Vol. 79, No. 3

Computational Design of Antiviral RNA Interference Strategies That Resist Human Immunodeficiency Virus Escape†

Joshua N. Leonard and David V. Schaffer*

The Department of Chemical Engineering and The Helen Wills Neuroscience Institute, University of California, Berkeley, California

Recently developed antiviral strategies based upon RNA interference (RNAi), which harnesses an innate cellular system for the targeted down-regulation of gene expression, appear highly promising and offer alternative approaches to conventional highly active antiretroviral therapy or efforts to develop an AIDS vaccine. However, RNAi is faced with several challenges that must be overcome to fully realize its promise. Specifically, it degrades target RNA in a highly sequence-specific manner and is thus susceptible to viral mutational escape, and there are also challenges in delivery systems to induce RNAi. To aid in the development of anti-human immunodeficiency virus (anti-HIV) RNAi therapies, we have developed a novel stochastic computational model that simulates in molecular-level detail the propagation of an HIV infection in cells expressing RNAi. The model provides quantitative predictions on how targeting multiple locations in the HIV genome, while keeping the overall RNAi strength constant, significantly improves efficacy. Furthermore, it demonstrates that delivery systems must be highly efficient to preclude leaving reservoirs of unprotected cells where the virus can propagate, mutate, and eventually overwhelm the entire system. It also predicts how therapeutic success depends upon a relationship between RNAi strength and delivery efficiency and uniformity. Finally, targeting an essential viral element, in this case the HIV TAR region, can be highly successful if the RNAi target sequence is correctly selected. In addition to providing specific predictions for how to optimize a clinical therapy, this system may also serve as a future tool for investigating more fundamental questions of viral evolution.

An estimated 40 million people worldwide are living with human immunodeficiency virus (HIV), which infected 5 million people and caused 3 million deaths in 2003 (65). In the United States alone, nearly 900,000 people are HIV positive (16). While the advent of highly active antiretroviral therapy (HAART) has made great strides in delaying the onset of AIDS, this drug cocktail approach has proven inadequate as a cure and is associated with a host of severe side effects, including a number of metabolic disorders, which can undermine patient compliance (61, 63). Furthermore, the growing problem of viral resistance to antiviral drugs threatens to reverse the progress that has been made in reducing AIDS-related morbidity and mortality (52). Still more ominous is the observation that resistant viral strains are emerging more rapidly than new drugs are being developed. Finally, the development of an effective AIDS vaccine has been elusive (17).

A promising alternative to HAART and vaccine development is the use of genetic therapies to inhibit viral reproduction by acting directly on genetic sequences. Antisense, ribozyme, and RNA decoy-based approaches have been studied for over a decade (reviewed in reference 45), and one approach using a lentiviral vector to express antisense RNA against the HIV envelope is currently undergoing phase I clinical trials (28, 43). This clinical strategy involves harvesting peripheral CD4+ T cells, transducing them with a gene therapy vector, and reintroducing them into the body. Alternatively, peripheral hematopoietic CD34+ cells could be harvested, transduced, amplified, and reintroduced into the bone marrow. The latter type of approach could be used to establish a continuous supply of multiple types of resistant cell populations, since CD34+ cells can differentiate into macrophages, T cells, and dendritic cells after being transduced with lentiviral vectors (3, 8, 21).

More recently, RNA interference (RNAi) has emerged as a robust and highly evolutionarily conserved mechanism for down-regulating gene expression through targeted mRNA cleavage (15; reviewed in reference 9). In this process, the RNase III family endonuclease DICER cleaves double-stranded RNA into duplexes of 21 to 22 bp, called short interfering RNA (siRNA). The antisense strand is then used as a guide by the RNAi-induced silencing complex to direct specific cleavage of mRNAs that share sequence identity with the siRNA (50). Recently, it was shown that siRNAs can be introduced directly into mammalian cells, either via synthetic delivery vectors conjugated to siRNA molecules or by using gene transfer vectors to deliver siRNA expression cassettes, to induce RNAi (41, 49). Such approaches have already been shown to block the in vitro replication of important pathogens, including poliovirus (20), hepatitis B and C viruses (reviewed in references 51), influenza virus (19), and HIV type 1 (HIV-1) (reviewed in references 4 and 40). It was also recently shown that an siRNA that reduces the expression of cellular coreceptors for HIV can effectively block viral entry (2). Furthermore, a combination of antiviral siRNAs, TAR decoy RNAs, and ribozymes directed against the HIV coreceptor CCR5 induced potent inhibition of multiple steps in the HIV life cycle (42).

Despite this early promise, the clinical application of RNAi

* Corresponding author. Mailing address: Department of Chemical Engineering and The Helen Wills Neuroscience Institute, University of California, 201 Gilman Hall, Berkeley, CA 94720-1462. Phone: (510) 643-5963. Fax: (510) 642-4778. E-mail: Schaffer@cchem .berkeley.edu.

faces a number of challenges. Recently, it was shown that certain types of siRNA can induce considerable nonspecific or off-target changes in gene expression, particularly when high dosages are used (36, 55, 60). Thus, it is possible that siRNA may be safely utilized only within a limited therapeutic window. More importantly, the exquisite sequence specificity of RNAi makes the problem of viral escape a significant potential obstacle, particularly when its target is a highly mutable entity such as HIV. This concern was recently exemplified when two in vitro studies demonstrated that HIV can evolve resistance to RNAi (6, 7). Together, these findings indicate that optimization is required in the design of effective antiviral RNAi therapies, and computational and theoretical approaches can aid in the analysis of antiviral therapies (10, 47, 64).

Previous theoretical studies of the general processes of viral evolution have yielded important findings; however, they have been centered around models that are significantly abstracted or otherwise removed from the specific case of HIV (compiled in reference 13), though some recent theoretical studies have incorporated various degrees of detail of retroviral biology (47, 53, 68). The most widely used HIV-specific model is based upon a set of deterministic differential equations that treat viruses and cells as compartmentalized populations (reviewed in reference 39). While these studies have contributed significantly to understanding general principles of viral evolution and HIV infection dynamics, these approaches are not suitable for studying HIV's evolution around RNAi, a phenomenon that is likely to be dominated by stochastic processes and for which specific molecular details must be considered.

In this study, we used a computational model of HIV replication to elucidate design principles for applying RNAi against HIV in a manner that delays or prevents viral escape. To realistically model this complex system, we have developed a novel type of computer model: a mechanistic agent-based stochastic simulation (MASS) that incorporates molecular-level details of HIV reproduction and susceptibility to RNAi. Here, we report our computational results on the predicted therapeutic efficacy of combination RNAi approaches, the importance of efficient siRNA delivery, and the significance of targeting RNAi to essential elements of the viral genome such as the TAR RNA domain. Our approach is appropriate for the analysis of combination therapies that incorporate RNAi (or other effector molecules) directed against relevant cellular targets such as cyclinT1 and CDK-9, which are required for Tat function (72). However, we have focused this initial study on the exclusive use of RNAi directed against viral sequences. The results of this quantitative approach highlight the importance of carefully designing RNAi therapies for the optimal inhibition of viral replication.

MATERIALS AND METHODS

Model development overview: MASS. We have constructed the first stochastic simulation, to our knowledge, that incorporates the detailed molecular-level mechanisms of retroviral infection and reproduction. Individual viral genomes are represented as discrete "agents," that is, data sets that encode viral sequences and properties of the corresponding viral particles. In each cycle of viral replication, a large ensemble of these virions randomly infects a population of cells and is then subjected to a series of steps that emulate the processes of reverse transcription (in which the two viral RNA genomes from each virion are converted into a single, chimeric, double-stranded DNA genome), integration of viral DNA into the host genome, and transcription of viral genomic RNA. This

infected cell population stably expresses a variety of antiviral siRNA molecules, as further described below. The fates of cells within this population are recorded and tracked individually to account for cell-to-cell variability in events, including expression of siRNAs and genetic mixing between viruses that coinfect a single cell. The cycle ends as the viral genomes in a cell are randomly paired and packaged into the next generation of diploid virions, at which point the process is repeated with a new population of uninfected cells.

The cell culture analog of a simulation cycle would be the serial transfer of viral supernatant onto uninfected cells at fixed time intervals. Conceptually, one simulation cycle also emulates the process of infectious particle release, death of previously infected cells, and infection of naïve cells in vivo. Each simulation was initiated by infecting a population of 15,000 cells with 150,000 wild-type virions, which led to a high degree of coinfection and, hence, genetic mixing in the subsequent generation. In addition, as is explained below, a viral population of this size is sufficiently large to guarantee that all single mutations and many double or higher mutation combinations are generated when the initial viral population undergoes reverse transcription for the first time. Consequently, the simulated situation is more analogous to the initiation of RNAi treatment in a patient who already harbors a large viral load than to the case of prophylactic RNAi therapy during early stages of infection in vivo, where the ratio of viruses to cells is 3 to 5 orders of magnitude lower than that used here (43). All steps were assumed to reach a steady state within each cycle, such that the dynamics during a single cycle were not relevant.

A number of steps during retroviral reproduction are determined by chance and were represented accordingly in the simulation. Viral infection of cells was implemented as a Poisson process, in which each virus randomly selected a cell to infect and multiple viruses could simultaneously infect any single cell. In addition, HIV rapidly generates genetic diversity during reverse transcription by producing on average of ~0.3 new point mutations and recombining (between its two RNA template genomes) ~3 times per cycle of replication (33, 44). While reverse transcribing, the simulation used these frequencies to randomly decide whether to point mutate and/or switch RNA templates (recombine) at every base copied into viral cDNA. Since the HIV genome is approximately 9 kb in length, our initial infection size of 150,000 virions yields ~5-fold coverage of single base mutations. Furthermore, the site at which HIV integrates into the host cell's genomic DNA influences the level of basal viral gene expression (34, 57). In our simulations, integration is modeled by randomly assigning an integration site factor according to a binomial distribution on the range {0 . . . 1}, which is proportional to the basal (Tat-free) transcription level (see the supplemental material for details).

Since the model accounts for mutation and recombination, the viral sequences rapidly and randomly diversify as the simulation progresses to generate a large, genetically mixed population often referred to as a viral quasispecies or swarm (12, 13, 35). As described in greater detail in the next section, some of these viral sequences can acquire mutations that render them resistant to the RNAi. Since this viral escape is a stochastic event, viral population dynamics vary between trials, even for a fixed set of conditions. However, after a sufficient number of cycles, each individual simulation results in either an "extinction," wherein the virus population drops to zero, or an "escape," in which case the virus population grows to some threshold level indicating that RNAi-resistant viruses have emerged and established an expanding infection. Thus, the antiviral efficacy of a given scenario was defined as the percentage of trials that resulted in an extinction of the virus population, i.e., successful inhibition by RNAi. In this study, each antiviral efficacy data point represents the outcome of at least 200 independent trials.

RNAi-mediated cleavage and viral escape. It has been shown by studies in mammalian cells that the level of RNAi-mediated cleavage induced in a cell depends on the dosage of siRNA delivered to (or expressed within) that cell (22, 56, 67). In our model, the level of RNAi induced in a given cell was represented as P_cleavage, which is the probability that a single locus of a given RNA transcript will be cleaved by the RNAi pathway. Depending on the experiment, cells within a population expressed uniform or differing P_cleavage values.

Based on various published results (24, 26, 46), we postulated a model for viral escape from RNAi in which a single base mismatch between the viral genome and siRNA sequences resulted in a reduced chance of degradation. A single mismatch near the center of the target reduced RNAi-mediated cleavage probability by 50%, a mismatch at an end of the target sequence reduced cleavage probability by 10%, and intermediate mismatches were assigned an intermediately reduced cleavage probability following a second-order polynomial. Two or more mismatches were scored as total immunity to RNAi (see the supplemental material), though this relationship could readily be adjusted to fit new experimental data.

Combination RNAi. For simulations in which multiple viral sequences were simultaneously targeted, we considered a situation in which RNAi efficacy was

divided equally between targets. Thus, if two viral loci were targeted simultaneously, each locus would be degraded one-half as effectively as would a single locus in the one-target case. An experimental analogy would be to use equimolar mixtures of different types of equally effective siRNA duplexes, where the total concentration of siRNA molecules was kept constant. It should be noted that under this assumption, the effective overall RNAi cleavage probability of a transcript is reduced by dividing the RNAi between multiple loci due to multiple cleavages of a single RNA molecule. For example, if the one-target RNAi cleaves 80% of target RNAs, the two-target RNAi cleaves only 64% of the RNAs (see the supplemental material). This model should be considered a worst-case scenario from a clinical standpoint, since cleaved transcripts are likely to be rapidly degraded and, therefore, this multiple cleavage problem might be less significant than our assumption asserts. We did not consider the case of asymmetrically divided RNAi (i.e., nonequimolar mixtures of different siRNA duplexes), but such a case is also addressable by this method.

Tat/TAR feedback. For simulations incorporating the HIV Tat/TAR feedback loop, a set of differential equations was derived to describe the relationship between the concentrations of viral genomes and Tat within a single cell (see the supplemental material). The steady-state solution of these equations was calculated for each cell during each cycle of replication. It should be noted that while the solution of these equations was deterministic for a given set of variables, the variables in the equations were stochastically determined for each viral genome and each cell. Though the TAR loop may play various roles in viral reproduction (18, 29, 48), we considered only its primary function, that of Tat-mediated transcriptional activation (transactivation) (14). These equations incorporated experimental data describing the relationship between basal HIV gene expression (presumably due to variable viral integration sites) and Tat-transactivated gene expression (34). The full derivation is described in the supplemental material, as are all relevant constants. The steady-state implicit solutions for the concentrations of viral RNA genomes derived from each provirus (G_i) and the total (G_T) are given by the following two equations:

$$G_i = CM_1(1 - P\_\text{cleavage})\left(\frac{IF_i}{1 - EMF_i}\right)\{1 + AMF_i[a\exp(-r1F_i) + b - 1]$$

$$[1 - \exp(-sCM_2G_T)]\} \qquad (1)$$

$$G_T = \sum_{i=1}^{\#\_proviruses\_in\_cell} G_i \qquad (2)$$

The genetic sequence and integration site of each viral genome is tracked individually, as indicated by the subscript i in various parameters. As described above, the basal (Tat-free) viral transcription level of a given provirus is dependent upon the site at which it integrates into the host genome, and IF_i (the integration site factor) is a nondimensional factor that characterizes this basal level. In each HIV RNA genome in the simulation, the capped 5′ G was numbered as position 1 and the entire TAR loop comprised positions 1 to 59, whose sequence was taken from the archetypical isolate pNL4-3 (1). The functions of HIV's TAR region have been analyzed in a wide variety of studies that employed detailed mutational analysis and forced evolution techniques (5, 11, 14, 18, 22, 25, 31, 58, 59, 69). In these, it was shown that bases 18 to 44 compose the essential TAR element necessary for Tat transactivation (25, 31, 58), and in our model, only mutations in this region were assumed to affect TAR function. Consequently, we did not include the role that upstream sequences within the TAR region play at the proviral DNA level (18, 48). Based on these previous studies, a set of quantitative rules was devised to describe the effect that mutations in various regions of the essential TAR element have on the ability of Tat to transactivate genomes carrying such mutations. Mutations in the *tat* gene were not considered. AMF_i (the amplification mutation factor) is a nondimensional factor that incorporates these rules into the governing equations. As discussed above, mutations in the target sequence of viral RNA can lead to reduced RNAi-mediated cleavage, and EMF_i (the escape mutation factor) is a nondimensional factor that incorporates these effects. A complete description of the calculation of these factors is included in the supplemental material.

A more subtle detail is that though our model treats the TAR element as a single locus, an HIV RNA genome actually contains two TAR sequences (one in each of the 5′ and 3′ long terminal repeats). The 5′ TAR plays the primary role in Tat-mediated transactivation, and due to the mechanism of reverse transcription, the 5′ TAR is preferentially used as a template for proviral TAR DNA. Even in the rare event that the 3′ TAR RNA is used as template for proviral DNA, the resulting provirus will still have two identical TAR sequences, rendering it indistinguishable from a provirus from a single 5′ TAR RNA template (38). For example, any unique mutations in the 3′ TAR RNA will be either lost

or, much more rarely, copied into both the 5′ and 3′ proviral TAR DNA loci after a single round of viral replication. In practice, this means that a virus harboring RNAi resistance-conferring mutations in the 3′ TAR RNA alone would only rarely and transiently replicate differently than would a virus harboring such mutations in both TAR sequences. Consequently, the treatment of TAR as a single locus is not expected to significantly impact our results either qualitatively or quantitatively.

This model contains only two adjustable parameters: CM_1 and CM_2. CM_1 influences the overall transcription level (and, therefore, the reproductive rate) of the viral genomes. It is related to the average basal (Tat-free) expression level of a viral genome. CM_2 indirectly determines the number of cells that, on average, will attain saturating or supersaturating concentrations of Tat protein, and increasing CM_2 makes the viral population, on average, less sensitive to Tat concentrations. The experiments in this study utilized values of CM_1 and CM_2 that yielded a rate of viral population growth that was consistent with our in vitro observations of a replication-competent HIV-based model vector (data not shown). Thus, we selected a moderate reproductive ratio, the multiplicative factor by which population size increases in each generation, of ~3-fold per cycle. Note that this value is not equivalent to the burst size, the average number of virions that a single infected cell produces before it dies. A range of combinations of CM_1 and CM_2 were tested, and the conclusions described below were not qualitatively sensitive to variations in these constants (see the supplemental material for a more detailed discussion of these constants).

Machines used. Simulations were run on either a PC with a Pentium 4 processor (1.7 GHz) or an Apple PowerPC G5 with dual 2-GHz processors. The simulations used either the Mersenne Twister (http://www-personal.engin .umich.edu/~wagnerr/MersenneTwister.html) or the RanrotB (http://www .agner.org/random/) random number generators, both of which are freely available online.

RESULTS

In the computational or theoretical study of evolution, one must choose a meaningful way to relate genetic changes to changes in reproductive ability (or fitness). In a limiting theoretical model of population genetics pioneered by Kimura, the neutral theory of evolution, it is assumed that the majority of mutations have negligible impacts on the fitness of the organism (or virus) (37). In reality, however, some mutations are likely to be more deleterious than are others, as indicated by the presence of highly evolutionary conserved regions (13). In our study of HIV's evolution in the presence of RNAi, a neutral evolution assumption is complicated by the fact that the function of each viral sequence targeted could impact viral fitness in a complex and nonobvious manner operating at the DNA, RNA, and/or protein level. Therefore, as an initial base case scenario, we simulated the evolution of viral genomes containing no functional sequences—only RNAi targets. Thus, for these neutral viral genomes, no negative effects were attributed to mutations, which served only to benefit the virus by facilitating escape from RNAi. We later built upon this neutral viral genome model by incorporating the deleterious effects that viral mutation has on an important viral function, the Tat/TAR transcriptional feedback loop.

Neutral viral genomes base case. (i) Combination RNAi. It is possible that viral resistance to RNAi might be circumvented by targeting multiple viral sequences simultaneously, thus reducing the chance that a fully resistant mutant could arise (4, 40). In clinical practice, however, there will probably exist some upper bound on the amount of siRNAs that can be delivered to or expressed in a cell. Transfer of presynthesized siRNAs to cells will be limited by delivery mechanisms and biological clearance, similar to the case of traditional drug delivery. Likewise, transduction of target cells by gene therapy vectors carrying shRNA or siRNA expression cassettes can

also be limited at several steps, both in vitro and in vivo (66). In addition, there is some evidence that multiple types of siRNA can compete for cellular RNAi machinery (26) and that high levels of certain siRNAs can induce off-target or nonspecific changes in gene expression via an inflammatory response (36, 55, 60). Thus, we first investigated how the use of an increasing number of RNAi targets with a fixed overall siRNA level would impact antiviral efficacy.

We first modeled a cell population in which every cell expressed siRNA such that the overall RNAi level, or P_cleavage (defined above), was the same in each cell. We compared situations in which the siRNA was directed against either one viral sequence or multiple, evenly spaced targets in the HIV genome, with P_cleavage held constant. As discussed above in "Model development overview," a worst-case scenario was used to model the application of multiple types of siRNAs—under this assumption, the cleavage efficiency of a single viral genome decreases incrementally with the number of RNAi targets used, due to the possibility of cleaving a single transcript more than once and thus wasting RNAi resources (described further in the supplemental material). Despite this fact, for every overall RNAi level tested, increasing the number of targets significantly increased the antiviral efficacy (Fig. 1A). Of particular interest is the observation that for a given overall RNAi level, the efficacy rose dramatically over a small change in the number of targets used. Moreover, as the overall RNAi level was increased, fewer targets were required to achieve this marked increase in efficacy. We also observed that while higher levels of RNAi more effectively led to viral extinction, in the cases where viral escape did occur, it occurred more rapidly under strong RNAi pressure (Fig. 1B).

(ii) Nonuniform RNAi expression. In the previous experiment, all cells in the population were assumed to express RNAi in an identical manner. However, in both the laboratory and future clinical practice, RNAi levels are likely to vary between cells within a single population due to limitations on the delivery of siRNA or siRNA-expressing expression vectors (66). To investigate the consequences of this heterogeneity, we first performed a series of simulations in which the cell population consisted of two distinct subpopulations—naïve cells (expressing no siRNA) and cells expressing siRNA (directed against a single viral target sequence) at a uniform level.

As shown in Fig. 2A, increasing the fraction of naïve cells, while holding P_cleavage constant in siRNA-expressing cells, led to a monotonic decrease in the antiviral efficacy, as would be anticipated. Interestingly, it was found that all data collapsed onto a single curve when the outcomes were plotted versus the mean P_cleavage of the cell population (the product of the fraction of cells expressing siRNA and the P_cleavage characterizing the RNAi level in this fraction) (Fig. 2B). Thus, the mean RNAi level was the most significant determinant of efficacy in this experiment.

To further investigate how cell-to-cell variability in RNAi impacts antiviral efficacy, we next posed a more realistic description of RNAi expression heterogeneity. Since we and other researchers have observed that the level of RNAi-mediated cleavage is strongly dependent on the dose of siRNA delivered or expressed (22, 56, 67), a series of simulations in which RNAi expression within a cell population varied according to a binomial distribution (the discrete analog of the
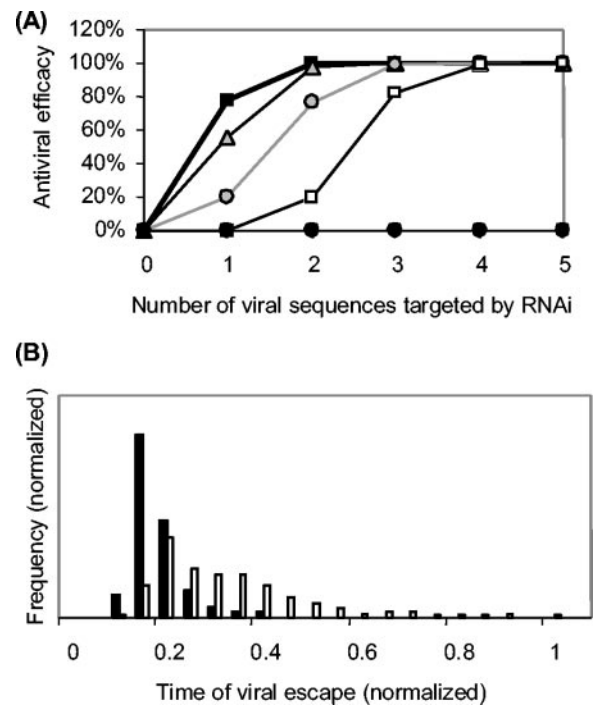

FIG. 1. Antiviral efficacy of the combination RNAi approach. (A) Evolution of neutral viral genomes simulated under conditions of combinatorial RNAi inhibition. In each series, the overall RNAi cleavage probability on a per-locus basis (P_cleavage) was held constant, while an increasing number of evenly spaced viral sequences was targeted for degradation. In this and all other figures, antiviral efficacy refers to the percentage of trials resulting in an extinction of the viral population. Each data point represents the result of at least 200 independent simulation trials. Symbols for different values of P_cleavage are as follows: 0.9 (black squares); 0.8 (grey triangles); 0.75 (grey circles); 0.7 (white squares); 0.6 (black circles). (B) Distribution of viral escape times for two values of P_cleavage: 0.8 (black) and 0.7 (white). Time until viral escape (in number of simulation cycles) was normalized to the maximum value depicted. Frequency was normalized to the total number of escape events.

Gaussian distribution) was run. The antiviral efficacy was compared for populations sharing the same mean RNAi level and differing only in the breadth (standard deviation) of the RNAi distribution (Fig. 3A).

As the RNAi distribution breadth increases, the cell population includes more cells of both higher and lower RNAi expression. Interestingly, for a mean P_cleavage of 0.8 the efficacy decreased with increasing breadth, whereas for a mean P_cleavage of 0.85 increasing breadth caused an increase in efficacy. Finally, when increasing the breadth for a mean P_cleavage of 0.75, the series passed through a minimum in efficacy; i.e., the efficacy first decreased then increased with increasing breadth (Fig. 3A). In contrast to the bipartite cell population scenario, the data series did not collapse onto a single curve when plotted against any parameter we tested (data not shown). Most significant is that the observation that efficacy did not appear to be highly sensitive to RNAi distribution breadth for any mean P_cleavage tested.

Functional viral genomes: Tat/TAR transactivation. As the next biologically realistic step beyond the noncoding, neutral genome base case, we incorporated the essential HIV function
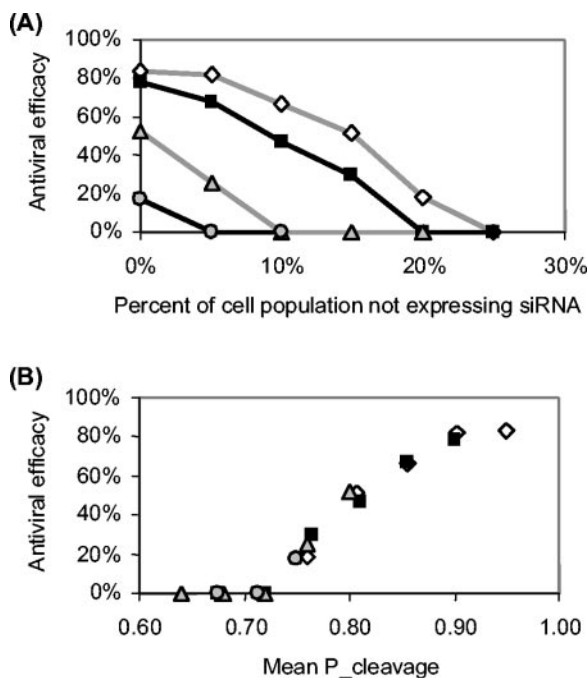
FIG. 2. Nonuniform RNAi: bipartite cell population and neutral viral genomes. (A) Evolution of neutral viral genomes in a cell population consisting of a mixture of cells expressing either uniform levels of siRNA (corresponding to P_cleavage) or no siRNA. In each series, P_cleavage was held constant while the fraction of nonexpressing cells was increased. (B) Data points from panel A plotted against the mean P_cleavage of the population, calculated by multiplying the value of P_cleavage in siRNA-expressing cells by the fraction of cells that expressed siRNA. Symbols for different values of P_cleavage are as follows: 0.95 (white diamonds); 0.9 (black squares); 0.8 (grey triangles); 0.75 (grey circles).
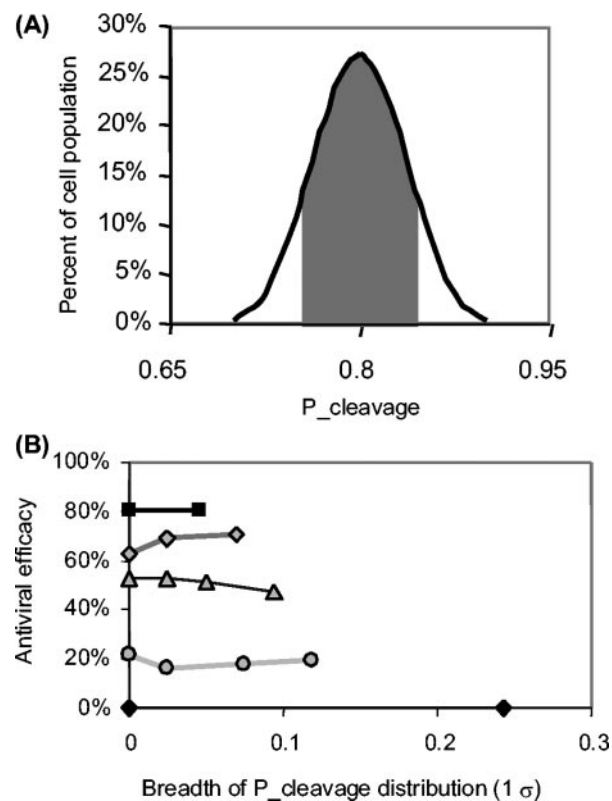

FIG. 3. Nonuniform RNAi: binomial distribution of siRNA expression and neutral viral genomes. (A) Depiction of binomial distribution, centered about a P_cleavage of 0.8. Shaded region is ±1 standard deviation (σ). (B) Evolution of neutral viral genomes on a cell population with binomially distributed siRNA expression. In each series, mean P_cleavage was held constant while the standard deviation of the distribution (σ) was increased. The right-most data point in each series represents the widest possible distribution for that particular value of mean P_cleavage. Symbols for different values of mean P_cleavage were as follows: 0.9 (black squares); 0.85 (grey diamonds); 0.8 (grey triangles); 0.75 (grey circles); 0.5 (black diamonds).

of Tat/TAR-mediated transcriptional activation. Though the TAR sequence contains a high degree of secondary structure, it has been shown to be at least partially susceptible to RNAi-mediated cleavage (30, 71). In addition, TAR is an attractive target because it exerts it functions directly through RNA, as described below, rather than by coding for a protein. Therefore, there is no need to postulate a relationship between protein sequence and protein function, which could lessen the quantitative nature of this study. Furthermore, this case can readily be modeled quantitatively using the plethora of empirical data that describe the effects of TAR mutation on viral function (5, 11, 14, 18, 22, 25, 31, 58, 59, 69). It can be anticipated that if a virus suffers fitness losses by mutating within a functionally important region, then this region is an attractive therapeutic target. Every HIV transcript begins with the noncoding TAR RNA element at its 5′ terminus, where the HIV-encoded Tat protein binds and subsequently recruits the cellular cofactors cyclin T1 and Cdk9. This protein complex subsequently hyperphosphorylates the large subunit of cellular RNA polymerase II, leading to enhanced transcriptional elongation (72). This Tat-TAR interaction yields a nonlinear positive transcriptional feedback loop and is crucial for viral propagation. Studies using directed mutagenesis of TAR have demonstrated that a small number of mutations is sufficient to render the virus replication incompetent (62). Therefore, targeting TAR with RNAi would force the virus to negotiate a delicate balance between loss of function and susceptibility to RNAi-mediated degradation.

In this set of simulations, we considered RNAi directed against various portions of TAR, though mutations in the *tat* gene were not considered. Based upon the literature (5, 11, 14, 18, 22, 25, 31, 58, 59, 69), a set of rules was formulated to describe the consequences of TAR mutations on the ability of Tat to stimulate transcriptional elongation (see the supplemental material). As might be expected, RNAi targeting the highly conserved "essential" TAR loop (bases 18 to 44) led to more efficient viral extinction than did that targeting nonfunctional sequences in the region, and targets that spanned the two types of region produced intermediate efficacies (Fig. 4). However, not all RNAi targets within the essential TAR loop were equally efficacious, and for lower overall levels of RNAi the antiviral efficacy became extremely sensitive to target location choice. Intriguingly, at one overall RNAi level tested, a shift in target sequence by only a few bases was sufficient to switch from completely efficacious to completely inefficacious. Finally, we noted that the antiviral efficacy appeared to be
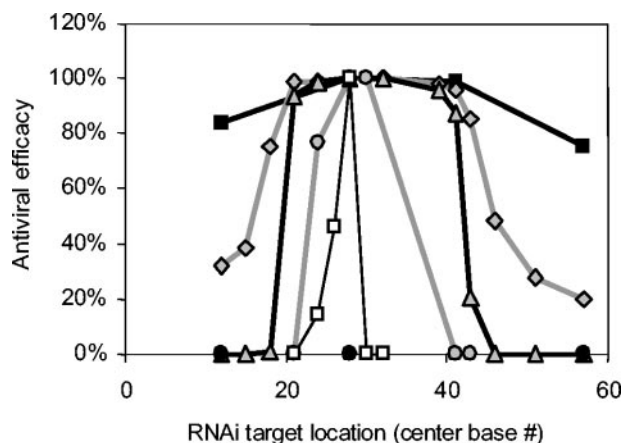
FIG. 4. Choice of RNAi target sequence in TAR, as shown in Tat/TAR functional viral genomes reproduced on cell lines that uniformly expressed siRNA against a single region of the TAR RNA. In each series, different 21-bp target sequences were used for a fixed value of P_cleavage, and antiviral efficacy was recorded; the location of the target sequence is reported as the center base of the target. Symbols for different values of mean P_cleavage were as follows: 0.9 (black squares); 0.85 (grey diamonds); 0.8 (grey triangles); 0.78 (grey circles); 0.77 (white squares); 0.76 (black circles).

more sensitive to the overall RNAi level than was the case for the noncoding genomes. This was most true at intermediately efficacious sites (Fig. 4).

We next tested the effects of heterogeneous RNAi expression on the replication of functional HIV genomes that possessed an active (at least before mutation) Tat/TAR loop. Based on the results on the effects of target location choice (Fig. 4), we selected three representative target sequences for analysis: a highly conserved site (bases 18 to 38), an intermediately conserved site (bases 31 to 51), and a nonconserved site (bases 47 to 67). We first simulated the previously described bipartite cell population, in which RNAi was expressed uniformly in one subpopulation, whereas the other subpopulation expressed no RNAi (naïve cells).

As shown in Fig. 5A to C, the antiviral efficacy again decreased as the fraction of cells lacking RNAi increased. Consistent with observations for the uniform cell population case (Fig. 4), the antiviral efficacy appeared to be more sensitive to both the overall RNAi level and to small changes in the fraction of naïve cells than was the case for neutral viral genomes (Fig. 2). In contrast to the neutral viral genome case, the antiviral efficacies did not collapse to a single curve when plotted against the mean P_cleavage of the cell population (Fig. 5D to F). In particular, for two simulations with the same mean population RNAi level, the antiviral efficacy was always higher when the fraction of naïve cells was smaller, indicating that the presence of naïve cells more adversely impacts antiviral efficacy against Tat/TAR functionalized genomes than was the case for neutral genomes. This lack of collapse was most severe when targeting a highly conserved sequence (bases 18 to 38), but collapse to single curve was partially obtained when targeting an intermediately conserved (bases 31 to 51) or nonconserved (bases 47 to 67) sequence (Fig. 5D to F).

We next simulated the replication of the functional Tat/TAR virions on a cell population in which RNAi expression was

binomially distributed. As previously discussed, this formulation more realistically emulates heterogeneity in siRNA delivery or expression. We chose the same highly conserved, intermediately conserved, and nonconserved RNAi targets as above (Fig. 5). In each case, a series of trials was performed where the siRNA expression distribution was increased in breadth (standard deviation) while holding the mean RNAi level constant, and we found that the consequences of increasing this breadth differed depending on the viral sequence targeted (Fig. 6). When targeting the most highly conserved sequence, efficacy increased with breadth of the distribution over a narrow range of P_cleavage values; outside of this range, efficacy was either 100 or 0%, independent of siRNA heterogeneity (Fig. 6A). When targeting a moderately conserved sequence, efficacy again increased with distribution breadth for lower values of P_cleavage. However, for larger values of P_cleavage, efficacy either decreased slightly or was unchanged as breadth increased (Fig. 6B). Finally, when targeting a nonconserved sequence, efficacy increased with breadth of the RNAi distribution unless P_cleavage was too low to have any antiviral efficacy (Fig. 6C). As was the case when targeting neutral genomes, these data series (for a given target sequence data set) did not collapse onto single curves when plotted against any parameter combination we tested (data not shown).

DISCUSSION

The increasing prevalence of multiply drug-resistant HIV strains, as well as continued problems with patient compliance due to side effects, has highlighted the need to improve or replace the use of HAART drug cocktails (52). RNA-based gene therapy offers a promising alternative, because side effects and the need for strict patient compliance to a drug regimen could be avoided. Also, such a strategy does not require the expression of foreign and potentially immunogenic proteins. In particular, RNAi provides a robust method for specifically inhibiting the expression of targeted cellular or viral genes, and it shows promise as a novel and broadly applicable approach to antiviral therapy. However, clinical application of RNAi faces several challenges, including the potential for viral escape (6, 7, 52), nonspecific RNAi effects (36, 55, 60), and limited delivery efficiency of siRNA duplexes or expression vectors. For all of these reasons, there is a significant need to optimize the design of antiviral RNAi therapy strategies.

In this study, we investigated several clinically relevant considerations and constraints in order to facilitate the design of anti-HIV RNAi-based therapies that delay or prevent the emergence of resistant viruses. This investigation was performed in silico using a novel type of MASS. Previous computational and theoretical formulations would not be well suited to address the scenarios simulated. The simulation more closely models the process of in vitro viral propagation, but the predictions can be extended to in vivo replication as well.

It has been suggested by a number of researchers that viral evolution around RNAi might be thwarted through the simultaneous application of a mixture of antiviral siRNAs targeting different viral and/or cellular sequences (4, 40). Our results demonstrate that increasing the number of RNAi targets, even while holding the total level of siRNA molecules constant,
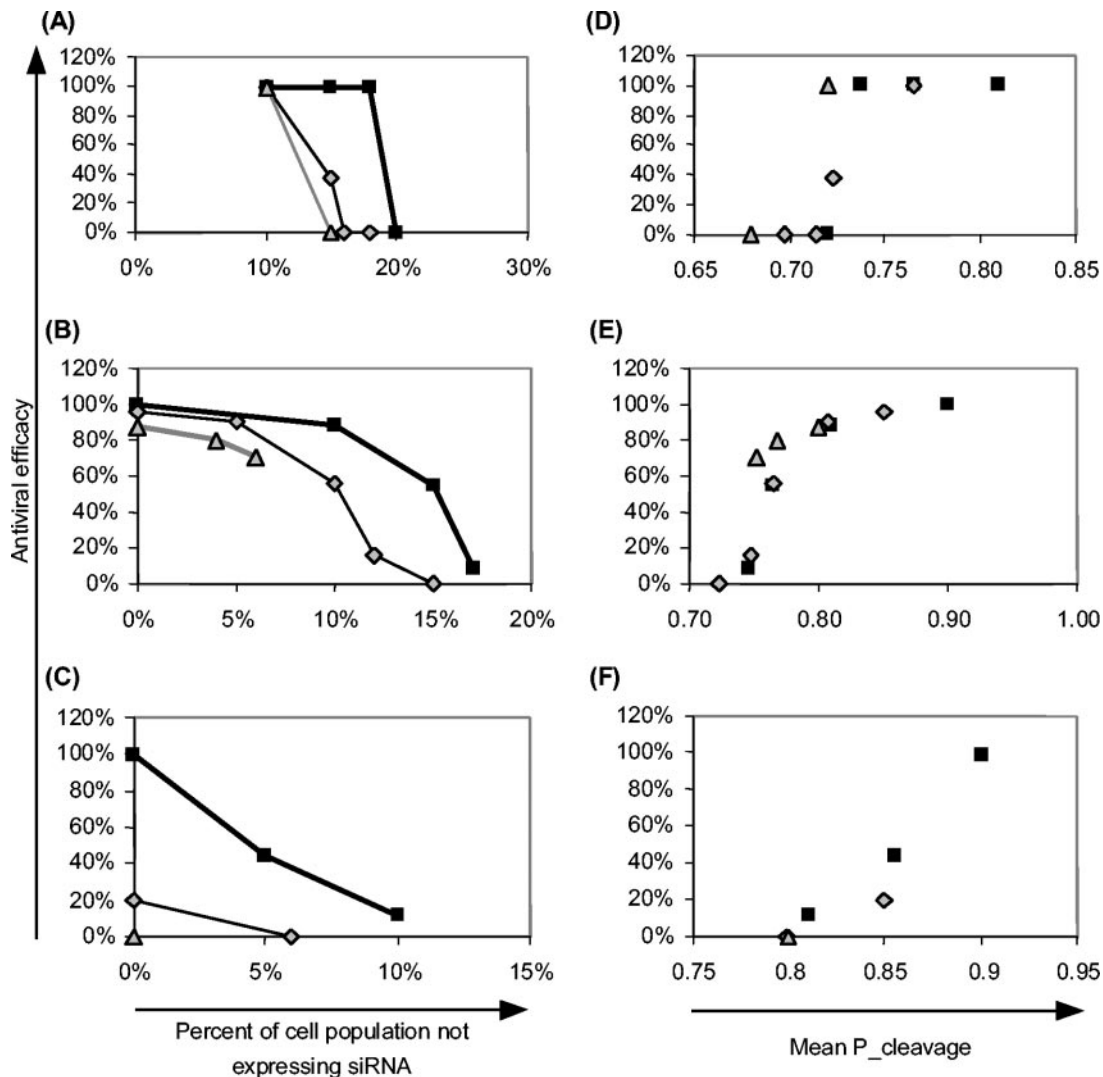
FIG. 5. Nonuniform RNAi: bipartite cell population and Tat/TAR functional viral genomes. (A to C) Similar to Fig. 2, Tat/TAR functional genomes replicated on a mixed population of cells expressing either uniform levels of siRNA (corresponding to P_cleavage) or no siRNA were analyzed. In each series, P_cleavage was held constant while the fraction of nonexpressing cells was increased. A highly conserved sequence (bases 18 to 38) (A), an intermediately conserved sequence (bases 31 to 51) (B), and a nonconserved sequence (bases 47 to 67) (C) were used. (D to F) Data points from panels A to C, respectively, plotted against the mean P_cleavage of the population, which was calculated by multiplying the value of P_cleavage in siRNA-expressing cells by the fraction of cells that expressed siRNA. Symbols for different values of mean P_cleavage are as follows: 0.9 (black squares); 0.85 (grey diamonds); 0.8 (grey triangles).

dramatically improves RNAi efficacy (Fig. 1A). However, this rapid improvement in efficacy occurred over a narrow range of target numbers. That is, as the overall siRNA expression level was increased, fewer targets were required to achieve the elevated efficacy. This result indicates that in practice, targeting a small set of viral sequences could be effective, as long as the siRNA delivery or expression is sufficiently strong. Additionally, we observed that stronger RNAi reduced the chance of viral escape, but when escape did occur, it occurred more rapidly (Fig. 1B). Most likely, stronger RNAi led to accelerated viral extinction, such that escape was only observed when it occurred early during the infection.

In clinical practice, RNAi will most likely be induced by either delivery of presynthesized siRNAs (complexed into a liposome or polyplex) or through the use of gene transfer

vectors carrying delivery siRNA expression cassettes for in vivo or ex vivo gene therapy. Regardless, target cell transduction efficiency will be limited to some extent, and our results suggest that this limitation could significantly compromise long-term antiviral efficacy. With a mixed pool of expressing and nonexpressing cells, we found that the cell population's overall mean level of RNAi was the single most important determinant of efficacy against neutral viral genomes (Fig. 2). The fact that all data collapse onto a single curve suggests the simple design principle that limited siRNA delivery can be at least partially compensated by strong RNAi induction in cells that are successfully transduced. However, the mean RNAi level was no longer a sole predictor of efficacy when Tat/TAR functionalized genomes were used. Specifically, given two scenarios with the same mean RNAi level (mean P_cleavage), antiviral effi-
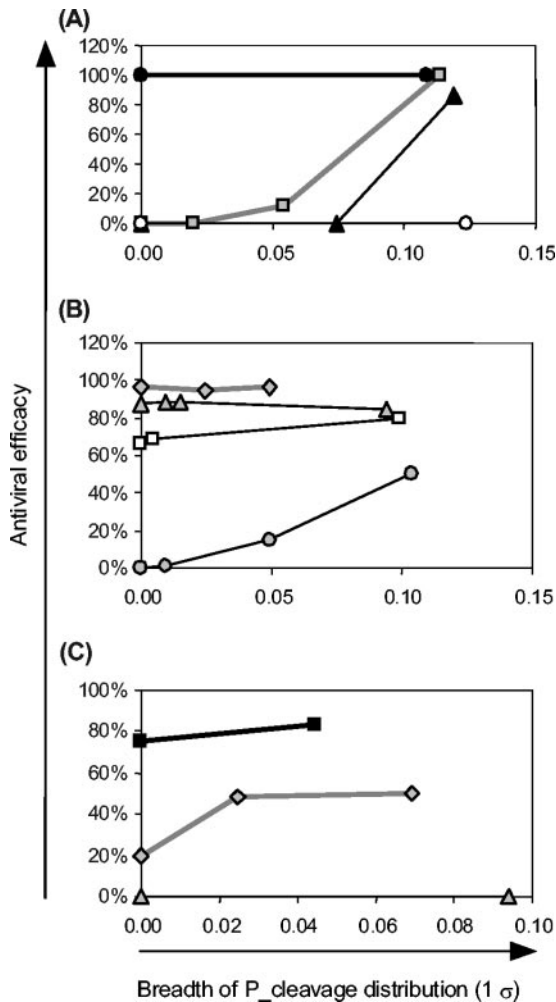
FIG. 6. Nonuniform RNAi: binomial distribution of siRNA expression and Tat/TAR functional viral genomes. Similar to Fig. 3, Tat/TAR functional genomes replicated on a cell population with binomially distributed siRNA expression. In each series, mean P_cleavage was held constant while the standard deviation of the distribution (σ) was increased. A highly conserved sequence (bases 18 to 38) (A), an intermediately conserved sequence (bases 31 to 51) (B), and a nonconserved sequence (bases 47 to 67) (C) were used. The right-most data point in each series represents the widest possible distribution for that particular value of mean P_cleavage. Symbols for different values of mean P_cleavage are as follows: 0.9 (black squares); 0.85 (grey diamonds); 0.8 (grey triangles); 0.79 (white squares); 0.78 (grey circles); 0.77 (black circles); 0.76 (grey squares); 0.75 (black triangles); 0.74 (white circles).

cacy was higher when more cells expressed siRNA, even at a lower level (Fig. 5). Consequently, it may be therapeutically important to ensure that a maximal fraction of the target cell population induces antiviral RNAi, even if this comes at the expense of lower levels of RNAi in those same cells.

We also analyzed the case where delivery was sufficient to induce some degree of antiviral RNAi in all cells, but the level of RNAi varied between cells, following a binomial distribution, another clinically relevant situation. When neutral viral genomes were propagated on these cells, we found that increasing the heterogeneity of siRNA delivery or expression led to small increases or decreases in efficacy, depending on the

overall level of siRNA present. However, the efficacy appeared to be rather insensitive to siRNA expression heterogeneity in general. In surprising contrast, antiviral efficacy against Tat/TAR functionalized genomes generally increased with the heterogeneity of RNAi expression, and in some cases the increase was dramatic (Fig. 6).

These differences between neutral and functional genome models might be explained by the existence of a threshold level of Tat feedback. If a given RNAi strategy can effectively reduce viral genetic expression below this threshold, Tat would be unable to effectively amplify HIV mRNA transcription and thereby Tat's own expression. If a sufficient fraction of the cell population were forced into this drastically reduced level of viral gene expression, the HIV population would not be able to adequately sample viral sequence space to produce and establish an escape mutant population. This interpretation could explain both the increase in efficacy with siRNA expression heterogeneity that was observed at low levels of RNAi as well as the modest decrease in efficacy observed for higher levels of RNAi; in other words, if a cell population existed mainly on one site of this threshold level of Tat feedback, the primary effect of broadening the RNAi distribution would be to push some cells onto the other side of the threshold. In parallel, viral genomes that coinfect a single cell likely cooperate by contributing Tat to activate each other's Tat/TAR feedback loops. If RNAi-mediated inhibition is sufficiently strong to drive viral expression down to the point where efficient viral replication depends on cooperative production of Tat, slight increases in viral inhibition could significantly increase the antiviral efficacy by reducing the occurrence of viral coinfection. This argument could explain why antiviral efficacy was more sensitive to increases in either the overall RNAi level or in the fraction of cells that produce RNAi when targeting functional viral genomes instead of neutral genomes. These results also demonstrate why it is important to incorporate key aspects of the viral life cycle into a model.

At any rate, it appears that therapeutic efficacy is more sensitive to the total fraction of cells that express some RNAi than to the heterogeneity of expression levels in a fully transduced cell population. Therefore, antiviral therapies should be designed to minimize the fraction of target cells that do not exhibit antiviral RNAi, since they could provide an avenue for viral replication and escape that can subsequently overwhelm and compromise the entire cell population. In clinical practice, it might be difficult to ensure that siRNA are delivered to or expressed in a majority of the target cell population. In some cases, though, the effective transduction efficiency can be increased if the transduced cells have a selective growth advantage. For example, in clinical trials using murine retrovirus-mediated gene therapy to treat congenital X-linked severe combined immunodeficiency, the cells transduced with a genetically corrected cytokine receptor were fortuitously positively selected and thus were able to repopulate the patients (23). Recently, other researches have developed a method for reversibly conveying a growth advantage to transduced lymphocytes, such that repopulation can be safely enhanced in situations where native cells are not efficiently cleared naturally (70). Considering the apparent significance of transduction efficiency in attaining antiviral efficacy, it might be worth pursuing more aggressive techniques for enhancing the expan-

sion of transduced cells, which might otherwise be deemed excessively risky.

In this study, we found that antiviral efficacy was highly sensitive to the choice of RNAi target sequence within TAR, such that a shift of several bases was sufficient to eliminate viral inhibition, especially under conditions of lower overall siRNA expression. It should be noted that not all RNA sequences are equally susceptible to RNAi-mediated degradation, and we have not yet incorporated such limitations (54). However, our results do suggest that when testing different siRNAs against conserved viral sequences (such as TAR), it is probably worthwhile to exhaustively test as many sequences as possible.

A number of assumptions were built into the model, and future development may include expanding the model to more accurately incorporate several aspects of the viral life cycle. For example, we assumed that the virus replicates in discrete generations, similar to passage in cell culture, but the model can readily be extended to account for intergenerational mixing of mutants in the viral swarm. This study incorporated the viral diversity-generating mechanisms of point mutation and recombination, but viral genomic deletions and insertions were not considered. Since the RNAi was directed against a highly conserved *cis*-acting sequence (i.e., TAR), major rearrangements in this region would probably lead to a dead-end viral genome. If truly nonessential viral sequences were targeted by RNAi, it is likely that the virus could simply remove these sequences through deletion (7). Consequently, by prohibiting deletions, the neutral case served as an extreme lower limit on the degree to which a target site is evolutionarily conserved. In any case, future work could incorporate deletions and insertions by allowing for a low rate of template switching between regions of imperfect homology during reverse transcription; such a modification would necessitate the postulation of a recombination probability function that incorporates the degree of homology between RNA template regions. In addition, this work considered only multiple targets within the HIV genome. In reality, combinatorial gene therapies targeting various aspects of the HIV life cycle, including both viral and cellular targets, may be highly effective (4, 40, 42). These viral and host functions, such as the use of cellular receptors CCR5 and CXCR4, as well as the cyclin T1 and Cdk9 cofactors for Tat transactivation, can readily be incorporated into the simulation to further explore which combinations of therapies may prove most effective, as well as provide insights into how best to enhance the overall therapy through the optimization of its individual antiviral elements. Finally, RNAi therapies would most likely be used for patients already infected with the virus, and the virus would therefore have already replicated and mutated its sequence, similar to the challenges faced by HAART (52). Future work may analyze how the preexistence of viral sequence diversity can impact therapeutic success.

We have developed a robust method for analyzing retroviral evolution in a manner that can incorporate a significant level of molecular detail, and we have shown that this approach can yield insights into the highly complex process by which replicating entities evolve around selective pressure. In addition to aiding the development of more effective RNAi therapies, this system could be used to study more fundamental aspects of viral evolution, such as the debated advantages of a mutator strain, the role of recombination in evolution, and the balance between advantageous, neutral, and deleterious mutations (12, 13, 27, 32). An experimental system to test the predictions of this work could establish the biomedical relevance and implications of this work, as well as serve as a model system for further analyzing mechanisms of viral evolution. This stochastic simulation environment thus provides a useful starting point for developing an enhanced molecular-level understanding of many medically and biologically important aspects of retroviral evolution.

REFERENCES

1. **Adachi, A., H. E. Gendelman, S. Koenig, T. Folks, R. Willey, A. Rabson, and M. A. Martin.** 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. J. Virol. **59:**284–291.
2. **Anderson, J., A. Banerjea, and R. Akkina.** 2003. Bispecific short hairpin siRNA constructs targeted to CD4, CXCR4, and CCR5 confer HIV-1 resistance. Oligonucleotides **13:**303–312.
3. **Banerjea, A., M. J. Li, G. Bauer, L. Remling, N. S. Lee, J. Rossi, and R. Akkina.** 2003. Inhibition of HIV-1 by lentiviral vector-transduced siRNAs in T lymphocytes differentiated in SCID-hu mice and CD34⁺ progenitor cell-derived macrophages. Mol. Ther. **8:**62–71.
4. **Berkhout, B.** 2004. RNA interference as an antiviral approach: targeting HIV-1. Curr. Opin. Mol. Ther. **6:**141–145.
5. **Berkhout, B., and K. T. Jeang.** 1989. *trans* activation of human immunodeficiency virus type 1 is sequence specific for both the single-stranded bulge and loop of the *trans*-acting-responsive hairpin: a quantitative analysis. J. Virol. **63:**5501–5504.
6. **Boden, D., O. Pusch, F. Lee, L. Tucker, and B. Ramratnam.** 2003. Human immunodeficiency virus type 1 escape from RNA interference. J. Virol. **77:**11531–11535.
7. **Das, A. T., T. R. Brummelkamp, E. M. Westerhout, M. Vink, M. Madiredjo, R. Bernards, and B. Berkhout.** 2004. Human immunodeficiency virus type 1 escapes from RNA interference-mediated inhibition. J. Virol. **78:**2601–2605.
8. **Davis, B. M., L. Humeau, and B. Dropulic.** 2004. In vivo selection for human and murine hematopoietic cells transduced with a therapeutic MGMT lentiviral vector that inhibits HIV replication. Mol. Ther. **9:**160–172.
9. **Denli, A. M., and G. J. Hannon.** 2003. RNAi: an ever-growing puzzle. Trends Biochem. Sci. **28:**196–201.
10. **Di Mascio, M., R. M. Ribeiro, M. Markowitz, D. D. Ho, and A. S. Perelson.** 2004. Modeling the long-term control of viremia in HIV-1 infected patients treated with antiretroviral therapy. Math. Biosci. **188:**47–62.
11. **Dingwall, C., I. Ernberg, M. J. Gait, S. M. Green, S. Heaphy, J. Karn, A. D. Lowe, M. Singh, and M. A. Skinner.** 1990. HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. EMBO J. **9:**4145–4153.
12. **Domingo, E.** 2001. Quasispecies and RNA virus evolution: principles and consequences. Landes Bioscience, Georgetown, Tex.
13. **Domingo, E., R. G. Webster, and J. J. Holland.** 1999. Origin and evolution of viruses. Academic Press, San Diego, Calif.
14. **Feng, S., and E. C. Holland.** 1988. HIV-1 tat trans-activation requires the loop sequence within tar. Nature **334:**165–167.
15. **Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello.** 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature **391:**806–811.
16. **Fleming, P. L., R. H. Byers, P. A. Sweeney, D. Daniels, J. M. Karon, and R. S. Janssen.** 2000. HIV prevalence in the United States, 2000. Abstr. 9th Conf. Retroviruses Opportunistic Infect. 2000, abstr. 11.
17. **Garber, D. A., G. Silvestri, and M. B. Feinberg.** 2004. Prospects for an AIDS vaccine: three big questions, no easy answers. Lancet Infect. Dis. **4:**397–413.
18. **Garcia, J. A., D. Harrich, E. Soultanakis, F. Wu, R. Mitsuyasu, and R. B. Gaynor.** 1989. Human immunodeficiency virus type 1 LTR TATA and TAR region sequences required for transcriptional regulation. EMBO J. **8:**765–778.
19. **Ge, Q., M. T. McManus, T. Nguyen, C. H. Shen, P. A. Sharp, H. N. Eisen, and J. Chen.** 2003. RNA interference of influenza virus production by directly targeting mRNA for degradation and indirectly inhibiting all viral RNA transcription. Proc. Natl. Acad. Sci. USA **100:**2718–2723.
20. **Gitlin, L., S. Karelsky, and R. Andino.** 2002. Short interfering RNA confers intracellular antiviral immunity in human cells. Nature **418:**430–434.
21. **Gruber, A., I. Chen, K. L. Kuhen, J. C. Wheat, P. Law, and F. Wong-Staal.**

2003. Generation of dendritic cells from lentiviral vector-transduced CD34⁺ cells from HIV⁺ donors. J. Med. Virol. **70:**183–186.

22. **Gupta, S., R. A. Schoer, J. E. Egan, G. J. Hannon, and V. Mittal.** 2004. Inducible, reversible, and stable RNA interference in mammalian cells. Proc. Natl. Acad. Sci. USA **101:**1927–1932.

23. **Hacein-Bey-Abina, S., A. Fischer, and M. Cavazzana-Calvo.** 2002. Gene therapy of X-linked severe combined immunodeficiency. Int. J. Hematol. **76:**295–298.

24. **Hamada, M., T. Ohtsuka, R. Kawaida, M. Koizumi, K. Morita, H. Furukawa, T. Imanishi, M. Miyagishi, and K. Taira.** 2002. Effects on RNA interference in gene expression (RNAi) in cultured mammalian cells of mismatches and the introduction of chemical modifications at the 3′-ends of siRNAs. Antisense Nucleic Acid Drug Dev. **12:**301–309.

25. **Hauber, J., and B. R. Cullen.** 1988. Mutational analysis of the *trans*-activation-responsive region of the human immunodeficiency virus type I long terminal repeat. J. Virol. **62:**673–679.

26. **Holen, T., M. Amarzguioui, M. T. Wiiger, E. Babaie, and H. Prydz.** 2002. Positional effects of short interfering RNAs targeting the human coagulation trigger tissue factor. Nucleic Acids Res. **30:**1757–1766.

27. **Holmes, E. C., and A. Moya.** 2002. Is the quasispecies concept relevant to RNA viruses? J. Virol. **76:**460–465.

28. **Humeau, L. M., G. K. Binder, X. Lu, V. Slepushkin, R. Merling, P. Echeagaray, M. Pereira, T. Slepushkina, S. Barnett, L. K. Dropulic, R. Carroll, B. L. Levine, C. H. June, and B. Dropulic.** 2004. Efficient lentiviral vector-mediated control of HIV-1 replication in CD4 lymphocytes from diverse HIV⁺ infected patients grouped according to CD4 count and viral load. Mol. Ther. **9:**902–913.

29. **Huthoff, H., and B. Berkhout.** 2001. Mutations in the TAR hairpin affect the equilibrium between alternative conformations of the HIV-1 leader RNA. Nucleic Acids Res. **29:**2594–2600.

30. **Jacque, J. M., K. Triques, and M. Stevenson.** 2002. Modulation of HIV-1 replication by RNA interference. Nature **418:**435–438.

31. **Jakobovits, A., D. H. Smith, E. B. Jakobovits, and D. J. Capon.** 1988. A discrete element 3′ of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by an HIV trans activator. Mol. Cell. Biol. **8:**2555–2561.

32. **Jenkins, G. M., M. Worobey, C. H. Woelk, and E. C. Holmes.** 2001. Evidence for the non-quasispecies evolution of RNA viruses [corrected]. Mol. Biol. Evol. **18:**987–994.

33. **Jetzt, A. E., H. Yu, G. J. Klarmann, Y. Ron, B. D. Preston, and J. P. Dougherty.** 2000. High rate of recombination throughout the human immunodeficiency virus type 1 genome. J. Virol. **74:**1234–1240.

34. **Jordan, A., P. Defechereux, and E. Verdin.** 2001. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. EMBO J. **20:**1726–1738.

35. **Kijak, G. H., V. Simon, P. Balfe, J. Vanderhoeven, S. E. Pampuro, C. Zala, C. Ochoa, P. Cahn, M. Markowitz, and H. Salomon.** 2002. Origin of human immunodeficiency virus type 1 quasispecies emerging after antiretroviral treatment interruption in patients with therapeutic failure. J. Virol. **76:**7000–7009.

36. **Kim, D. H., M. Longo, Y. Han, P. Lundberg, E. Cantin, and J. J. Rossi.** 2004. Interferon induction by siRNAs and ssRNAs synthesized by phage polymerase. Nat. Biotechnol. **22:**321–325.

37. **Kimura, M.** 1994. Population genetics, molecular evolution, and the neutral theory; selected papers. University of Chicago Press, Chicago.

38. **Klaver, B., and B. Berkhout.** 1994. Premature strand transfer by the HIV-1 reverse transcriptase during strong-stop DNA synthesis. Nucleic Acids Res. **22:**137–144.

39. **Layden, T. J., J. E. Layden, R. M. Ribeiro, and A. S. Perelson.** 2003. Mathematical modeling of viral kinetics: a tool to understand and optimize therapy. Clin. Liver Dis. **7:**163–178.

40. **Lee, N. S., and J. J. Rossi.** 2004. Control of HIV-1 replication by RNA interference. Virus Res. **102:**53–58.

41. **Leirdal, M., and M. Sioud.** 2002. Gene silencing in mammalian cells by preformed small RNA duplexes. Biochem. Biophys. Res. Commun. **295:**744–748.

42. **Li, M. J., G. Bauer, A. Michienzi, J. K. Yee, N. S. Lee, J. Kim, S. Li, D. Castanotto, J. Zaia, and J. J. Rossi.** 2003. Inhibition of HIV-1 infection by lentiviral vectors expressing Pol III-promoted anti-HIV RNAs. Mol. Ther. **8:**196–206.

43. **Lu, X., Q. Yu, G. K. Binder, Z. Chen, T. Slepushkina, J. Rossi, and B. Dropulic.** 2004. Antisense-mediated inhibition of human immunodeficiency virus (HIV) replication by use of an HIV type 1-based vector results in severely attenuated mutants incapable of developing resistance. J. Virol. **78:**7079–7088.

44. **Mansky, L. M., and H. M. Temin.** 1995. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. J. Virol. **69:**5087–5094.

45. **Michienzi, A., D. Castanotto, N. Lee, S. Li, J. A. Zaia, and J. J. Rossi.** 2003. RNA-mediated inhibition of HIV in a gene therapy setting. Ann. N. Y. Acad. Sci. **1002:**63–71.

46. **Miller, V. M., H. Xia, G. L. Marrs, C. M. Gouvion, G. Lee, B. L. Davidson,**

and **H. L. Paulson.** 2003. Allele-specific silencing of dominant disease genes. Proc. Natl. Acad. Sci. USA **100:**7195–7200.

47. **Muller, V., and S. Bonhoeffer.** 2003. Mathematical approaches in the study of viral kinetics and drug resistance in HIV-1 infection. Curr. Drug Targets Infect. Disord. **3:**329–344.

48. **Ou, S. H., F. Wu, D. Harrich, L. F. Garcia-Martinez, and R. B. Gaynor.** 1995. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. J. Virol. **69:**3584–3596.

49. **Paddison, P. J., A. A. Caudy, and G. J. Hannon.** 2002. Stable suppression of gene expression by RNAi in mammalian cells. Proc. Natl. Acad. Sci. USA **99:**1443–1448.

50. **Parrish, S., J. Fleenor, S. Xu, C. Mello, and A. Fire.** 2000. Functional anatomy of a dsRNA trigger: differential requirement for the two trigger strands in RNA interference. Mol. Cell **6:**1077–1087.

51. **Radhakrishnan, S. K., T. J. Layden, and A. L. Gartel.** 2004. RNA interference as a new strategy against viral hepatitis. Virology **323:**173–181.

52. **Rambaut, A., D. Posada, K. A. Crandall, and E. C. Holmes.** 2004. The causes and consequences of HIV evolution. Nat. Rev. Genet. **5:**52–61.

53. **Reddy, B., and J. Yin.** 1999. Quantitative intracellular kinetics of HIV type 1. AIDS Res. Hum. Retrovir. **15:**273–283.

54. **Reynolds, A., D. Leake, Q. Boese, S. Scaringe, W. S. Marshall, and A. Khvorova.** 2004. Rational siRNA design for RNA interference. Nat. Biotechnol. **22:**326–330.

55. **Scacheri, P. C., O. Rozenblatt-Rosen, N. J. Caplen, T. G. Wolfsberg, L. Umayam, J. C. Lee, C. M. Hughes, K. S. Shanmugam, A. Bhattacharjee, M. Meyerson, and F. S. Collins.** 2004. Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells. Proc. Natl. Acad. Sci. USA **101:**1892–1897.

56. **Scherr, M., K. Battmer, A. Ganser, and M. Eder.** 2003. Modulation of gene expression by lentiviral-mediated delivery of small interfering RNA. Cell Cycle **2:**251–257.

57. **Schroder, A. R., P. Shinn, H. Chen, C. Berry, J. R. Ecker, and F. Bushman.** 2002. HIV-1 integration in the human genome favors active genes and local hotspots. Cell **110:**521–529.

58. **Selby, M. J., E. S. Bain, P. A. Luciw, and B. M. Peterlin.** 1989. Structure, sequence, and position of the stem-loop in tar determine transcriptional elongation by tat through the HIV-1 long terminal repeat. Genes Dev. **3:**547–558.

59. **Simm, M., W. Chao, O. Pekarskaya, P. Sova, P. Gupta, R. Balachandran, and D. J. Volsky.** 1996. Genetic variability and function of the long terminal repeat from syncytium-inducing and non-syncytium-inducing human immunodeficiency virus type 1. AIDS Res. Hum. Retrovir. **12:**801–809.

60. **Sledz, C. A., M. Holko, M. J. de Veer, R. H. Silverman, and B. R. Williams.** 2003. Activation of the interferon system by short-interfering RNAs. Nat. Cell Biol. **5:**834–839.

61. **ter Hofstede, H. J., D. M. Burger, and P. P. Koopmans.** 2003. Antiretroviral therapy in HIV patients: aspects of metabolic complications and mitochondrial toxicity. Neth. J. Med. **61:**393–403.

62. **Verhoef, K., M. Koper, and B. Berkhout.** 1997. Determination of the minimal amount of Tat activity required for human immunodeficiency virus type 1 replication. Virology **237:**228–236.

63. **Vigouroux, C., S. Gharakhanian, Y. Salhi, T. H. Nguyen, N. Adda, W. Rozenbaum, and J. Capeau.** 1999. Adverse metabolic disorders during highly active antiretroviral treatments (HAART) of HIV disease. Diabetes Metab. **25:**383–392.

64. **Weinberger, L. S., D. V. Schaffer, and A. P. Arkin.** 2003. Theoretical design of a gene therapy to prevent AIDS but not human immunodeficiency virus type 1 infection. J. Virol. **77:**10028–10036.

65. **World Health Organization.** 2004. AIDS epidemic update: December 2003. [Online.] http://www.who.int/hiv/pub/epidemiology/epi2003/en/.

66. **Wiethoff, C. M., and C. R. Middaugh.** 2003. Barriers to nonviral gene delivery. J. Pharm. Sci. **92:**203–217.

67. **Wiznerowicz, M., and D. Trono.** 2003. Conditional suppression of cellular genes: lentivirus vector-mediated drug-inducible RNA interference. J. Virol. **77:**8957–8961.

68. **Wodarz, D., and M. A. Nowak.** 2002. Mathematical models of HIV pathogenesis and treatment. Bioessays **24:**1178–1187.

69. **Wu, F., J. Garcia, D. Sigman, and R. Gaynor.** 1991. tat regulates binding of the human immunodeficiency virus trans-activating region RNA loop-binding protein TRP-185. Genes Dev. **5:**2128–2140.

70. **Yam, P. Y., J. Wu, M. C. Jensen, J. A. Zaia, and J. K. Yee.** 2004. Abstr. Am. Soc. Gene Ther. 7th Annu. Meet., abstr. 369, p. 71.

71. **Yoshinari, K., M. Miyagishi, and K. Taira.** 2004. Effects on RNAi of the tight structure, sequence and position of the targeted region. Nucleic Acids Res. **32:**691–699.

72. **Zhu, Y., T. Pe'ery, J. Peng, Y. Ramanathan, N. Marshall, T. Marshall, B. Amendt, M. B. Mathews, and D. H. Price.** 1997. Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. Genes Dev. **11:**2622–2632.