

FIDELITY OF HIV-1 REVERSE TRANSCRIPTASE: THE “TRUTH” IS OUT THERE...*IN VIVO*!

Michael E. Abram<sup>1</sup>, Andrea L. Ferris<sup>1</sup>, and Stephen H. Hughes<sup>1</sup>

<sup>1</sup>HIV Drug Resistance Program, National Cancer Institute, Frederick, MD, USA

The relatively rapid genetic variation of HIV-1 observed in patients has been ascribed to the low fidelity/high mutation rate of reverse transcriptase (RT). While RT may be error prone in its conversion of the retroviral genome into dsDNA, the contribution of the host RNA polymerase II to the overall mutation rate *in vivo* has not been determined. It has been suggested, based largely on *in vitro* data, that errors made by HIV-1 RT often involve a misalignment/slippage mechanism, which causes a large number of frameshift mutations. Here we describe an efficient system that improves upon existing technology, to analyze the sequence of *in vivo*-derived products of HIV replication in greater detail. Using an HIV-1 vector that is both IN- and Env- we recover, after a single round of replication,  $>4 \times 10^4$  unintegrated HIV DNAs per infection. The presence of the *lacZ $\alpha$*  reporter gene makes it easy to identify viral DNAs that contain mutations in the reporter. All of the mutants we recover are analyzed by DNA sequencing. This makes it possible to determine the overall error rate, identify mutational hotspots, and determine the frequency of different types of mutations. In contrast to previous studies, mutation frequency was based on the number of actual mutations divided by the total number of colonies. We found the forward *in vivo* mutation rate of HIV-1 to be  $1.2 \times 10^{-5}$  mutations/bp/cycle; 3-fold lower than previously reported, the rate we measured is close to the rate reported for other retroviruses. While the mutation rate was not affected by the orientation of *lacZ $\alpha$*  in the vector, the number and types of mutations at each nucleotide position differed considerably. Base substitutions, not frameshifts, were the most frequent mutations observed, with G-to-A transitions predominating 2:1 over C-to-T transitions. A significant subset of the mutant sequences contained multiple of A-to-G transitions, including doublets and spaced singlets. Multiple mutations were analyzed separately in the determination of the overall error rate. Mutations in the dNTP-binding pocket that had overall mutation rates similar to wild-type showed pronounced differences in the mutational hotspots in *lacZ $\alpha$* . These data suggest that this system can be used to determine types of errors made during HIV replication *in vivo*, which will make it possible to determine how RT mutations affect fidelity.