa to lower error rates without much changing the overall shapes of the curves. What is essential is that the population of the master and of individual mutant types already changes quite conspicuously below the error threshold but is ''all or none'' above the threshold. However, the model chosen is entirely unrealistic if we want to apply it to our discussion of virus quasispecies. Let me discuss the parameters that determine the shape of the curves, and we shall see more clearly what to expect in situations closer to those of real viruses.

(*i*) The Most Obvious Parameter Is the Error Rate

First, a uniform fidelity q for all positions, and hence a uniform error rate $1 - q$, is absolutely fictitious. Grand-Pérez *et al.* (8) emphasize that different regions of the virus sequences must have different error rates, and it has long been known that it is not only the kind of base and its nearest neighbors that make individual positions more or less variable, ranging from exceedingly conservative positions to ''hot spots.'' This is the more so if mutations are enhanced by using drugs that resemble base analogues. Because the probability of mutants is sequence-specific, we must introduce two indices: one for the particular sequence S_k and another one for the positions ''i '' within each sequence S_k. Hence the fidelity of a position thus characterized is q_{ik} and its corresponding error rate $(1 - q_{ik})$. The overall fidelity of reproducing any given sequence S_k is then the product $q_{1k}q_{2k} \ldots q_{Nk}$ or $\prod_{i=1}^{N_k} q_{ik} \equiv \hat{q}_k^{N_k}$, where $\hat{q}_k$ is the geometric mean of all q_{ik} of sequence k, whereas N_k is its length expressed as the number of nucleotides. The geometric mean differs from the arithmetic by weighting more sensitively individual elements. Although some elements can reach zero without much changing the value of the arithmetic mean, they must not do so for the geometric mean, where the product becomes zero if one of its elements is zero. This property of the geometric mean has important consequences for the possible relevance of certain singular mutations. The theory, on the other hand, does not change its formal structure by the introduction of averages. Instead of a uniform fidelity q or error rate $(1 - q)$ (abscissa in Fig. 1), we now use the averages $\hat{q}_k$ and call the sequence-specific overall fidelity $\hat{q}_k^{N_k} \equiv Q_k$. Because $\hat{q}_k$ for any realistic sequence is very close to one, meaning a small average error rate $(1 - \hat{q}_k)$, as found between 10^{-3} and 10^{-5} for RNA viruses, a fairly precise approximation to Q_k is given by $e^{-N_k(1-\hat{q}_k)}$. The exponent is the average number of errors per sequence and can yield Q_k values appreciably smaller than one. The conclusion of this paragraph is that the probability distribution of mutants in a quasispecies is not at all uniform; single mutants may reproducibly appear orders of magnitude more frequently than others, of which some may have only a sporadic existence. This all happens below the error threshold and may produce quite nonuniform mutant distributions close to the error threshold, although the all-or-none nature of the error threshold is first realized after it is crossed.

(*ii*) The Error Threshold Is Not Solely Determined by the Average Error Rate

Of equal importance is the fitness landscape within the quasispecies distribution. Let us take the example represented by Fig. 1, in which one of the 10^6 different sequences, called S_m, is clearly distinguished by a (massive) selective advantage. At the critical error rate $(1 - \hat{q}_m)$, the master sequence S_m has a fidelity $Q_m = e^{-N_m(1-q_m)}$ with a value well below 1, which then requires a corresponding selective advantage of the master sequence, relative to the average of its mutant distribution. Any mutation occurring in the master sequence reduces its frequency of occurrence, whereas any mutation occurring in the rest of the quasispecies (including the mutations that come about in the master sequence) produces some other member of this mutant distribution. In other words, mutations in the rest of the distribution do not lower their total number. The master sequence must therefore be at least σ_m times more efficient in its reproduction to make up for the loss caused by its mutation rate, such that $\sigma_m Q_m > 1$. In the above example, σ_m can be easily calculated. For realistic nonuniform distributions, which, in addition, may contain several neutral master copies, the σ functions, although clearly definable by the eigenvalues and eigenvectors of the exact solutions, would be calculable only if the details of the fitness landscapes are known. Experimental data, such as those presented in the papers under consideration, are therefore most important.

We now see that it is not only the irregularity in the mutant distribution but also the contribution of each mutant to the reproductivity of the quasispecies that is rated by natural selection. A particular mutation that appears to occur very frequently may be either neutral or, if it involves a ''strategic'' position, deleterious, or under certain environmental changes, such as those produced by the immune response of the host, it may be advantageous for the virus. These situations cause quite dramatic differences for different viruses, as is known for polio virus and HIV (15). The equivalent importance of σ_m and Q_m is obvious from the symmetric condition $\sigma_m Q_m > 1$. It is somewhat obscured in the standard expression defining the error threshold: $(1 - \hat{q}_m) > \ln \sigma_m / N_m$ which follows from $\sigma_m Q_m$ if one substitutes for Q_m the exponential $e^{-N_m(1-\hat{q}_m)}$. The logarithmic dependence on σ_m in the error threshold relation seems to belittle its influence, because the logarithm of numbers clearly larger than one does not depart far from one. However, σ values may be very close to one, i.e., equal to $1 + \varepsilon$, where $\varepsilon \ll 1$. Then the logarithm is a very small number: $\ln(1 + \varepsilon) \approx \varepsilon$ for $\varepsilon \ll 1$. In irregular fitness landscapes and irregular mutant distributions, σ_m is a complicated function of the variables concealed in the averages σ_m and Q_m. As in the case of fidelity Q_m, much of the detail appearing in the precise form of eigenvalues and eigenvectors gets lost in the averaging procedures leading to σ_m values. Selectivity is different for different

fitness classes of mutants, the extremes being neutrals ($\sigma \to 1$) and nonreproducible deleterious mutants ($\sigma \to \infty$). In addition, there might be large fluctuations even for a given virus in the same environment, when very rare (but important) mutations occur stochastically in different temporal sequences.

(*iii*) Virus Infection Involves More Complex Operations than Just the Replication of RNA or DNA

The genomes of even the smallest viruses encode several functions, of which replication is only one, albeit a very important one. Yet what is finally weighted for selection is the virus, overall performance in the infection process. We have studied the kinetics of infection with the example of the bacteriophage Q_β [Eigen *et al.* (16)]. The experiments carried out by M. Gebinoga (17) involved pulse-like infection. Samples of host cells were incubated with virus for defined lengths of time, i.e., quenching the infection after 1, 2, 3, etc., minutes up to the total time interval between infection and lysis of the host cells. The quenched samples were carefully treated with toluene to remove the outer (lipid) membrane of the host cells. The remaining murein sacculus, which is not penetrable for larger molecules such as proteins and nucleic acids or for organelles, was perfused with radioactively labeled nucleoside triphosphates and amino acids. The kinetics of both RNA and protein synthesis then were recorded. The rates, extrapolated to the starting point, which reflects the profile of the process *in vivo*, are plotted as a function of time in Fig. 2. As seen, at the moment of infection the rate of RNA formation is zero, because specific replicase is not yet available, but protein synthesis (using the translation machinery of the host) is active right from the start. After about 10 min, enough replicase has been formed so that both replication and translation can now compete, leading to a sharp (i.e., hyperbolic rather than exponential) increase of both protein and RNA concentration, as expected for a nonlinear "hypercyclic" (18) mechanism. At the same time, replicase formation is down-regulated by the binding of coat protein (which acts as regulator) to the replicase gene. The burst of synthesis comes to an abrupt halt when the amount of RNA present has increased sufficiently to block all available binding sites provided by the host ribosome population as well as those of the (ultimately constant) replicase concentration—very much as in an "end-point titration." From now on, both RNA and coat protein production proceeds at a constant rate until the host cell lyses, about 40 min after infection. This is a highly regulated mechanism, which yields about 10,000–20,000 complete virus particles (each consisting of one plus strand of RNA and 180 coat proteins that form its icosahedral shell). Because of error accumulation, only less than 10% of the viruses produced are viable. Nevertheless, this is more than sufficient to maintain the autocatalytic growth nature of the overall process, described by the phenomenological rate equations.

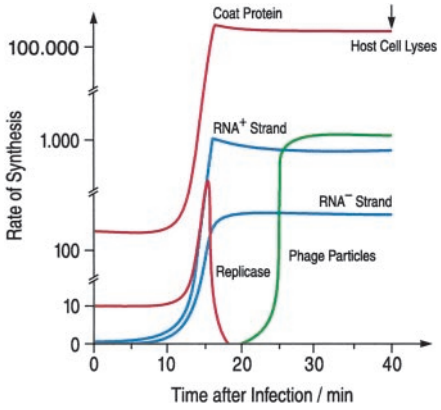


Fig. 2. Kinetic profile for the infection of *E. coli* by phage Q_β. The total infection cycle lasts about 40 min. The ordinate refers to the rates (i.e., the number of particles formed per minute and per host cell) that are plotted as functions of time. Red lines represent proteins (coat protein and replicase); blue lines, RNA (plus and minus strands); and the green line, complete virus particles.

What I wanted to show is the complex nature of the overall process of virus infection, which differs from one type to another but in no case is just simple replication. Coming back now to the objective of this commentary, I would like to emphasize three points:

(*i*) Very close to the error threshold, the mutant spectrum, as Fig. 1 demonstrates, becomes quite diverse. Before the phase transition occurs, master sequence and low-error mutants become a minority even in the unrealistic model case of uniform error rates. The wild type keeps its distance from this point of transition to maintain robust stability. The application of mutagens may change this situation in an uncontrolled way.

(*ii*) The resulting error spectrum involves neutral, deleterious, and also—under the new conditions—advantageous mutants. In which manner they are effective, before the total information "melts away" completely, depends on the particular type of virus.

(*iii*) The mutant spectrum expresses itself in a spectrum of phenotypic functions that include all processes involved in the complex infection mechanism. Hence, error catastrophe is intimately linked to all functions involved, because it depends on both (realistically quite complex) parameters Q and σ.

The paper of Grande-Pérez *et al.* (8) shows what is to be done to cope with this situation. Theory cannot remove complexity, but it shows what kind of "regular" behavior can be expected and what experiments have to be done to get a grasp on the irregularities. This is more true in biology than in any other field of the physical sciences. The work on the error threshold opens a new paradigm for how to fight viruses, namely not by inhibiting their replication but rather by favoring it with an increased rate of mutation. At first this procedure seems to challenge the virus to escape immune protection, but at the same time, it may cause the virus to lose all its pathogenic information. The paper quoted makes it plain that a lot of experimental work has to be done for each particular type of virus, presenting what is certainly one of the great challenges of the 21st century.

1. Eigen, M. (1971) *Naturwissenschaften* **58,** 465–523.
2. Crotty, S., Cameron, C. E. & Audino, R. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 6895–6900.
3. Loeb, L. A., Essigman, J. M., Kazazi, F., Zhang, J., Rose, K. D. & Mullins, J. I. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 1492–1497.
4. Loeb, L. A. & Mullins, J. I. (2000) *AIDS Res. Hum. Retroviruses* **13,** 1–3.
5. Domingo, E. (2000) *Virology* **270,** 251–253.
6. Holland, J. J. & Domingo, E. (1998) *Virus Genes* **16,** 13–21.
7. Eigen, M. (1993) *Med. Res. Rev*. **4,** 385–398.
8. Grande-Pérez, A., Sierra, S., Castro, M. G., Domingo, E. & Lowenstein, P. R. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 12938–12943.
9. Kimura, M. & Ohta, T. (1971) *Population Genetics* (Princeton Univ. Press, Princeton, NJ).
10. Hamming, R. W. (1980) *Coding and Information Theory* (Prentice–Hall, Englewood Cliffs, NJ).
11. Eigen, M., McCaskill, J. & Schuster, P. (1988) *J. Phys. Chem.* **92,** 6881–6891.
12. Eigen, M., McCaskill, J. & Schuster, P. (1989) *Adv. Chem. Phys.* **75,** 149–263.
13. Schuster, P. & Swetina, J. (1988) *Bull. Math. Biol.* **50,** 635–650.
14. Tarazona, P. (1992) *Phys. Rev. A* **45,** 6038–6050.
15. Eigen, M. & Nieselt-Struwe, K. (1990) *AIDS* **4** (Suppl. 1), 585–593.
16. Eigen, M., Biebricher, C. K., Gebinoga, M. & Gardiner, W. C. (1991) *Biochemistry* **30,** 11005–11018.
17. Gebinoga, M. (1991) Ph.D. thesis (Technical University of Braunschweig, Braunschweig, Germany).
18. Eigen, M. & Schuster, P. (1979) *The Hypercycle: A Principle of Natural Self Organization* (Springer, Berlin).

Molecular indetermination in the transition to error catastrophe: Systematic elimination of lymphocytic choriomeningitis virus through mutagenesis does not correlate linearly with large increases in mutant spectrum complexity

A. Grande-Pérez*, S. Sierra†, M. G. Castro*‡§, E. Domingo†, and P. R. Lowenstein*‡§¶

*Molecular Medicine and Gene Therapy Unit, Department of Medicine, University of Manchester, Manchester M13 9PT, United Kingdom; †Centro de Biología Molecular Severo Ochoa, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain; ‡Gene Therapeutics Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048; and §Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA 90095

Communicated by Manfred Eigen, Max Planck Institute for Biophysical Chemistry, Goettingen, Germany, July 18, 2002 (received for review January 5, 2002)

Studies with several RNA viruses have shown that enhanced mutagenesis resulted in decreases of infectivity or virus extinction, as predicted from virus entry into error catastrophe. Here we report that lymphocytic choriomeningitis virus, the prototype arenavirus, is extremely susceptible to extinction mutagenesis by the base analog 5-fluorouracil. Virus elimination was preceded by increases in complexity of the mutant spectra of treated populations. However, careful molecular comparison of the mutant spectra of several genomic segments suggests that the largest increases in mutation frequency do not predict virus extinction. Highly mutated viral genomes have escaped detection presumably because lymphocytic choriomeningitis virus replicates at or near the error threshold, and genomes in the transition toward error catastrophe may have an extremely short half-life and escape detection with state-of-the-art cloning and sequencing technologies.

RNA viruses replicate with an extremely high mutation rate (10^{-3} to 10^{-5} misincorporations per nucleotide copied) because of the low fidelity of RNA polymerases and reverse transcriptases, which lack proofreading repair activities (1–3). Coupled to recombination, high yields and short replication cycles, large varieties of related mutant genomes, defined as viral quasispecies, represent the genomic structure of RNA virus populations (4). The theoretical quasispecies is a steady-state organized distribution of error copies of the most fit (or master sequence) in a particular environment.

The fast generation of a large repertoire of mutant virus genomes has been interpreted as the molecular mechanism underlying the rapid evolution of RNA viruses during *in vivo* infections (5–9) and has important clinical implications. For example, it has been shown that genomes resistant to chemotherapeutic agents can already be detected in an HIV-infected person before any exposure to these agents (10). Computer simulations of the behavior of quasispecies predict that during genome replication, the mutation rate can increase only up to a threshold value without compromising the production of infectious genomes. Crossing this threshold will lead to the complete loss (or "melting") of the genetic information. This is called error catastrophe of replication, and its induction would lead to an elimination of the infection; i.e., any viral genomes produced would contain so many misincorporations that their genomes could not replicate further (11, 12). The transition into error catastrophe would convert a productive infection into an abortive one (13).

Considerable evidence suggests that riboviruses exhibit a genome mutation rate that is very close to the predicted threshold for entry into error catastrophe of replication. Consistent with this hypothesis, exposure of RNA viruses to mutagenic agents leads frequently to abortive infections. Thus, when examined, the mutation frequency of poliovirus and vesicular stomatitis virus could not be increased at defined loci by more that 3-fold by using chemical mutagenesis [e.g., 5-fluorouracil (FU), 5-azacytidine (AZC), ethylmethanesulfonate, or nitrous acid (14)]. In similar experiments the retrovirus mutation rate was increased by 13-fold after treatment with the nucleoside analog AZC (15). In another study 3′-azido-3′-deoxythymidine (AZT) was shown to increase the rate of HIV-1 mutation by a factor of 7 in a single round of replication (16).

Further evidence in support of mutagenesis-induced virus elimination has been obtained with HIV-1 by using deoxynucleoside analogs to increase viral genome mutagenesis (17). Those authors coined the term "lethal mutagenesis" to describe mutagen-induced loss of viral infectivity and suggested this could provide a new approach to treating HIV-1 and other RNA virus infections. The same authors also proposed to induce lethal mutagenesis through the incorporation of mutagenic ribonucleosides into the HIV-1 genome by the host cell RNA polymerase, rather than by the viral reverse transcriptase to avoid resistant viruses (18).

Recently, we passaged foot-and-mouth disease virus (FMDV) in the presence of FU and AZC and characterized the viral genomes through intensive sequencing analysis. Occasionally, but not systematically, FMDV was driven into extinction concomitantly with an increase in mutation frequency and mutant complexity (19). In a follow-up study we demonstrated that reducing viral loads and viral fitness systematic extinctions of FMDV were achieved accompanied by an increase in mutant spectrum complexity (20).

In a more recent study Crotty *et al.* (21) suggested that ribavirin's antiviral activity may be exerted through lethal mutagenesis. They described that ribavirin caused a large (99.3%) loss of viral genome infectivity, concomitantly with only a 9.7-fold increase in mutagenesis, and suggested that ribavirin forced poliovirus into error catastrophe (22).

Lymphocytic choriomeningitis virus (LCMV), a noncytopathic arenavirus in cell culture and a natural pathogen of wild mice, has been an important model to study persistent infections and to define the role of the immune system in viral pathogenesis (23–29). LCMV has a single-stranded RNA genome of negative polarity composed of a large (L) and a small (S) RNA segment (30). The complete nucleotide sequence reveals an ambisense coding strategy on both RNA segments encoding two proteins in opposite orientations: a nucleocapsid protein (NP) and an envelope glycoprotein

Abbreviations: 5-FU, 5-fluorouracil; AZC, 5-azacytidine; FMDV, foot-and-mouth disease virus; LCMV, lymphocytic choriomeningitis virus; GP, glycoprotein; NP, nucleocapsid protein; pfu, plaque-forming unit.

¶To whom reprint requests should be addressed. E-mail: lowensteinp@cshs.org.

 www.pnas.org/cgi/doi/10.1073/pnas.182426999

(GP) on the small RNA segment (31, 32), and an RNA polymerase and a small zinc-finger protein (Z) on the large RNA segment (30, 33). LCMV is an excellent model to study the generation and selection of viral variants in cultured cells and *in vivo*. The *in vivo* biological significance of LCMV variants has been a matter of study, but the complexity of LCMV quasispecies has not been examined in detail (34–36).

Serial passages of LCMV quasispecies in the presence of increasing concentrations of the nucleoside analogs AZC and FU demonstrated that systematic elimination of the virus can be obtained with FU in a dose-dependent fashion. Despite achieving systematic elimination of LCMV infections *in vitro*, and up to 15-fold increases in mutation frequency during viral replication in the presence of mutagens, the systematic virus elimination did not correlate with the large increases in the number of misincorporations per nucleotide, but did correlate with our inability to amplify any viral sequences.

Materials and Methods

Cells and Virus. Baby hamster kidney cells (BHK-21) were grown as described (19). Vero p68 cells were maintained in DMEM supplemented with 5% FCS and 2% glutamine. LCMV ARM 53 b is a triple plaque-purified clone from ARM CA, passaged four times in BHK-21 cells. A stock virus (p0) was prepared by infecting BHK-21 monolayers (1×10^7 cells) with 0.01 plaque-forming unit (pfu) of LCMV per cell.

Virus Infections. Conditions for a standard infection were selected to obtain high virus titers ($\geq 10^8$ pfu/ml). Semiconfluent [2×10^6 cells in 100-mm diameter dishes (Greiner, Nurtingen, Germany)] monolayers of BHK-21 were infected with 0.01 pfu per cell (0.001 pfu per cell for the low viral load infections in serial passages in the presence of AZC) in 5 ml of DMEM supplemented with 10% FCS, 10% tryptose phosphate broth, 2% L-glutamine, 0.52% glucose, and 50 μg/ml gentamicin. Viral supernatants were harvested 48 h postinfection, clarified by centrifugation at 2,500 rpm for 30 min at 4°C, and stored at −80°C.

For various preextinction virus populations (virus infectivity $\leq 10^3$ pfu/ml) (passages 1 and 2 of LCMV grown in 100 μg/ml of FU, and passage 9 of virus grown in 25 μg/ml of FU) 15 flasks of semiconfluent BHK-21 monolayers (1.6×10^7 cells per 182-cm^2 flask) were infected and incubated for 48 h in 10 ml of infection medium per flask. Virus sequences could be only amplified from passage 1 in 100 μg/ml. No specific LCMV sequences could be amplified from passage 2 in 100 μg/ml or from passage 9 in 25 μg/ml. Virus infectivity was determined by plaque assay on Vero cell monolayers as described (37). Values shown are the mean of three titrations.

Drug Treatment. Preparation of AZC and FU stock solutions, determination of BHK-21 cell viability, and procedures for infections in the presence of mutagen nucleoside analogs have been described (19).

RNA Extraction, Reverse Transcription–PCR, Nucleotide Sequencing, and Calculation of Mutant Spectrum Complexity. RNA was extracted with Trizol (Sigma) from supernatants of infection following the manufacturer's protocol. Virus suspensions with low infectivity ($\leq 10^3$ pfu/ml) were concentrated 20-fold by ultrafiltration (Centricon Plus-20 Biomax-100, Millipore) before RNA extraction, and RNA was resuspended in 1/10 the volume of control samples.

RNAs were amplified in triplicate by reverse transcription–PCR by using avian myeloblastosis virus reverse transcriptase and *Pfu* DNA polymerase (both from Promega). As a control to confirm that an excess of template was being amplified, a 1/10 dilution of the RNA was processed in parallel. cDNAs were purified with a Wizard PCR purification kit (Promega) and cloned into pGEM-T Easy Vector (Promega) that was used to transform *Escherichia coli* DH5α to obtain molecular clones. Purified cDNA and molecular clones were subjected to cycle sequencing with Big Dye chemistry (Applied Biosystems; Perkin–Elmer), and the products obtained were analyzed with an ABI 377 automated sequencer (Perkin–Elmer).

Three LCMV genomic regions were amplified by reverse transcription–PCR and sequenced by using the following oligonucleotide primers: in the small genomic RNA the GP genomic region, using primers GP247 (forward 5′-GTG GCA TGT ACG GTC TTA AGG-3′) and GP759 (reverse 5′-GGT ATT GGT AAC TCG TCT GGC-3′); and the NP region with primers NP2223 (forward 5′-GCA TTG TCT GGC TGT AGC TTA-3′) and NP2743 (reverse 5′-CAA TGA CGT TGT ACA AGC GC-3′). In the large RNA, the polymerase region was amplified with primers L3654 (forward 5′-AGT TTA AGA ACC CTT CCC GC-3′) and L4233 (reverse 5′-TGT TGA GGG TTC CAC AGA GC-3′). All of the nucleotide positions are given in the viral (genomic) sense. Sequences were retrieved from GenBank although a reconstruction of large genomic RNA was performed by using the data from the two published sequences (accession nos. M27693 and J04331).

Total cDNA (for consensus sequences) and molecular clones obtained from passages 1 and 9 of control, FU, and AZC treatments of the virus as well as p0 (our initial stock) were sequenced. The consensus sequence of p0 was compared with the published sequence and four nucleotide substitutions in the GP gene (C467 → T, G603 → A, C606 → G, and G607 → C) were found that accounted for two amino acid changes (D176 → N and R177 → A) of the GP1 protein, whereas one nucleotide substitution was found in the NP gene at position A2290 → G. In the mutagenized populations only one nucleotide substitution was found in the consensus sequence of passage 9 of 5 μg/ml AZC. All of our calculations are based on mutations found by alignment of consensus sequences compared with that of p0, and, within each quasispecies analyzed, only mutations not present in the corresponding consensus sequence were counted. An average of 15 clones per sample were sequenced (7,000–11,000 nt per sample in each genomic region were screened for substitutions, deletions, or insertions). Sequences were compared by using BLAST TWO sequences (http://www.ncbi.nlm.nih.gov/gorf/bl2.html) and SEWER (http://iubio.bio.indiana.edu/webapps/SeWeR/).

The complexity of the quasispecies was analyzed by means of two parameters: (*i*) mutation frequency and (*ii*) Shannon entropy. The first represents the number of mutations per nucleotide, and the second is a measure of the number of different molecules in the mutant spectrum of the quasispecies. To calculate the mutation frequency of each quasispecies, repeated mutations were counted only once. Shannon entropy is calculated with the formula: $-\Sigma_i [(p_i \times \ln p_i)/\ln N]$ in which p_i is the frequency of each sequence in the quasispecies and N is the total number of sequences compared (38). Shannon entropy value of 0 indicates that all of the sequences are identical whereas a value of 1 indicates that every molecule differs from the others in its nucleotide sequence.

Results

Genomic Viral Quasispecies Structure of LCMV: Experiments with FU. We characterized the genomic structure of LCMV during serial passages in the presence of the mutagenic base analogs FU or AZC. We measured and compared viral infectivity and mutation frequencies of mutagenized and nonmutagenized LCMV populations. The behavior of LCMV quasispecies in serial passages in the presence of increasing concentrations of FU is illustrated in Fig. 1*A*. Infectivity values after a single passage (given in Fig. 1*B*) show the dose–response effect of FU on LCMV replication. At low concentrations (10, 15, 20 μg/ml) LCMV replicates at a lower infectivity (≈3–4 log decrease) but it is not lost, even after 13 passages. An increase in FU to 25 μg/ml forces LCMV to move toward extinction in 11 passages. At 50, 75, and 100 μg/ml of FU, we uncovered a

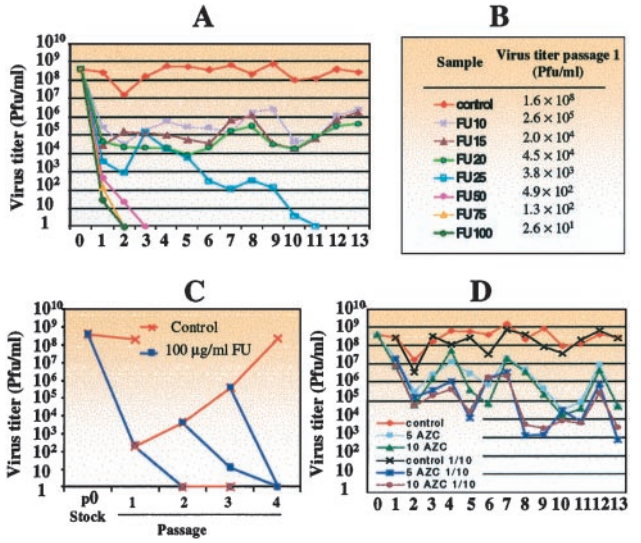
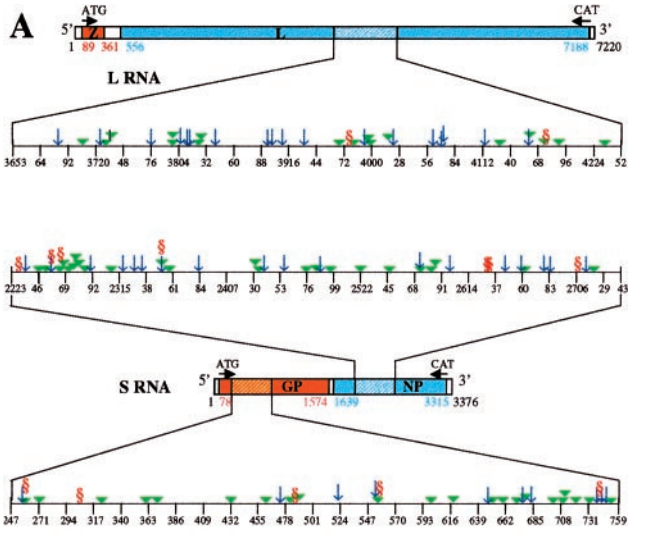


Fig. 1. Evolution of LCMV during serial passages in the presence of FU (*A–C*) and AZC (*D*). Virus serial passages were obtained by infection of BHK-21 monolayers with 0.01 pfu per cell of LCMV (p0, stock) in the absence (control) or presence of the indicated concentrations of nucleoside analogs (FU and AZC in μg/ml). Each supernatant was assayed for virus infectivity and used as inoculum for the following passage. (*A*) LCMV dose–response to FU. (*B*) Virus infectivity titers of the first passage in the presence of increasing concentrations of FU; each value is the mean of three separate experiments. (*C*) Recovery of LCMV infectivity after mutagenesis with 100 μg/ml FU. Blue lines indicate virus passages in medium containing 100 μg/ml FU; red lines represent incubations without mutagen (control). (*D*) Evolution of LCMV in the presence of AZC. Multiplicity of infection in series 1/10 of control, 5 μg/ml AZC, and 10 μg/ml AZC was 0.001 pfu/cell per passage.



B

	L		GP		NP	
	n	%	n	%	n	%
Control	2	14	5	36	7	50
FU	18	39	10	22	18	39
AZC	17	27	22	35	24	38

L (polymerase), GP (glycoprotein) and NP (nucleoprotein).

Fig. 2. (*A*) Schematic diagram showing the positions of mutations found in the genomic regions of control and mutagenized LCMV quasispecies. Each single-stranded genomic segment encodes for two proteins in an ambisense orientation. Positions of start and end codons for each ORF as well as mutations within the L (polymerase), GP, and NP genes are given in the genomic sense. Mutations (base substitutions, deletions, or insertions) found in the analyzed virus populations (see Table 1) are represented as § for untreated virus (control), ↓ for FU-treated virus, and ▼ for AZC-treated virus, irrespective of the passage at which they were found. The hashed areas represent the sequenced genomic regions. (*B*) Distribution of mutations by genomic regions in LCMV control and nucleoside analog-treated populations. Numbers (*n*) of mutations and percentages of the total for each treatment are given.

drastic effect: virus infectivity could not be detected after either two or three passages.

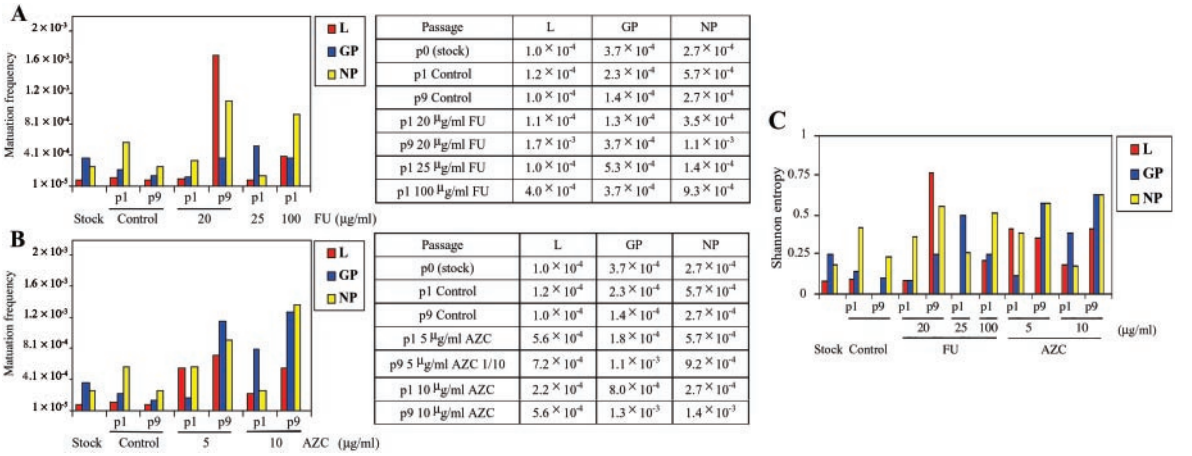
At the dose of 50, 75, and 100 μg/ml of FU, virus extinction was achieved systematically in just two passages in 3–5 separate experiments. However, total recovery of the infectivity of LCMV after a $\geq$ 6 log reduction in viral titer in 100 μg/ml of FU was obtained in three further serial passages in the absence of mutagen (Fig. 1*C*). At any recovery passage, the virus could still be eliminated systematically by further addition of 100 μg/ml of FU in 1–2 further passages. Once the virus infectivity is effectively undetectable, even serial blind passages did not allow viral titers to re-emerge, confirming complete elimination of LCMV infectivity.

Studies with AZC. Evolution of LCMV during serial passages in the presence of high doses of AZC showed that at any of the two concentrations tested virus elimination was not observed in at least 13 passages (Fig. 1*D*). The highest dose of AZC not to affect the survival of BHK cells was 10 μg/ml (19). Previously, we showed that for FMDV passaged in the presence of AZC the lowering of the viral load enhanced the frequency of extinction events (19). To test whether reducing viral load in the presence of AZC could enhance the elimination of LCMV, we serially passaged the virus at a 10-fold dilution of the multiplicity of infection of control and AZC-treated series. In these experiments a 10-fold reduction of the multiplicity of infection did not lead to systematic viral extinction over 13 passages.

Uneven Distribution of Mutations in the LCMV Genome. LCMV has two genomic single-stranded RNAs, large and small. Each one encodes for two ORFs arranged in an ambisense orientation (Fig. 2*A*). Large RNA (7.2 kb) harbors the genes for the polymerase or L protein and also a small zinc finger protein. The small RNA (3.3 kb) encodes for the GP precursor of GP1 and GP2 and NP. We have sequenced three genomic regions of molecular clones of LCMV populations passaged in the presence or absence of FU and AZC and compared with the consensus (average) sequence. The locations of the mutations found shown in Fig. 2*A* indicate that there are great differences in the number of mutations detected between regions. In naive LCMV populations (see Fig. 2*B*) mutations in the NP and GP regions are 3.5-fold and 2.5-fold more abundant than in the L (polymerase) region, indicating that the polymerase is a conserved region and suggesting that mutations in the L region could be less tolerated than in the other two. Although the overall frequency of mutations was augmented in all regions of the LCMV genome, both mutagens induced a proportionally higher increase of mutations in the L region. Strikingly, of the two mutagens tested, AZC was able to induce the highest number of mutations in all regions analyzed. Nevertheless, AZC-treated virus was not eliminated even when viral load was decreased by 10-fold (see Fig. 1*D*).

Molecular Landscape of LCMV's Genomes During Entry into Error Catastrophe. To measure the genetic complexity of the viral quasispecies in mutagenized LCMV populations we calculated the mutation frequencies and Shannon entropy values for p0 (stock) virus and passages 1 and 9 for LCMV control and mutagen-treated populations (Fig. 3).

In the absence of published data on mutation frequencies for untreated LCMV populations we did calculate those first. In control viruses, the values for any of the three regions analyzed ranged between 1.0×10^{-4} and 5.7×10^{-4} substitutions per



Passage	L	GP	NP
p0 (stock)	1.0×10^{-4}	3.7×10^{-4}	2.7×10^{-4}
p1 Control	1.2×10^{-4}	2.3×10^{-4}	5.7×10^{-4}
p9 Control	1.0×10^{-4}	1.4×10^{-4}	2.7×10^{-4}
p1 20 μg/ml FU	1.1×10^{-4}	1.3×10^{-4}	3.5×10^{-4}
p9 20 μg/ml FU	1.7×10^{-3}	3.7×10^{-4}	1.1×10^{-3}
p1 25 μg/ml FU	1.0×10^{-4}	5.3×10^{-4}	1.4×10^{-4}
p1 100 μg/ml FU	4.0×10^{-4}	3.7×10^{-4}	9.3×10^{-4}

Passage	L	GP	NP
p0 (stock)	1.0×10^{-4}	3.7×10^{-4}	2.7×10^{-4}
p1 Control	1.2×10^{-4}	2.3×10^{-4}	5.7×10^{-4}
p9 Control	1.0×10^{-4}	1.4×10^{-4}	2.7×10^{-4}
p1 5 μg/ml AZC	5.6×10^{-4}	1.8×10^{-4}	5.7×10^{-4}
p9 5 μg/ml AZC 1/10	7.2×10^{-4}	1.1×10^{-3}	9.2×10^{-4}
p1 10 μg/ml AZC	2.2×10^{-4}	8.0×10^{-4}	2.7×10^{-4}
p9 10 μg/ml AZC	5.6×10^{-4}	1.3×10^{-3}	1.4×10^{-3}

Fig. 3. Genetic heterogeneity of LCMV quasispecies replicated in the presence of mutagenic nucleoside analogs. Mutation frequencies (*A* and *B*) and Shannon entropy (*C*) were calculated for the indicated LCMV passages in the absence (control) or presence of FU (*A* and *C*) and AZC (*B* and *C*). The mutation frequencies (*A*, tables) were calculated by dividing the number of mutations found in the viral population (and not present in its consensus sequence) by the total number of nucleotides sequenced (7,000–11,000 nt). The normalized Shannon entropy (*B*) was calculated as $-\Sigma_i [(p_i \times \ln p_i)/\ln N]$ in which p_i is the frequency of each sequence in the quasispecies and N is the total number of sequences compared. Genomic regions analyzed were: L (polymerase), NP, and GP. Passage 9 in the presence of 5 μg/ml AZC is from the 1/10 dilution series.

nucleotide. After nine passages in the absence of drugs, mutation frequencies were between 1.0×10^{-4} and 2.7×10^{-4} substitutions per nucleotide. The highest mutation frequency detected corresponded to the NP region of passage 1 and the lowest was for L in agreement with previous observations (Table 1) that NP was the most variable and L the most conserved region (34–36). Shannon entropy values correlated with the detected mutation frequencies and were usually less than 0.25 except for the NP region of passage 1, which showed a higher mutant complexity (Fig. 3).

When LCMV populations were serially passaged in FU we uncovered large variations in mutation frequencies between regions. Unexpectedly, passage 9 of virus treated with 20 μg/ml FU showed a 16.8-fold increment in mutation frequency in the L region whereas in the GP region it was just 2.6-fold and 4.1-fold in the NP region. These results demonstrate, first, that it is of vital importance to analyze different areas in the genomic RNA because mutation frequencies can be region specific. If we just take into account the GP region we will assume that FU can increase the mutation frequency only 2.6-fold instead of nearly 17-fold. Second, high increments in mutation frequency in a conserved genomic region encoding for the polymerase did not abolish replication of the virus (see Fig. 1*A*). Intriguingly, the LCMV quasispecies is accepting at least 17-fold more mutations than the control in this particular region. The infectious titer of this virus population is over 10^4 pfu/ml, 3–4 logs less than the control viral quasispecies, and did not become extinct for at least four additional passages in the presence of FU (Fig. 1*A*).

We were only able to analyze molecular clones from one pre-extinction population, the one treated with 100 μg/ml FU. Just one passage away from extinction the mutation frequency of the polymerase increased 3.4-fold compared with the 1.6-fold found for the GP and NP regions.

For AZC-treated populations, values of mutation frequencies were similar for the three genomic regions and showed an increment of 5.1- to 9-fold compared with the control after nine passages for any of the concentrations tested. The L region was the less mutated of the three, suggesting that the type of mutagen used affects evolution of the virus. For all of the mutagen-treated populations Shannon entropies were higher than the controls and correlated with the variations of mutation frequencies, the maximum observed for the population treated for nine passages with 20 μg/ml FU.

Importantly, in two pre-extinction populations (100 μg/ml of FU on passage 2 and 25 μg/ml on passage 9) specific LCMV sequences could not be amplified. This failure to amplify LCMV-specific sequences could not be overcome, even after large scale-up production of potentially infectious virus. In these populations mutation frequencies possibly became so high as to include mutations in the primer binding sites, and highly mutated sequences were highly unstable, given the proximity of RNA virus replication to the error threshold (4–6, 11–14, 39–43). New methods are needed to attempt to detect and analyze highly mutated, minority LCMV genomes that may transiently be produced during the transition toward error catastrophe.

Mutations Induced by FU and AZC in LCMV Populations. The types of mutation that were found in control, FU-treated, and AZC-treated populations are shown in Table 1. The most frequent types of mutation induced by FU were transitions U → C and A → G followed by C → U, which are expected because of the chemical structure of the FU and the effect that this analog has on the cellular levels of dNTPs. In the AZC-treated populations, transversions C → G were the most frequent followed by C → A, G → C, and G → U, in agreement with previous evidence (39). It is remarkable that in just one passage in the presence of 100 μg/ml of FU we have detected 11 transitions caused by this mutagen but only two in the sample treated with 20 μg/ml, which can account for the deleterious effect of 100 μg/ml FU on the virus population. Taken together, these results confirm that these base analogs are introducing mutations in the viral RNA. However, the high mutation frequency and Shannon entropy values for the polymerase (L) genomic region of the quasispecies of passage 9 in 20 μg/ml FU did not correlate systematically and linearly with elimination of the quasispecies.

Discussion

Lethal mutagenesis induced by treatment with nucleoside analogs has been suggested as a novel way of treating HIV infections (17). The basis of this new antiviral therapy resides in the fact that increasing the already high mutation rates of RNA viruses would lead to the loss of the genetic message (melting of information) of the virus. Forcing the virus to cross the error catastrophe threshold would convert a productive into an abortive infection because of an excess of mutations. This is a transition predicted by quasispecies theory (4, 11–13). Base or nucleoside analogs, in particular FU and AZC, have been used to induce mutagenesis in several RNA viruses (14, 19, 44). Several studies have shown that nucleoside analog-induced loss of viral infectivity paralleled moderate increases in

Table 1. Types of mutations found in LCMV populations subjected to chemical mutagenesis with 5-FU and AZC

Mutation	No. mutagen			5-FU				AZC			
				20 μg/ml		25	100	5 μg/ml		10 μg/ml	
	p0	p1	p9	p1	p9	p1	p1	p1	p9	p1	p9
A → C	0	0	0	0	0	0	0	0	0	0	0
A → G	1	1	0	0	13	0	2	1	0	0	1
A → U	0	1	0	0	0	0	0	0	1	0	0
G → A	1	0	0	1	0	1	0	0	1	0	0
G → C	0	0	0	0	0	0	0	2	5	1	4
G → U	1	1	1	0	0	0	0	0	1	1	4
C → A	0	0	0	0	0	1	0	1	5	3	5
C → G	0	0	0	0	0	0	0	1	9	2	9
C → U	0	2	0	0	0	0	4	2	0	0	1
U → A	0	0	0	1	0	0	0	0	0	0	0
U → C	2	2	1	2	13	2	5	2	1	1	2
U → G	1	0	1	0	0	0	0	0	0	0	0
Deletions	0	0	0	0	1	1	1	0	1	0	0
Insertions	0	0	0	0	0	0	0	1	0	0	0
Total	6	7	3	4	27	5	12	10	24	8	26
No. of mutations per nt* ($\times 10^{-3}$)	0.24	0.2	0.13	0.16	1.1	0.2	0.52	0.45	0.91	0.39	1

Alternating shading is used to group passage numbers in the presence of individual mutagens' concentration.
*Mutations found in total (regardless of the genomic region) divided by the total number of nucleotides sequenced in each passage.

mutation frequency during single infection events (21, 22) or after serial passages (17–20). Thus, it has been suggested that RNA viruses replicate near the error catastrophe threshold, and more recent experiments have been taken as evidence that this threshold is effectively crossed by treating poliovirus with ribavirin (22). However, although in these experiments ribavirin eliminated 99% of poliovirus genome infectivity, approximately 10^7 pfu were still detected, which is higher than titers found for many other viruses even in the absence of any treatment.

We have now explored extinction mutagenesis of the prototype arenavirus, LCMV. To provide detailed evidence of the molecular landscape during the transition of RNA virus replication into error catastrophe we have analyzed LCMV viral quasispecies passaged in the presence of the nucleotide analogs AZC and FU. Serial passages in the presence of increasing doses of FU showed very different outcomes of LCMV quasispecies evolution under the selective pressure of the mutagen. Within a narrow range of mutagen concentration we observed the systematic extinction of virus passaged in concentrations of 50 μg/ml FU or above, progressive loss of infectivity toward extinction after 11 passages in 25 μg/ml FU, and a decrease in replication ability of the virus passaged in 10–20 μg/ml FU. When grown in the presence of AZC, virus extinction was not observed either in 13 passages at any concentration of AZC or by decreasing the virus load 10-fold. This finding is in contrast with previous results that documented that FMDV can occasionally be driven to extinction by AZC (19). Higher concentrations of AZC could not be used because they induced unacceptably high toxicity in BHK cells.

Nucleotide sequence analyses of the three genomic regions of virus quasispecies treated in the presence or absence of mutagen showed no correlation with virus elimination. Increases in mutation frequency in the regions analyzed preceded viral extinction, and our results do not exclude that in other genomic regions mutations accumulated in a way as to direct viral extinction. The polymerase gene of virus passaged nine times in 20 μg/ml of FU, a virus that was not eliminated by mutagen treatment, showed the highest mutation frequency, 15-fold more than passage 9 of control virus. Conversely, the virus treated with 100 μg/ml of FU, a pre-extinction population, only showed an increment of 3.4-fold for the same region compared with the control. Treatment of LCMV with 5 μg/ml or 10 μg/ml AZC was not able to drive the virus to extinction, although very high and comparable numbers of nucleotide misincorporations were detected.

Error catastrophe (4, 13, 45, 46) is predicted to involve accumulation of errors in successive replication rounds until the genomic information is entirely lost (4, 46). In this article, we show evidence that LCMV treated with 100 μg/ml FU can be systematically driven into error catastrophe. After one passage in the presence of the mutagen the virus infectivity was seriously compromised, and an additional blind passage in the presence of the drug was necessary to effectively eliminate LCMV infectivity. However, a second passage in the absence of mutagen allowed the virus to regain full infectivity as predicted by viral quasispecies theory. Although the molecular basis of entry into error catastrophe of replication can be found only in the genetic material of the virus, failure to amplify highly mutated sequences precludes such direct demonstration. Thus, our data strongly suggest that the relatively lower increases in viral mutation frequencies that can be obtained experimentally (e.g., up to 17-fold) may reflect only those genomes that can still be amplified.

Previous studies on mutagenesis of HIV and poliovirus with nucleoside base analogs calculated mutation frequencies focusing their analyses on just one genomic region. Our results have shown that mutations in the LCMV populations were not evenly distributed throughout the virus genome. We observed areas where both amounts and positions of mutations introduced by FU and AZC differ from one gene to the other. Thus, mutation frequencies from just one genomic region might lead to wrong conclusions about the

molecular scenario during mutagenesis of LCMV, and ideally the whole genome ought to be sequenced.

Differences in mutation frequencies among several genomic regions subjected to enhanced mutagenesis may be determined in part by differences in the tolerance of these individual genomic regions to additional mutations, which in turn may depend on the particular mutagenic agent used. Some of the mutated genomes analyzed might have already accumulated extremely debilitating mutations that could compromise their viability. Further studies are needed to clarify these possibilities.

In Table 1 we show the types of mutations found in the three genomic regions of each sample analyzed. Considering just the transitions induced by FU the number of mutations in the FU-treated populations would be two transitions for passage 1 of 20 μg/ml FU and passage 1 of 25 μg/ml FU, 26 transitions for the passage 9 of 20 μg/ml FU, and 11 transitions for passage 1 of 100 μg/ml FU. This finding suggests that the dose–response effect we had detected in the loss of infectivity caused by FU treatment is also found in the number of mutations typically induced by this base analog in the viral genome. This could explain that the virus treated with 100 μg/ml is driven to extinction in just two passages because the mutations introduced in the genome are proportional to the amount of mutagen available in the cell.

The most abundant types of mutation found in the AZC-treated population were transversions C $\rightarrow$ G, C $\rightarrow$ A, G $\rightarrow$ C, and G $\rightarrow$ U, which are the reported mutations induced by AZC (19). The fact that in passage 9 of 10 μg/ml AZC treatment four times more G $\rightarrow$ U mutations were found could be caused by an excess of AZC. Thus, our results indicate that the nucleoside analog treatment is effectively introducing mutations into the viral genomic LCMV RNA.

Amino acid replacements corresponding to mutations found in control as well as in FU- and AZC-treated LCMV populations are shown in Table 2, which is published as supporting information on the PNAS web site, www.pnas.org. All of the amino acid replacements were exclusive to mutagenized virus populations when compared with those observed in the control virus populations. AZC-induced amino acid replacements were more evenly distributed in the three genomic regions, whereas for FU, the L and GP regions accepted fewer amino acid replacements than NP. In particular, at passage 9 of LCMV in 20 μg/ml FU only about 30% of the mutations found in the L or GP regions led to amino acid replacement compared with the 70% observed for the NP region, indicating that amino acid changes are more tolerated in NP. The fact that only 28% of the mutations found in the L region of passage 9 in 20 μg/ml FU were nonsynonymous indicates that although this region can accept higher number of mutations the majority do not alter the amino acid sequence of the polymerase. Comparison of the degree of acceptability of amino acid substitutions (47) suggests an increase in drastic amino acid substitutions in mutagenized populations.

Our data demonstrate conclusively that an RNA virus can be systematically and efficiently driven into extinction by nucleoside analogs, presumably by introducing mutations in the viral RNA. Although the easiest interpretation of our results is that they indicate that FU drives LCMV quasispecies into extinction because of increased mutagenesis, direct evidence of melting genomic information is still lacking. We postulate that genome mutagenesis will affect primer binding sites used to amplify viral genomes and that mutation frequencies obtained thus reflect viral genomes at the gates of error catastrophe. Because RNA virus populations replicate close to error catastrophe (4–6, 11–14, 39–43, 45, 46), genomes including a substantially larger number of mutations than found in the pre-extinction populations, may have a fleeting existence, and constitute a minority in the hypermutated RNA genome population. Proof of the existence of such molecules on their way to melting of genetic information (4, 11, 13, 46) will require development of analytical techniques to gain access to such replication-incompetent intermediates.

We are very thankful to J. C. de la Torre for detailed information regarding the LCMV experimental system and primer sequence information, P. Borrow for providing us with LCMV and p68 Vero cells, N. Sevilla for primer sequence information, and L. Hall, P. Fullwood, and M. Bond for help with sequencing. This work was supported by a Sir Henry Wellcome Award for Innovative Research (053995), The Wellcome Trust (to P.R.L. and E.D.), an International Research Collaboration grant from The Wellcome Trust (049862) (to P.R.L. and E.D.), and grants from the Spanish Ministry of Education (DGES) (PM97-0060-C02-01 and BMC 2001-1823-C02–01) and Fundación Ramón Areces (to E.D.). General support was gratefully received from the British Hearth Foundation and National Health Service R & D Levy. From 1993 to 2000, P.R.L. was a Research Fellow of The Lister Institute of Preventive Medicine, Bushey Heath, U.K.

1. Steinhauer, D. A., Domingo, E. & Holland, J. J. (1992) *Gene* **122,** 281–288.
2. Temin, H. M. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6900–6903.
3. Mansky, L. M. (1998) *J. Gen. Virol.* **79,** 1337–1345.
4. Eigen, M. & Biebricher, C. K. (1988) in *RNA Genetics*, eds. Domingo, E., Holland, J. J. & Ahlquist, P. (CRC, Boca Raton, FL), pp. 211–245.
5. Domingo, E., Escarmís, C., Sevilla, N., Moya, A., Elena, S. F., Quer, J., Novella, I. S. & Holland, J. J. (1996) *FASEB J.* **10,** 859–864.
6. Domingo, E. & Holland, J. J. (1997) *Annu. Rev. Microbiol.* **51,** 151–178.
7. Domingo, E. (2000) *Virology* **270,** 251–253.
8. Holland, J. & Domingo, E. (1998) *Virus Genes* **16,** 13–21.
9. Nowak, M. A. & McMichael, A. J. (1995) *Sci. Am.* **273,** 58–65.
10. Nájera, I., Holguín, A., Quiñones-Mateu, M. E., Muñoz-Fernández, M. A., Nájera, R., López-Galíndez, C. & Domingo, E. (1995) *J. Virol.* **69,** 23–31.
11. Swetina, J. & Schuster, P. (1982) *Biophys. Chem*. **16,** 329–345.
12. Nowak, M. & Schuster, P. (1989) *J. Theor. Biol*. **137,** 375–395.
13. Eigen, M. & Schuster, P. (1979) *The Hypercycle: A Principle of Natural Self-Organization* (Springer, Berlin).
14. Holland, J. J., Domingo, E., de la Torre, J. C. & Steinhauer, D. A. (1990) *J. Virol.* **64,** 3960–3962.
15. Pathak, V. K. & Temin, H. M. (1992) *J. Virol.* **66,** 3093–3100.
16. Mansky, L. M. & Bernard, L. C. (2000) *J. Virol.* **74,** 9532–9539.
17. Loeb, L. A., Essigmann, J. M., Kazazi, F., Zhang, J., Rose, K. D. & Mullins, J. I. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 1492–1497.
18. Loeb, L. A. & Mullins, J. I. (2000) *AIDS Res. Hum. Retroviruses* **13,** 1–3.
19. Sierra, S., Dávila, M., Lowenstein, P. R. & Domingo, E. (2000) *J. Virol.* **74,** 8316–8323.
20. Pariente, N., Sierra, S., Lowenstein, P. R. & Domingo, E. (2001) *J. Virol.* **75,** 9723–9730.
21. Crotty, S., Maag, D., Arnold, J. J., Zhong, W., Lau, J. Y. N., Hong, Z., Andino, R. & Cameron, C. E. (2000) *Nat. Med.* **6,** 1375–1379.
22. Crotty, S., Cameron, C. E. & Andino, R. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 6895–6900.
23. Salvato, M. S. (1993) in *The Arenaviridae*, ed. Salvato, M. (Plenum, New York), pp. 133–156.
24. Peters, C. J., Buchmeier, M. J., Rollin, P. E. & Ksiazek, T. G. (1996) in *Virology*, eds. Fields, B. N., Knipe, D. M. & Howley, P. M. (Lippincott–Raven, Philadelphia), pp. 1521–1551.
25. Oldstone, M. B. A. (2002) *Curr. Top. Microbiol. Immunol.* **263,** 83–117.
26. Salvato, M., Shimomaye, E., Southern, P. & Oldstone, M. B. (1988) *Virology* **164,** 517–522.
27. Salvato, M., Borrow, P., Shimomaye, E. & Oldstone, M. B. A. (1991) *J. Virol.* **65,** 1863–1869.
28. Evans, C. F., Borrow, P., de la Torre, J. C. & Oldstone, M. B. A. (1994) *J. Virol.* **68,** 7367–7373.
29. Oldstone, M. B. A. (1991) *J. Virol.* **65,** 6381–6386.
30. Salvato, M. S. & Shimomaye, E. M. (1989) *Virology* **173,** 1–10.
31. Riviere, Y., Ahmed. R., Southern, P. J., Buchmeier, M. J., Dutko, F. J. & Oldstone M. B. A. (1985) *J. Virol.* **53,** 966–968.
32. Southern, P. J., Mandaleshwar, K. S., Riviere, Y., Jacoby, D. R., Buchmeier, M. J. & Oldstone, M. B. A. (1987) *J. Virol.* **157,** 145–155.
33. Salvato, M., Shimomaye, E. & Oldstone, M. B. A. (1989) *Virology* **169,** 377–384.
34. Matloubian, M., Kolhekar, S. R., Somasundaram, T. & Ahmed, R. (1993) *J. Virol.* **67,** 7340–7349.
35. Dockter, J., Evans, C. F., Tishon, A. & Oldstone, M. B. A. (1996) *J. Virol.* **70,** 1799–1803.
36. Sevilla, N., Kunz, S., Holz, A., Lewicki, H., Homann, D., Yamada, H., Campbell, K. P., de la Torre J. C. & Oldstone, M. B. A. (2000) *J. Exp. Med.* **192,** 1249–1260.
37. Ahmed, R., Simon, R. S., Matloubian, M., Kolhekar, S. R., Southern, P. J. & Freedman, D. M. (1988) *J. Virol.* **62,** 3301–3308.
38. Volkestein, M. V. (1994) *Physical Approaches to Biological Evolution* (Springer, Berlin).
39. Pathak, V. K. & Temin, H. M. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 6024–6028.
40. Drake, J. W. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 4171–4175.
41. Drake, J. W. & Holland, J. J. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 13910–13913.
42. Holland, J. J., de la Torre, J. C. & Steinhauer, D. A. (1992) *Curr. Top. Microbiol. Immunol*. **176,** 1–20.
43. Lee, C. H., Gilbertson, D. L., Novella, I. S., Huerta, R., Domingo, E. & Holland, J. J. (1997) *J. Virol.* **71,** 3636–3640.
44. Pringle, C. R. (1970) *J. Virol.* **5,** 559–567.
45. Domingo, E., Biebricher, C., Holland, J. J. & Eigen, M. (2001) *Quasispecies and RNA Virus Evolution: Principles and Consequences* (Landes, Austin, TX).
46. Schuster, P. & Stadler, P. F. (1999) in *Origin and Evolution of Viruses*, eds. Domingo, E., Webster, R. G. & Holland, J. J. (Academic, San Diego), pp. 1–24.
47. Feng, D. F., Johnson, M. S. & Doolittle, R. F. (1985) *J. Mol. Evol*. **21,** 112–125.