Journal of NeuroVirology, 7: 235–249, 2001
 ISSN 1355–0284/01 $12.00+.00

HIV-1 LTR C/EBP binding site sequence configurations preferentially encountered in brain lead to enhanced C/EBP factor binding and increased LTR-specific activity

Heather L Ross,[1] Suzanne Gartner,[2] Justin C McArthur,[2] John R Corboy,[3] John J McAllister,[1] Scott Millhouse,[1] and Brian Wigdahl[1]

[1]*Department of Microbiology and Immunology, The Pennsylvania State University, College of Medicine, Hershey, Pennsylvania, USA;* [2]*Department of Neurology, The Johns Hopkins School of Medicine, Baltimore, Maryland, USA; and* [3]*Department of Neurology, The University of Colorado, Health Sciences Center, Denver, Colorado, USA*

Recent studies have shown that two CAAT/enhancer binding protein (C/EBP) sites are critically important for efficient human immunodeficiency virus (HIV) type 1 (HIV-1) replication within cells of the monocyte/macrophage lineage, a primary cell type infected by HIV-1 and a potentially important vehicle for transport of virus to the central nervous system (CNS). Given the relevance of HIV-1 LTR sequence variation with respect to HIV-1 replication within monocyte populations and the important role that monocyte tropism likely plays in HIV-1 infection of the brain, C/EBP site sequence variation was examined within peripheral blood- and brain-derived LTR populations. Brain-derived LTRs commonly possessed a C/EBP site I configuration (6G, comprised of a thymidine to guanosine substitution with respect to the clade B consensus sequence at position 6 of C/EBP site I) that leads to enhanced binding of C/EBP proteins over that observed with the HIV-1 clade B consensus sequence at this site. In contrast, the 6G C/EBP site I configuration appeared infrequently within sequenced peripheral blood-derived LTRs. In addition, C/EBP site II was even more highly conserved in brain-derived HIV-1 LTR populations than site I. This was not the case with peripheral blood-derived LTR C/EBP site II sequences. The high degree of C/EBP site II conservation in brain-derived LTRs was likely important in LTR regulation since the clade B consensus sequence conserved at C/EBP site II recruited high amounts of C/EBP family members. Transient transfection analyses indicated that conservation of the strong C/EBP site II in brain-derived LTRs was likely due to important interactions with Tat. Overall, brain-derived HIV-1 LTRs preferentially contained two highly reactive C/EBP binding sites, which may suggest that these sites play important roles in LTR-directed transcription during invasion and maintenance of HIV-1 in the central nervous system. *Journal of NeuroVirology* (2001) **7,** 235–249.

Introduction

The human immunodeficiency virus (HIV) type 1 (HIV-1)-associated acquired immune deficiency syndrome (AIDS) dementia complex (HADC) comprises a series of cognitive, motor, and behavioral dysfunctions that affect an estimated 20% of AIDS patients (McArthur *et al*, 1993; Lipton and Gendelman, 1995). More sensitive neuropsychological testing has indicated that the number of patients with this progressive neurodegenerative disorder is likely greater (Ho *et al*, 1989; Geleziunas *et al*, 1992), and central nervous system (CNS) pathologic investigations report abnormalities at time of autopsy in about 90% of cases (Ho *et al*, 1989; Geleziunas *et al*, 1992). This has suggested that HIV-1 infection of the brain is a common occurrence and often leads to

Address correspondence to Brian Wigdahl, Department of Microbiology and Immunology (H107), The Pennsylvania State University College of Medicine, 500 University Drive, P. O. Box 850, Hershey, PA 17033, USA. Fax: (717) 531-5580. E-mail: bwigdahl@psu.edu

Received 20 June 2000; revised 3 November 2000; accepted 19 December 2000.

progressive neurologic dysfunction. In more recent years, the number of cases of HADC has declined with the advent of highly active antiretroviral therapy (HAART) (Perelson *et al*, 1997).

Investigations into HIV-1 gene regulation within cells of the monocyte/macrophage lineage is important, because cells of this lineage may produce a bulk of potentially neurotoxic proteins and infectious virus in the CNS during the generation of HIV-specific neurologic dysfunction. HIV-1 replication can lead to cytokine and cell signaling dysregulation, which impacts neurologic function. In addition, proviral transcription can lead to the production of neurotoxic proteins such as Tat, Nef, and gp120. Therefore, identifying the cellular and viral factors that regulate HIV-1 gene expression within cells of the monocyte/macrophage lineage is clearly an important component of understanding the complex disease process associated with HIV-1 CNS infection.

In general, proteins from the C/EBP transcription factor family appear to be intimately involved in the regulation of gene expression in myelomonocytic cells. The promoter elements of many monocyte-expressed genes contain C/EBP binding sites and cell signaling events within cells of the monocyte/macrophage lineage can significantly up-regulate C/EBP expression. Therefore, it is not surprising that this family of transcription factors appears to play a very prominent role in HIV-1 replication within cells of the monocyte/macrophage lineage. Previous studies have shown that C/EBP β could *trans*-activate the HIV-1 long terminal repeat (LTR) in transient transfection analyses and that the LTR contained three binding sites for this protein (Tesmer *et al*, 1993). C/EBP factors have been demonstrated to *trans*-activate the HIV-1 LTR in the U-937 promonocytic cell line (Henderson *et al*, 1995). Furthermore, site-directed mutagenesis studies have established that LTR-directed transcription in these cells required at least one functional C/EBP site (located at either −107 to −118 or −167 to −175 relative to the transcriptional start site). Additional studies indicated that both C/EBP binding sites were required for replication of an infectious HIV-1 molecular clone in U-937 cells, as well as in primary macrophage populations. However, they were dispensable for replication of the HIV-1 infectious molecular clone in selected T cell lines and primary T cell populations (Henderson *et al*, 1996; Henderson and Calame, 1997).

The viral genome evolves as a consequence of viral replication during the course of disease progression (Pang *et al*, 1992), and quickly leads to phenotypic alterations in progeny virus. These sequence changes can occur throughout the viral genome, including the different transcription factor binding sites of the LTR, leading to alterations in binding site reactivity and variations in viral LTR activity. One study that demonstrates the potential impact of LTR sequence variation on HIV-1-specific gene expression was reported by Corboy and coworkers (1992). In this study, transgenic mice were constructed carrying LTRs driving beta galactosidase (β-gal) expression. Two of the LTRs were derived from HIV-1 strains obtained from the frontal lobe (JR-FL) and cell-free cerebrospinal fluid (JR-CSF) of patient JR. The third LTR utilized in this study was obtained from the T-tropic HIV-1 IIIB/LAV strain. Only animals containing CNS-derived LTRs directed β-gal expression within the CNS. Expression in these animals was particularly evident in neurons. In addition, the two CNS-derived LTRs, which differed from one another at only 16 bp positions in the U3 region, differed slightly in the neuroanatomical locations of β-gal expression. These observations suggested that the LTR may play a role in viral gene expression in specific cell types, and consequently, yield a selective advantage for replication of certain HIV-1 quasispecies in the brain and likely within specific neuroanatomical brain locations.

We propose that HIV-1 LTR sequence variation plays a critical role in determining viral replication efficiency and may consequently impact overall viral replication in a cell type-specific manner. Given the critical role C/EBP sites play in HIV-1 replication within cells of the monocyte/macrophage lineage and the prevalence of HIV-1-infected monocytic cells within the CNS, the impact of LTR sequence variation in C/EBP sites I and II and its impact on DNA-protein complex formation was examined. Sequence variation in HIV-1 LTR C/EBP sites I and II within PCR-amplified, brain-derived viral LTRs was compared to similarly obtained LTR sequences derived from peripheral blood. These comparisons have indicated that brain-derived LTRs commonly possess two strong C/EBP binding sites, while very often peripheral blood-derived LTRs only contain one highly reactive site. In particular, a commonly encountered sequence variant of the clade B consensus C/EBP site I (6G, containing a thymidine to guanosine substitution at position 6) led to an increased reactivity for C/EBP factors. In addition, the clade B consensus C/EBP site II was extremely well conserved in the brain-derived LTRs as compared to the peripheral blood-derived LTRs. The conserved clade B C/EBP site II exhibits a very high reactivity for C/EBP transcription factors and appears to be important for Tat *trans*-activation in the context of an LTR obtained from the brain-derived YU-2 strain of HIV-1, but not other T-tropic or dual-tropic viruses. These results indicate one factor that may influence HIV-1 replication efficiency in specific brain cell populations is the presence of a relatively high affinity C/EBP binding site II and a specific C/EBP site I configuration. This may not be surprising because IL-6 is an important activator of C/EBP transcription (Akira *et al*, 1990) and elevated levels of IL-6 have been reported in the CSF of HIV-1-infected individuals (Gallo *et al*, 1989; Perrella *et al*, 1992).

Results

Brain-derived HIV-1 LTRs exhibit unique patterns of sequence variation at C/EBP sites I and II when compared to the same sites in peripheral blood-derived LTRs

Due to the limited number of brain-matched peripheral blood and tissue specimens available for the sequence analysis and the absence of complete clinical information obtained from many individuals within this study, the major objective of this study was to identify sequence changes in the C/EBP binding sites that were more conserved in the brain-derived LTRs compared to peripheral blood-derived sequences previously reported. Specifically, sequence alterations that would be expected to alter C/EBP factor reactivity towards either or both C/EBP site I or II and might consequently affect C/EBP-dependent transcription were the primary focal point for the DNA-protein biochemistry and functional analyses presented.

The PCR-amplified proviral sequences derived from brain represented a total of 164 LTR sequences derived from 18 individuals. The number of brain-derived LTRs examined per patient ranged from 3 to 21 and are summarized in Figure 3. Similar analyses were conducted using HIV-1 sequences collected by PCR amplification of provirus within uncultured peripheral blood mononuclear cells (PBMCs). The PBMC-derived LTR sequences were pooled from three large published studies, each of which contained multiple LTRs derived from four or more individuals that resided within North America. This analysis represented a total of 328 sequences derived from 40 individuals (Michael *et al*, 1994; Estable *et al*, 1996; Kirchhoff *et al*, 1997). The number of PBMC-derived LTRs examined per patient ranged from 3 to 42 and are summarized within each reference. Unfortunately, the status of patient brain disease was not reported in any of these three studies, so we were unable to establish any correlations between the two groups of LTRs (brain-derived versus PBMC-derived) and disease progression.

In our analysis, sequence changes that related to shared properties between patients, such as cellular and tissue targets infected by HIV-1, rather than patient-specific phenomena, such as secondary infections and drug treatment, are likely to be observed more frequently on an interpatient basis. Based on this reasoning, equal weight was placed on each individual, so the sequence analysis was conducted by first examining the LTRs of each individual separately to calculate the frequency of a nonconsensus change at each nucleotide position as a percent of total LTRs examined from the individual. The mean percentages of all individuals were then determined to obtain a mean frequency of nonconsensus change subsequently designated as the mean nonconsensus frequency (NCF) at each nucleotide position of C/EBP sites I and II, as well as the activating transcription factor/cyclic AMP response element binding protein (ATF/CREB) and NF-κB binding sites. These values were then plotted as a bar graph above the clade B consensus sequence (Figure 1A–D). Nucleotide changes most often observed were apparent nucleotide substitutions, although insertions and deletions did occur at certain positions and these too were included as nucleotide changes in this analysis.

To examine sequence variation within C/EBP sites I and II, the sequences were aligned and compared to the consensus sequence of HIV-1 clade B within this region. For comparative purposes, ATF/CREB and nuclear factor kappa B (NF-κB) binding sites, which flank C/EBP site I (relative positions are diagrammed in Figure 1) were also included in the analyses, because previous studies have shown that moderate sequence variation is observed at the ATF/CREB binding site and the NF-κB consensus sequence is typically very well conserved in peripheral blood-derived HIV-1 LTR sequences.

At each nucleotide position for the four sites examined, the most common nucleotide matched the consensus for HIV-1 clade B. This was not surprising as clade B is the most common HIV-1 subtype in North America. The NF-κ B binding site, in particular, was extremely well conserved as would be expected (Figure 1A). Almost all variation observed occurred at the first two nucleotides of the sequence with the NCF at the first nucleotide equal to only 3.2% for both brain- and PBMC-derived LTRs and 4% and 0.6% at the second position, respectively. As could be expected, the less conserved ATF/CREB sequence exhibited only moderate variability at several nucleotides (Figure 1B). The most obvious region of variability occurred at position 7, where the NCF was 21% for the PBMC-derived ATF/CREB sequences and 12.7% for the brain-derived sequences. There were no significant differences between the brain- and PBMC-derived LTR populations, however, with respect to the specific base pair (bp) alterations that occurred at position 7 of the ATF/CREB binding site. Most of the bp alterations at position 7 involved a T to C or T to G substitution in both LTR populations.

In examining the NCF for C/EBP site I (Figure 1C), an obvious point of variability not apparent with any of the other binding sites was observed. At bp position 6 (indicated by an asterisk), the NCF was 42.5% for the brain-derived LTRs, and only 5.8% for the PBMC-derived LTRs. Specific differences at bp position 6 will be discussed later. A second point of moderate variability was observed at bp position 3, where the NCF at this position in C/EBP site I within LTRs obtained from brain was 22.6%, while the NCF at this position in C/EBP site I within LTRs derived from peripheral blood was 10.7%. There were no significant differences between the brain- and PBMC-derived LTR populations, however, with respect to the specific bp alterations that occurred at position 3. Most alterations at this position in both

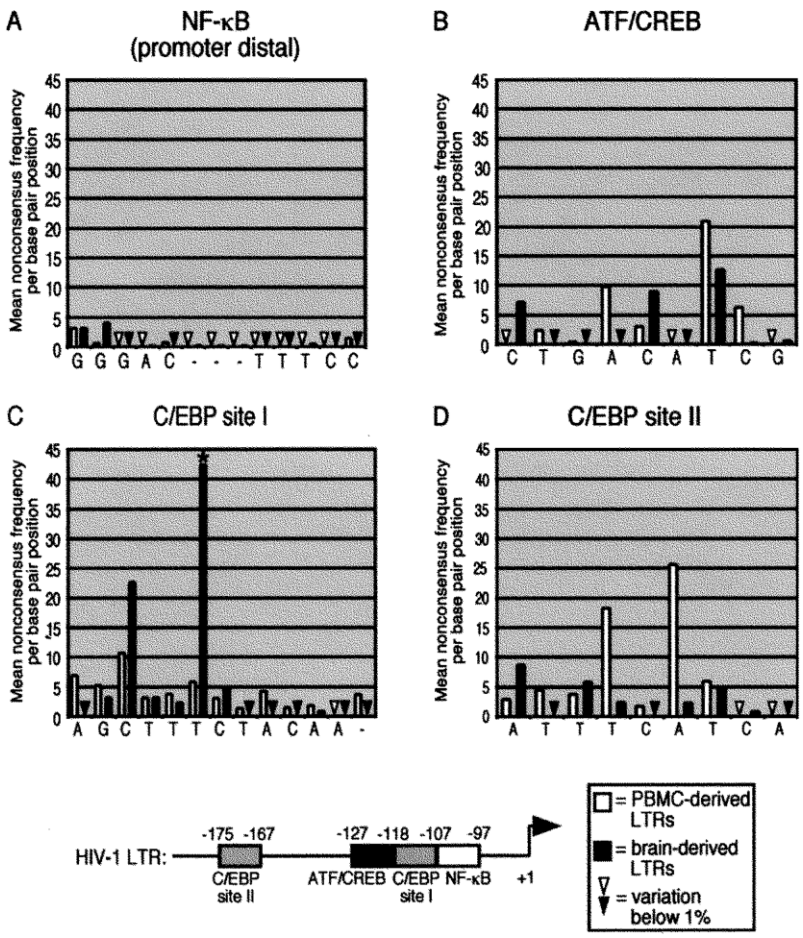

Figure 1 Brain-derived LTRs exhibit unique sequence configurations at C/EBP sites I and II when compared to peripheral blood-derived LTRs. LTR sequences derived from each individual were aligned and compared to the published clade B consensus sequence. The frequency of occurrence of a nonclade B nucleotide was calculated at each position for each individual as a percent of total LTRs examined. The mean nonconsensus frequency (NCF) was calculated from these values at each nucleotide position for the population of individuals and these values were plotted above the consensus sequence. These analyses were conducted for the HIV-1 LTR (**A**) NF-κB, (**B**) ATF/CREB, (**C**) C/EBP site I, and (**D**) C/EBP site II binding sites. The position of each nucleotide is identified below the consensus binding site sequence for each binding site examined. The single asterisk identifies a common guanine substitution that is frequently found in brain-derived LTRs. The analysis of brain-derived LTRs represented a total of 164 sequences derived from 18 individuals, while the analysis of PBMC-derived LTRs represented a total of 328 sequences derived from 40 individuals (Estable *et al*, 1996; Kirchhoff *et al*, 1997; Michael *et al*, 1994). The number of brain-derived LTRs examined per patient ranged from 3 to 21 and are summarized in Figure 3. The number of PBMC-derived LTRs examined per patient ranged from 3 to 42 and are summarized as described (Estable *et al*, 1996; Kirchhoff *et al*, 1997; Michael *et al*, 1994).

LTR populations involved a C to T or C to A bp substitution. All other nucleotide positions within C/EBP site I exhibited minimal sequence variation with respect to the clade B consensus sequence at this site.

Given that both C/EBP sites I and II have been shown to be involved in the regulation of HIV-1 gene expression and replication in monocytes, sequence variation within C/EBP site II was also examined. This analysis led to the identification of another striking difference between the two LTR populations (Figure 1D). The brain-derived LTRs exhibited very little sequence variation at any position within C/EBP site II, while the C/EBP site II in PBMC-derived LTRs exhibited variation across the sequence with two moderate peaks of nucleotide divergence at positions 4 and 6. The only point of C/EBP site II variation in LTRs derived from brain was the NCF of 8.8% at the first position of the binding site. However, the PBMC-derived LTRs exhibited an NCF of 18.3% at position 4 and an NCF of 25.6% at position 6.

C/EBP sites I and II are conserved to different degrees in brain- and peripheral blood-derived LTRs when compared to the clade B consensus sequence

Because the pattern of C/EBP site I and II conservation was so different, the mean NCF for the entire binding site for each of the four transcription factors was determined (Figure 2). Based on this analysis, it was determined that neither the ATF/CREB nor NF-κB binding site differed between the two LTR populations when compared to the clade B consensus

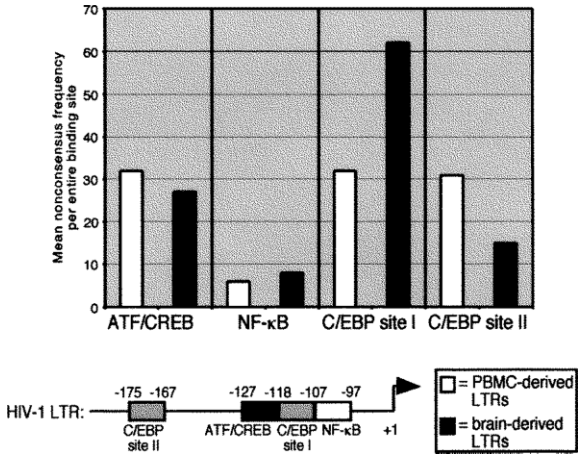

Figure 2 Conservation of C/EBP site I and II sequences differs between LTRs obtained from brain and peripheral blood. LTR sequences derived from each individual were aligned and compared to the clade B consensus sequence. The frequency of occurrence of a nonclade B binding site was calculated for each individual as a percent of total LTRs examined. The mean NCF was calculated from these values at each binding site for the population of individuals and these values were plotted above the consensus sequence. These analyses were conducted for the NF-κB, ATF/CREB, C/EBP site I, and C/EBP site II binding sites. The analysis of brain-derived LTRs represented a total of 164 sequences derived from 18 individuals, while the analysis of PBMC-derived LTRs represented a total of 328 sequences derived from 40 individuals (Estable *et al*, 1996; Kirchhoff *et al*, 1997; Michael *et al*, 1994). The number of brain-derived LTRs examined per patient ranged from 3 to 21 and are summarized in Figure 3. The number of PBMC-derived LTRs examined per patient ranged from 3 to 42 and are summarized as described (Estable *et al*, 1996; Kirchhoff *et al*, 1997; Michael *et al*, 1994).

sequence. The brain-derived LTRs exhibited a mean NCF of 27% for the ATF/CREB binding site, while the PBMC-derived LTRs exhibited a mean NCF of 32%. The mean NCF for the NF-κB binding site was low for both the brain and PBMC populations, 8% and 6%, respectively.

However, the two LTR populations differed greatly when conservation of the two C/EBP binding sites was examined. For C/EBP site I, the mean NCF for the peripheral blood-derived LTR population was 32% (the same mean NCF value observed for the ATF/CREB binding site within the same LTR population). However, the brain-derived LTRs exhibited a mean NCF of 62%, exhibiting almost double the amount of sequence variation at C/EBP site I when compared to peripheral blood LTR populations (statistically significant difference; $P = .027$). With regard to C/EBP site II, the level of conservation was reversed. The mean NCF of the peripheral blood-derived LTR C/EBP site II was 31%, almost identical to the amount of sequence variation observed at C/EBP site I and the ATF/CREB site within the peripheral blood LTR population. However, the brain-derived LTR population exhibited a C/EBP site II mean NCF of only 15%, less than half the variation observed with the PBMC-derived LTR populations (statistically significant difference; $P = .056$). For the LTRs obtained from brain tissue, only the NF-κB binding site (of the four sites examined) was more conserved than C/EBP site II in these analyses.

Sequence variation at C/EBP site I of brain-derived LTRs is often represented by a single C/EBP binding site I configuration

To investigate the actual bp alterations that constituted the 42.5% NCF at position 6 of C/EBP site I, all brain-derived LTRs that deviated from the clade B consensus sequence at position 6 were examined. Most alterations detected involved a T to G change (as indicated previously, this site has been designated the 6G C/EBP site I). Brains containing LTRs with the 6G configuration were only scored positive for this configuration if they did not contain any variant nucleotides other than the T to G change at position 6. Of all brain-derived LTRs containing a bp alteration at position 6, 73% of the LTRs possessed a 6G binding site configuration (AGC TTG CTA CAA; data not shown). Given that 42.5% of a given patient's brain-derived LTRs contained a bp change at position 6 (Figure 1C) and 73% of these were the 6G configuration, 31% of any given patient's LTRs obtained from brain contained the 6G C/EBP site I configuration. Conversely, when a search was performed for the 6G site I configuration in the 328 LTRs derived from peripheral blood, only 16 of the LTRs (from 3 of 40 patients) contained the 6G site I configuration, suggesting that the frequency of this configuration within the peripheral blood was very low.

The majority (60%) of the C/EBP site I configurations encountered in brain-derived LTRs were either the clade B consensus sequence or the 6G variant. Both LTR site I configurations were represented within about the same number of individuals (data not shown). The clade B consensus sequence site I configuration was found within 61% of individuals while the 6G binding site appeared in 56% of all individuals. Within positive brains, the 6G site I configuration constituted 61% of an individual's C/EBP site I LTR sequences, while the clade B sequence configuration accounted for 49% of an individual's site I LTR sequences.

Distribution of common C/EBP site I and II configurations within each brain analyzed

The distribution of the 6G and clade B consensus C/EBP site I LTR sequences within all available brain specimens analyzed was also examined for C/EBP site I (Figure 3A), and compared to the distribution of the commonly encountered clade B consensus C/EBP site II sequence (Figure 3B). In only one of the brains examined (brain 510) was an LTR not detected that did not contain either the 6G or clade B consensus sequence at C/EBP site I (Figure 3A). Several of the HIV-1-positive brains (brains 2435, 2597, and 2710) contained only LTRs with the 6G

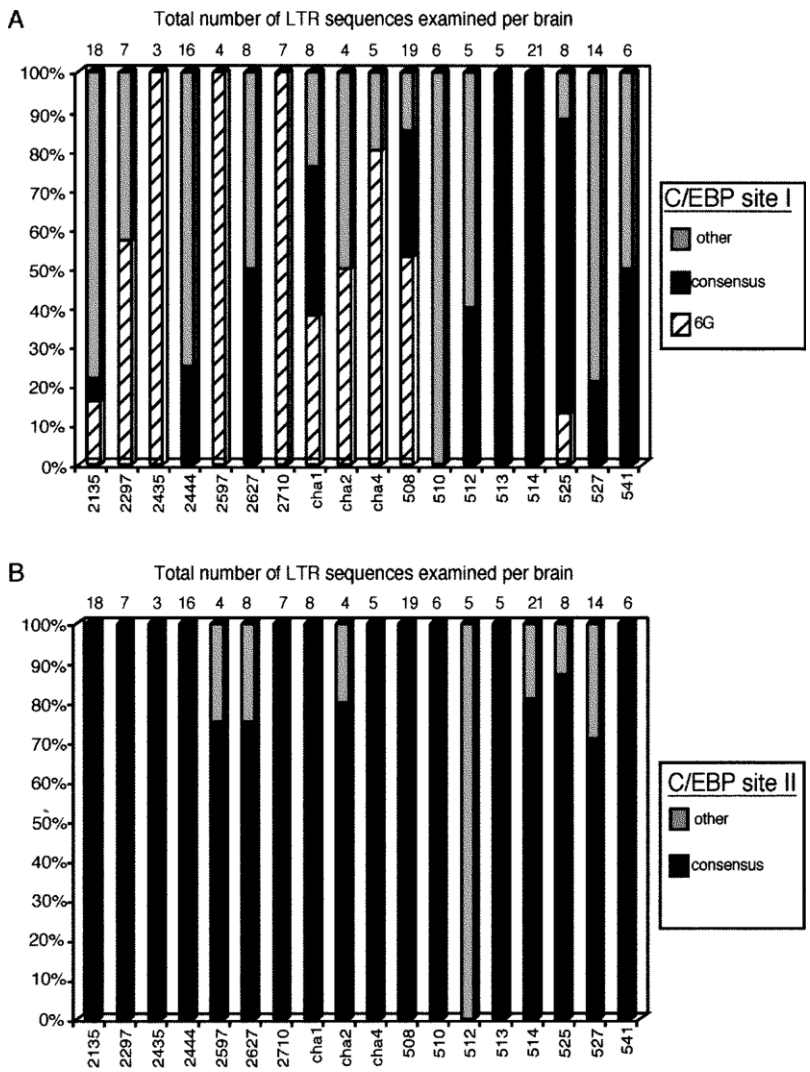

Figure 3 Distribution of predominant C/EBP binding sites I and II in each HIV-1-positive brain. (**A**) The distribution of 6G and clade B consensus C/EBP site I configurations is illustrated for each brain investigated. 'Consensus' designates the clade B C/EBP site I consensus sequence, while '6G' designates a T to G nucleotide substitution at position 6 of the clade B consensus sequence of C/EBP site I. "Other" designates any sequence that does not fall into one of these two categories. (**B**) The distribution of the consensus clade B consensus C/EBP site II configuration is illustrated for each brain investigated. All configurations that deviated from the clade B consensus sequence were designated as "other." The total number of sequences analyzed per brain is shown above the appropriate bar.

C/EBP site I configuration. Conversely, brains 513 and 514 only contained LTRs with the clade B consensus sequence. The remaining brains contained different levels of LTRs with the two C/EBP site I sequence configurations. The pattern of sequence distribution with respect to C/EBP site II was strikingly different (Figure 3B). Eleven of the brains examined only contained LTRs with the clade B consensus C/EBP site II configuration. Of the remaining six brains, only brain 512 did not contain a single LTR with the clade B consensus sequence at C/EBP site II. The remaining five brains exhibited about an 80% retention of LTRs with the clade B C/EBP site II configuration. Overall, the high degree of conservation of the clade B C/EBP site II configuration and the commonly encountered variations observed at C/EBP site I in brain-derived LTRs was strikingly different than the C/EBP sites observed in the peripheral blood LTR population.

Apart from the interesting point of variability at C/EBP site I, position 6, another point of variability was observed at C/EBP site I position 3, with an NCF of 22.6% in the brain-derived LTR population. When the variation at position 3 was examined, 28% of the bp changes at this site led to a configuration (3T: AGT TTT CTA CAA) with diminished reactivity toward C/EBP proteins (data not shown). Studies by other investigators have shown that C/EBP sites I and II can functionally compensate for one another within the HIV-1 LTR when one of the two C/EBP binding sites is inactivated and examined in functional analyses (Henderson and Calame, 1997; Henderson *et al*, 1996). Therefore, the C/EBP site II sequences

for the 14 brain-derived sequences that contained the weak 3T C/EBP site I were examined. When the corresponding C/EBP site II sequences were examined, almost all of the variants had the consensus clade B sequence (ATT TCA TCA), which exhibits an increased ability to bind C/EBP factors. Two other sequences were observed at C/EBP site II (6G: ATT TCG TCA) and (6G8T: ATT TCG TTA), both of which recruit high amounts of C/EBP protein. Therefore, of the brain-derived 3T C/EBP site I sequences that exhibited reduced ability to bind C/EBP factors, all of them appeared to have a compensatory strong C/EBP site II.

C/EBP sites preferentially encountered in brain-derived LTRs exhibit increased ability to bind C/EBP factors

To characterize the effect the common 6G C/EBP site I binding site configuration may have on C/EBP factor binding, relative to the also common clade B C/EBP site I, EMS analyses were performed. Radiolabeled double-stranded oligonucleotide probes that represented either the 6G C/EBP site I (Figure 4A, lanes 1–5) or the clade B consensus C/EBP site I (Figure 4A, lanes 6–10) were reacted with nuclear extract prepared from the U-937 monocytic cell line. These reactions were also incubated with antisera directed against the C/EBP activator proteins C/EBP α, C/EBP β, and C/EBP ε, in antibody supershift EMS analyses. It was apparent that there was stronger DNA-protein complex formation with the 6G variant (Figure 4A, lane 1) as compared to the clade B consensus C/EBP site I (Figure 4A, lane 6). The supershift EMS analyses confirmed this observation (Figure 4A, lanes 3 and 4 compared to lanes 8 and 9). Although only a moderate amount of C/EBP α was supershifted with the DNA-protein complexes formed with the clade B consensus sequence (Figure 4A, lane 8), a much greater amount of C/EBP protein was supershifted with the DNA-protein complexes formed with the 6G variant probe (Figure 4A, lane 3). Additionally, C/EBP β was barely detectable in complexes formed with the clade B consensus sequence (Figure 4A, lane 9), while a significant amount of C/EBP β was shifted from the complexes formed with the 6G C/EBP site I variant probe (Figure 4A, lane 4). This was important because C/EBP β is a more efficient *trans*-activator than C/EBP α (data not shown). These studies illustrated that brain-derived LTRs that frequently possess the 6G C/EBP site I configuration exhibited an enhanced ability to bind C/EBP factors than was observed with the clade B consensus C/EBP site I.

To determine the ability of the highly conserved C/EBP site II consensus clade B sequence variant to recruit C/EBP factors, an oligonucleotide representing the consensus C/EBP site II was reacted with U-937 nuclear factors in EMS analyses, in parallel with the 6G C/EBP site I oligonucleotide (Figure 4B). A much more abundant DNA-protein complex was observed with the conserved C/EBP site II sequence than was observed with the 6G C/EBP site I. Antibody supershift EMS analyses utilizing antisera directed against four of the common C/EBP *trans*-activators (Figure 4C) were performed. As shown, readily detectable quantities of C/EBP α, C/EBP β, and C/EBP ε were observed utilizing the U-937 nuclear extract. However, no C/EBP δ binding was detected (the C/EBP δ antibody has been shown to form supershifted complexes when reacted with extract derived from the THP-1 monocytic cell line; data not shown).

To further characterize the relative binding affinities of the commonly encountered clade B consensus C/EBP site II and the 6G C/EBP site I and to demarcate the binding sites, DNase I protection analysis was conducted using a radiolabeled probe (204 bp) that contained both of these binding sites (Figure 4D). The radiolabeled probe was reacted with increasing amounts of His-tagged C/EBP β (Figure 4D, lanes 1– 5) and the regions of protein protection were compared with probe incubated only with DNase I (Figure 4D, lane 6). By comparison to the reaction containing only DNase I, even with the lowest amounts of purified C/EBP β, a strong region of DNase I protection was observed at the clade B consensus C/EBP site II. Conversely, a region of protection was not observed at the 6G C/EBP site I until higher amounts of protein were added. These results indicate that the clade B consensus C/EBP site II has a much higher relative affinity for C/EBP β than does the 6G C/EBP site I.

In summary, the 6G C/EBP site I sequence has a higher relative affinity for C/EBP proteins than does the clade B C/EBP site I. However, the clade B C/EBP site II has an even higher relative affinity for C/EBP factor binding than the 6G C/EBP site I. Given that the clade B consensus sequence is highly conserved at the C/EBP site II, it can be concluded that a vast majority of brain-derived LTRs possess a C/EBP binding site II that is highly reactive with respect to binding members of the C/EBP family of transcription factors.

Brain-derived LTRs contain C/EBP site II configurations that are almost always highly reactive with respect to C/EBP proteins (a vast majority of LTRs contain the clade B consensus site II configuration)

Given the high level of clade B consensus C/EBP site II conservation in brain-derived LTRs, the binding phenotypes of the remaining C/EBP site II configurations encountered in brain were examined, to ascertain whether conservation of a strong C/EBP site II, and not just conservation of the clade B configuration, was found in the brain-derived LTRs. Ten C/EBP site II configurations were encountered in the brain sequence database, which differed from the clade B consensus site (Figure 5A). The configurations were named according to the position at which there was a bp substitution or deletion (del) and the nucleotide that was substituted. The 10 variants

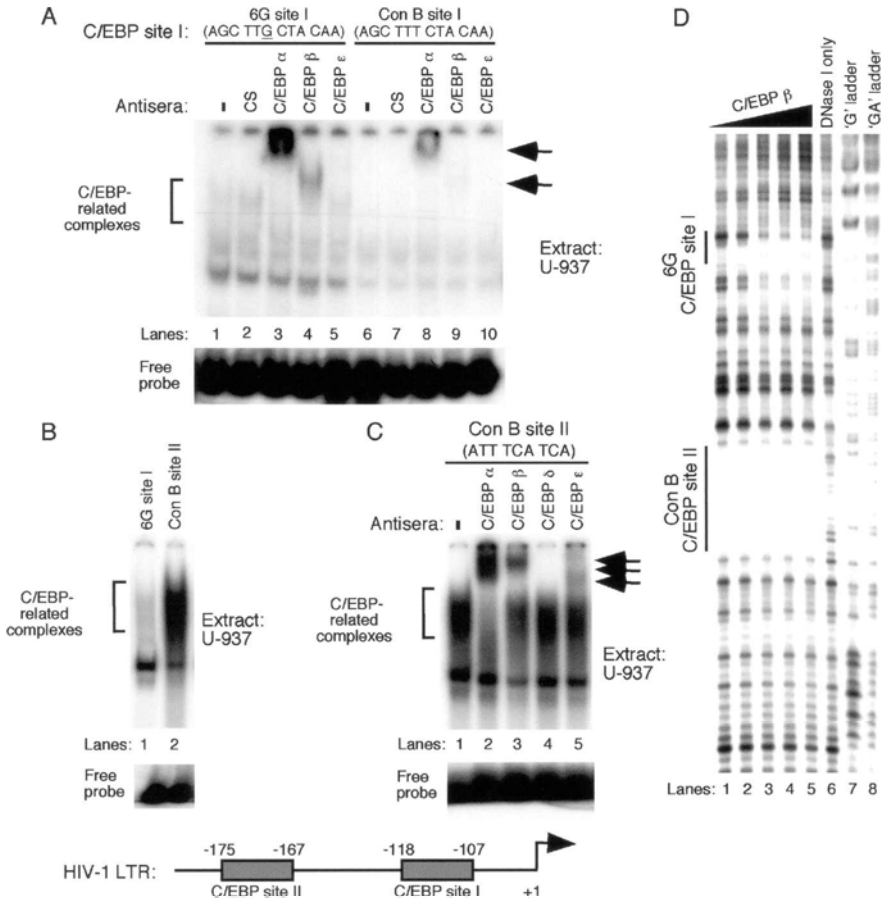

Figure 4 C/EBP sites preferentially encountered in brain-derived LTRs exhibit increased ability to bind C/EBP factors. (**A**) C/EBP site I oligonucleotides representing the 6G (lanes 1–5) and the clade B consensus sequence (lanes 6–10) were reacted with U-937 nuclear extract. (**B**) For comparison, the clade B consensus C/EBP site II and the 6G C/EBP site I were reacted with U-937 nuclear extract. (**C**) Antisera to C/EBP α, C/EBP β, C/EBP δ, and C/EBP ε were added to the EMS reactions to identify individual members of the C/EBP transcription factor family, while control rabbit immunoglobulin (CS) was added to lanes 2 and 7 in panel A. Arrows to the right of the panels indicate supershifted complexes. Free probe is indicated at the bottom of panels A, B, and C. (**D**) DNase I protection analysis was conducted using a 200-bp probe that contained the 6G C/EBP site I and the clade B consensus C/EBP site II. The probe was reacted with increasing amounts of His-tagged C/EBP β (lanes 1–5) and regions of DNase I protection were compared with probe incubated with DNase I only (lane 6). Sequencing ladders of the probe were also included to confirm the regions of DNase I protection.

were present at different degrees within the brain-derived LTR population, with the clade B consensus representing 88.6% of the 164 total brain-derived LTRs.

Of the 10 brain-derived C/EBP site II configurations that deviated from the clade B consensus sequence, patient information was known for the LTRs that contained six of the variant sequences. The variant configurations 7G, 1del, 1G, and 1T (Figure 5A) were all encountered in the brain-derived LTRs for which dementia status was not known. However, the remaining configurations (1C4C7C, 1C3C, 6G8T, 6G, 1C3C7C, and 4C) accounted for 9% of the C/EBP site II configurations encountered in all of the brain-derived LTRs. Interestingly, these six sequences were all derived from the LTRs of four patients (512, 514, 525, and 527; Figure 3) who were known to exhibit clinical signs of dementia. No sequence variants were encountered in the three patients known to exhibit no clinical signs of dementia (patients 508, 510, and 513; Figure 3). This limited information suggests that sequence variability at C/EBP site II should be considered when considering factors that correlate with the development of clinical dementia. There are very likely other factors involved in determining CNS disease progression.

To characterize the relative binding affinities of the brain-derived sequence variants, double-stranded radiolabeled oligonucleotide probes spanning each of the brain-derived C/EBP site II variants, including the clade B consensus sequence (Figure 5B, lanes 2–12) were reacted with U-937 nuclear extract which had been stimulated with IL-6 (to facilitate more efficient visualization of the C/EBP-containing DNA-protein complexes) in EMS analyses. Included in the analyses was a probe spanning

A

Brain-derived C/EBP site II configurations		Frequency within brain-derived LTRs
Con B	ATT TCA TCA	88.6%
7G	ATT TCA GCA	0.6%
1 del	-TT TCA TCA	0.6%
1G	GTT TCA TCA	0.6%
1T	TTT TCA TCA	0.6%
1C4C7C	CTT CCA CCA	0.6%
1C3C	CTC TCA TCA	1.8%
6G8T	ATT TCG TTA	0.6%
6G	ATT TCG TCA	2.4%
1C3C7C	CTC TCA CCA	1.2%
4C	ATT CCA TCA	2.4%

B

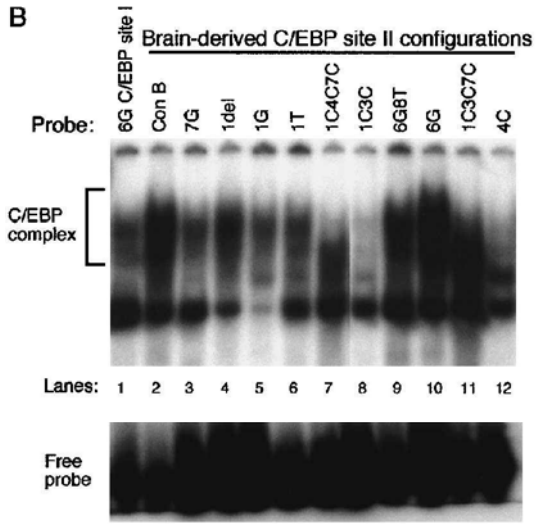

Figure 5 Brain-derived non-clade B C/EBP site II configurations exhibit a wide array of binding phenotypes but a high percentage are highly reactive with respect to binding C/EBP factors. (**A**) Ten C/EBP site II configurations other than the clade B consensus sequence were encountered within the brain-derived LTR population at very low frequencies. (**B**) Oligonucleotide probes representing all brain-derived C/EBP site II configurations, as well as a probe spanning the highly reactive 6G C/EBP site I, were reacted with IL-6-induced U-937 nuclear extract in EMS analyses. Brackets to the left of the panel indicate the C/EBP-related complexes. Free probe is indicated at the bottom of panel B.

the highly reactive 6G C/EBP site I (Figure 5B, lane 1), commonly encountered in brain. It was apparent that two of the variants, 6G (Figure 5B, lane 10) and 6G8T (Figure 5B, lane 9), both retained strong binding sites at C/EBP site II. The remainder of the variants exhibited moderate to greatly reduced reactivity toward C/EBP proteins. Interestingly, two of the variants, 1C4C7C (Figure 5B, lane 7) and 1C3C7C (Figure 5B, lane 11), exhibited strong DNA-protein complex formation of altered mobility, which contained only low levels of C/EBP protein.

Of the six C/EBP site II sequence variants (1C4C7C, 1C3C, 6G8T, 6G, 1C3C7C, and 4C) that constitute 9% of the total C/EBP site II sequences and were derived from patients exhibiting clinical signs of dementia, differing levels of C/EBP recruitment were encountered. Two of the variants (6G and 6G8T, representing 3% of the total brain-derived LTRs) recruited high amounts of C/EBP, two variants (4C and 1C3C, representing 4.2% of the total brain-derived LTRs) recruited low levels of C/EBP, and two variants (1C4C7C and 1C3C7C, representing 1.8% of the total brain-derived LTRs) recruited a protein other than C/EBP. Cumulatively, these results indicate that 93% of the 164 brain-derived LTRs analyzed were highly reactive with respect to C/EBP protein binding at site II. In addition, increased variability at C/EBP site II may correlate with clinical dementia and often leads to alterations in C/EBP factor recruitment at C/EBP site II.

C/EBP site II strength impacts basal and Tat-induced LTR activity in a HIV-1 strain-dependent manner

To examine the role of C/EBP site II in the context of LTRs derived from HIV-1 molecular clones that exhibit different cellular tropisms, luciferase reporter plasmids were constructed that were under the control of HIV-1 LTRs derived from the T-tropic LAI, dual tropic 89.6, and brain-derived, M-tropic YU-2 HIV-1 infectious molecular clones. The parental LAI and YU-2 LTR-luciferase constructs both possess strong clade B consensus configurations at C/EBP site II, while 89.6 possesses a weakly reactive 4C configuration (Figure 5B, lane 12). LTR-luciferase reporters were also constructed with each of the three parental backbones containing a C/EBP site II (5G7G: ATT TGA GCA) that exhibited very weak reactivity with respect to binding C/EBP proteins.

Each of these constructs were transiently transfected into U-937 cells in the absence or presence of HIV-1 *trans*-activator protein Tat co-expression (Figure 6). It was immediately obvious that the three different parental backbones exhibited divergent activities with respect to both basal and Tat-induced *trans*-activation (Figure 6A). The LTR derived from the T-cell line adapted LAI strain exhibited both the highest basal and Tat-induced activity when compared to LTRs derived from either 89.6 or YU-2. The luciferase construct containing the LTR derived from 89.6 exhibited a basal activity that was slightly below that of the LAI LTR, while the YU-2 LTR exhibited about a 30% reduction in basal activity as compared to the other parental LTR constructs. The Tat-induced activity obtained with each of the three LTR constructs was even more highly divergent. While the LAI LTR activity increased about 30-fold following Tat induction, the 89.6 LTR only exhibited about a 20-fold increase, and the YU-2 LTR was *trans*-activated by about 10-fold. Therefore, the parental constructs exhibited very divergent LTR activities when transfected into the monocytic U-937 cells in the absence or presence of Tat.

When C/EBP site II was converted to a weakly reactive binding site (5G7G) within the context of the three parental LTR backbones, basal activity was impacted with the LAI and 89.6 LTRs, but not with the

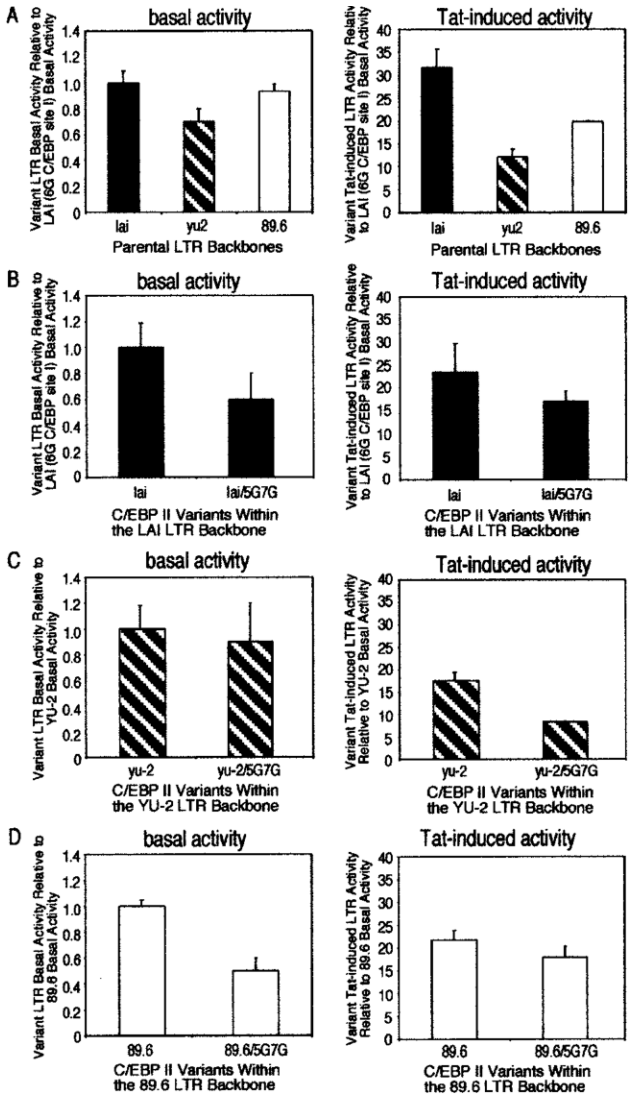

Figure 6 C/EBP site II sequence variation impacts basal and Tat-induced *trans*-activation in a HIV-1 strain-dependent manner. (**A**) LTRs derived from the LAI, YU-2, and 89.6 strains of HIV-1, were transiently transfected into U-937 cells to compare basal and Tat-induced LTR activity of each of these parental backbones. The weak 5G7G C/EBP site II configuration was also substituted into each of the parental backbones, (**B**) LAI, (**C**) YU-2, and (**D**) 89.6, and LTR activities were determined in the absence or presence of Tat. The luciferase activity of each of the chimeric LTRs was normalized to an arbitrary activity level of 1 for the LAI (6G) LTR. Each transient transfection was done in duplicate, with three independent experiments.

YU-2 LTR (Figure 6B–D). The LAI and 89.6 LTRs exhibited about 40% and 60% reductions in basal activity, respectively. This pattern was reversed with respect to Tat-induced activity (Figure 6B–D). Tat-induced activity with the LAI and 89.6 LTRs resulted in a minor reduction in activity, while the YU-2 LTR exhibited about a 50% reduction in response to Tat *trans*-activation. This indicated that a strong C/EBP site II phenotype was not important for basal activity of brain-derived LTRs, but was critical for highly efficient Tat *trans*-activation.

Discussion

HIV-1 has been demonstrated to exist as rapidly evolving quasi-species that have been described as swarms of genetic variants existing simultaneously within an infected individual (Pang *et al*, 1992). The composition of the most abundant variants within an individual changes over time as a result of frequent mutations introduced by the reverse transcription process, recombination between diploid genomic RNA, and strong selective pressures of the host immune response. This results in a dynamic interplay between the virus and host. Differences between blood- and CNS-derived HIV-1 strains have been documented in studies that examined sequence variation within the envelope V3-V5 region (Korber *et al*, 1994). It was suggested that monocyte/macrophage tropism of brain-derived viruses provided a biological constraint that resulted in conservation of specific amino acids within viral envelope proteins.

Whereas genetic variation within the antigenic V3 loop of the envelope gene has been suggested to be driven by the host immune system (Delwart *et al*, 1994; Lukashov *et al*, 1995), and conservation at positions in the V3 region may reflect the cell type or tissue tropism of HIV-1 variants (Korber *et al*, 1994), the forces that drive genetic variation and conservation of LTR sequences are less well understood. Studies involving LTR sequences derived from multiple tissues in one individual have demonstrated that LTRs derived from specific tissues can be classified phylogenetically as distinct variants, suggesting that LTRs may evolve in a tissue compartmentalized manner (Ait-Khaled *et al*, 1995). Evidence that LTR sequences derived from brain are functionally distinct from blood-derived LTRs also exists.

Transgenic mice were constructed that carry different HIV-1 LTRs attached to β-gal expression constructs. The three LTRs employed in the studies were obtained from lymph node (HIV-1 strain IIIB/LAV) or the CNS (HIV-1 strains JR-CSF and JR-FL). Only animals carrying the CNS-derived LTRs directed β-gal expression within the CNS, particularly in neurons. In addition, the specific neuroanatomical locations of β-gal expression differed slightly between the divergent CNS-derived LTRs (Corboy *et al*, 1992). However, phylogenetic analysis of a number of LTRs derived from brain tissue of four individuals did not demonstrate any simple association between specific nucleotides of brain-derived sequences and the expression results obtained with the CNS-derived LTRs in the transgenic analyses (Corboy and Garl, 1997), making it difficult to ascribe a CNS-specific function to specific *cis*-acting element sequence configurations within the LTR.

Sequence analysis reported herein has indicated that C/EBP site sequence variation was very different between LTRs obtained from peripheral blood and those obtained from brain. The peripheral blood-derived LTR populations exhibited a level of

sequence variation at these sites comparable to that of the ATF/CREB binding site (Figures 1 and 2). However, based on the sequence variation observed with LTR populations obtained from brain, both C/EBP binding sites appear to have important roles in HIV-1 replication within the brain. The clade B consensus C/EBP site II that exhibited a very high reactivity with C/EBP monocytic nuclear factors (Figure 4) was very well conserved in brain-derived LTRs, with only 15% of any given individual's LTRs deviating from the clade B consensus sequence. This was half of the variation observed at C/EBP site II in peripheral blood-derived LTRs. In fact, the site II variation approached that observed with the highly conserved NF-κB binding site (which exhibited an NCF of 8% in brain-derived LTRs). Furthermore, when the brain-derived C/EBP site II variant configurations that deviate from the clade B consensus site were examined, 93% of the total brain-derived LTRs exhibited a highly reactive C/EBP site II with respect to binding members of this transcription factor family.

In contrast, brain-derived LTRs exhibited far more variation at C/EBP site I as compared to peripheral blood-derived sequences. While LTRs derived from peripheral blood contained a NCF of 32%, brain-derived LTR sequences exhibited a NCF of 62%. In particular, a NCF of 42% was observed at C/EBP site I, position 6, and 73% of these changes were to the same 6G variant configuration. When the impact of the 6G variation at C/EBP site I was compared to the clade B consensus sequence at site I, the 6G configuration exhibited an enhanced capability of binding C/EBP factors (Figure 4). The well-conserved clade B C/EBP site II exhibited an even greater ability to bind C/EBP factors as compared to both the C/EBP site I clade B consensus and the more reactive 6G C/EBP site I configuration (Figure 4). Thus, brain-derived LTRs often possess two very strong C/EBP binding sites, an observation not observed with peripheral blood-derived LTRs. Furthermore, when a weakly reactive binding site was observed at C/EBP site I, a highly reactive site was always observed at C/EBP site II.

Interestingly, when the C/EBP site II of a T-cell tropic (LAI), dual-tropic (89.6), and brain-derived M-tropic (YU-2) strain was mutated to a weak binding site, a differential impact on basal and Tat-induced LTR activity was observed in transient expression analyses. A weakly reactive C/EBP site II significantly decreased basal *trans*-activation of the LAI and 89.6 LTRs, but did not impact the YU-2 LTR. Conversely, a weakly reactive C/EBP site II significantly inhibited Tat-induced YU-2 activity but did not significantly impact either the LAI or 89.6 Tat *trans*-activation. These results suggest that the strong conservation of the highly reactive clade B C/EBP site II within brain-derived LTRs may be critical for highly efficient Tat-dependent replication of HIV-1 within the brain. It should be noted that the Tat protein used in these studies was derived from the HIV-1 IIIB strain. It is probable that the HIV-1 Tat protein and the HIV-1 LTR co-evolve. Thus, while the Tat protein derived from IIIB did not *trans*-activate the YU-2-derived LTR in these studies, it is possible that the Tat protein derived from the HIV-1 YU-2 strain would *trans*-activate the YU-2 LTR to a greater degree.

One could predict that C/EBP site sequence variation would have more impact on HIV-1 replication in brain, given that viral replication is primarily detected in cells of the monocyte/macrophage lineage and C/EBP factors appear to play such an important role in normal monocyte/macrophage gene regulation. Most peripheral blood-derived LTRs are obtained from virus replicating in lymphocytes rather than from cells of the monocyte/macrophage lineage. In addition, C/EBP proteins are not expressed to significant levels within lymphocytic populations, rendering C/EBP binding site sequence variation of less importance to a PBL-derived virus. However, the C/EBP binding sites are still well conserved in the peripheral blood LTRs, to the same extent as the ATF/CREB binding site (Figure 2).

Blood samples from all individuals from whom brain LTRs were obtained were not available for analysis. It will be interesting to determine the extent of C/EBP site conservation in the peripheral blood of infected individuals, to investigate whether the sequence patterns of the two C/EBP sites observed in brain-derived viruses are brain-specific, or whether they are also observed in peripheral blood (particularly in the monocyte/macrophage compartment).

This leads to the question of whether or not the selectivity for strong C/EBP binding sites is observed in monocyte/macrophage populations that are circulating in the blood before they infiltrate the brain. To address this question, one would need to sequence a number of PCR-amplified LTR clones or viral clones derived from monocyte and lymphocyte populations present in peripheral blood, as well as from brain tissue from a large number of individuals. Although access to this type of database is not yet available in an extensive format, we have characterized the reactivity of C/EBP sites I and II for C/EBP proteins, for several published infectious molecular clones (data not shown). The HIV-1 infectious molecular clones were either derived from the CNS or were isolated from non-CNS compartments including, peripheral blood, infant lung tissue, and lymph node (Benson *et al*, 2000). The four clones that were isolated from CNS tissue (brain tissue or CSF) include: SF162, JR-FL, JR-CSF, and YU-2. The 10 clones that were isolated from non-CNS reservoirs include: BAL, AD8, LAI/IIIB, MN, ELI, RF, NDK, NL43, SF2, and 89.6. The repertoire of non-CNS-derived clones was diverse and included molecular clones that have been characterized as T-tropic, M-tropic, and dual-tropic. In addition, the molecular clones have been characterized as either syncytium inducing (SI) or non-syncytium inducing (NSI) and have been shown to utilize a variety of coreceptors, including but not limited to CXCR-4 and CCR5.

Differences in binding phenotype of C/EBP sites I and II were observed between the different infectious molecular clone populations. The viral clones derived from the CNS mimicked the results observed with the PCR-amplified brain sequences examined herein. All four clones possessed a strong C/EBP site II (three of which were the clade B configuration). None of the four contained the highly reactive 6G configuration at C/EBP site I, but three of the clones contained the also prominent clade B consensus C/EBP site I, which recruited a moderate level of C/EBP protein. Only SF162 contained a divergent sequence at C/EBP site I, that slightly decreased the level of C/EBP protein recruited. In contrast, of the 10 molecular clones derived from non-CNS reservoirs, four of the clones exhibited greatly reduced C/EBP binding at C/EBP site II. In addition, four of the clones also exhibited negligible binding at C/EBP site I.

These results would seem to indicate that the selection of two highly prevalent strong C/EBP binding sites within an LTR population does not occur until after the virus reaches the CNS, since the C/EBP site configurations encountered in the CNS-derived molecular clones recruited greater amounts of C/EBP protein than the C/EBP binding sites encountered in the clones isolated from non-CNS reservoirs. Given that 93% of the brain-derived LTRs possessed a C/EBP site II that was highly reactive with respect to binding C/EBP proteins, it was clear that a strong C/EBP site II was either required for entry into the brain or was required for maintenance of the virus in the CNS. Additional studies will need to be performed to address this issue in greater detail.

With respect to viral replication within the peripheral blood, it is also probable that viral strains with highly reactive binding sequence configurations at C/EBP sites I and II have an evolutionary advantage over those that do not, since early stages of HIV-1 infection are characterized by M-tropic, NSI viruses, while T-tropic, SI viral isolates evolve later (Choe et al, 1996; Deng *et al*, 1996; Doranz *et al*, 1996; Dragic *et al*, 1996; Feng *et al*, 1996). Those viral isolates that are better equipped to establish a highly productive infection in monocyte/macrophage populations may be better suited to establishing a productive infection in a new host. Given the impact of a weakly reactive C/EBP site II on Tat-induced replication of the M-tropic YU-2 isolate within the U-937 monocytic cell line (Figure 6C), a strong C/EBP site II appears to be important for viral replication within the context of a brain-derived M-tropic HIV-1 strain. Future studies must be conducted analyzing additional clinical samples, viral isolates, PCR-generated sequences, and patient information. This would facilitate the development of a better understanding of the role C/EBP proteins play in virus replication within both the peripheral blood and CNS, and their potential involvement in disease progression within both of these compartments.

Materials and methods

Cell culture and nuclear extract preparation

The U-937 (ATCC# CRL-1593.2) human monocytic cell line was grown in RPMI 1640 medium. The media was supplemented with 10% heat-inactivated fetal bovine serum, antibiotics (penicillin, streptomycin, and kanamycin, each at a concentration of 0.04 mg/mL), L-glutamine (0.3 mg/mL), and sodium bicarbonate (0.05%). The cells were maintained at 37 °C in 5% CO_2 at 90% relative humidity. U-937 cells treated with recombinant human IL-6 (Genzyme) were prepared by treating cells for 24 h prior to the preparation of nuclear extracts, with (875 U/ml per 10^6 cells). Nuclear extracts were prepared as described (Dignam *et al*, 1983).

Oligonucleotide synthesis and radiolabeling

Complementary single-stranded oligonucleotides corresponding to the published C/EBP sequences within the HIV-1 LTR were synthesized (Macromolecular Core Facility, Penn State College of Medicine, Hershey, PA) and annealed by brief heating at 100 °C followed by slow cooling to room temperature. Blunt-ended, double-stranded oligonucleotides were end-labeled using [γ-^{32}P] ATP and T4 polynucleotide kinase as described (Promega). The sequence of the probes used in these studies were as follows: Con B site I: TCG AGC TTT CTA CAA GGG; 6G site I: TCG AGC TTG CTA CAA GGG; and Con B site II: CTA GCA TTT CAT CAC GTG. The flanking sequences on either side of the C/EBP binding sites were those present in the LAI LTR. The C/EBP binding sites are indicated as underlined sequences.

DNase I protection analyses

The probe used in the DNase I protection analyses spanned nucleotides −40 to −240 and contained the HIV-1 strain LAI sequence. The probe was generated by radiolabeling the upstream primer (5′-GGG CTC GCC ACT CCC CAG TCC CGC CCA GGC-3′) in a kinase reaction using γ-p^{32}. The labeled primer and the downstream primer (5′-GAG CCT GCA TGG AAT GGA TGA CCC TGA GAG-3′) were then included in a PCR reaction with a plasmid that contained the LAI LTR. The resulting 200-bp radiolabeled DNA product was isolated using the Qiaquick PCR purification kit (Qiagen). The DNase I protection reactions were performed as described (Promega Core Footprinting System) and subjected to electrophoresis on an 8% polyacrylamide sequencing gel. The probes were also sequenced by the Maxam–Gilbert chemical sequencing procedure (Maxam and Gilbert, 1980) and the sequencing reactions were subjected to electrophoresis in parallel with the DNase I

protection reactions to verify the position of protein binding.

Electrophoretic mobility shift (EMS) analyses
EMS binding reactions were performed as described (McAllister *et al*, 2000) and subjected to electrophoresis (30 mA and 200 V) in a 5% high ionic strength native polyacrylamide gel. For supershift EMS analyses, the reactions were incubated with antibody (2 μg/μl) obtained from Santa Cruz Biotechnology, Inc.

Plasmids and site-directed mutagenesis
The luciferase reporter gene constructs containing the LTRs derived from the LAI and YU-2 HIV-1 molecular clones were constructed as described (McAllister *et al*, 2000). The parental 89.6 LTR was obtained by Pfu-mediated PCR amplification of the strain 89.6 HIV-1 molecular clone (Kim *et al*, 1995). The 89.6 molecular clone was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH; p89.6 provided by Ronald G. Collman, University of Pennsylvania. Specifically, primers TGCACTGCAGACTGGAAGGGCT and CCCTCTAGACTGCTAGAGAT were used to amplify and subsequently clone the LTRs into the Xba I and Pst I sites of a modified multiple cloning region within the pGL3-Basic Vector, generating YU-2-Luc and 89.6-Luc. The three parental luciferase constructs were then used as templates for site-directed mutagenesis. The QuickChange mutagenesis kit (Stratagene) was used to construct the weakly reactive C/EBP site II 5G7G configuration (ATT TGA GCA) in each of the parental backbones (altered nucleotides as compared to the consensus clade B site are underlined). All plasmids used in these studies were sequenced to verify the C/EBP binding site sequence configurations (Penn State College of Medicine Macromolecular Core Facility).

Protein purification
We utilized a polyhistidine-tagged C/EBP β construct (C/EBP β-BD-pRSET A) obtained from Dr. Edward Maytin (Lerner Research Institute, Cleveland, OH). This plasmid was transformed into *Escherichia coli* strain BL21 (DE3) pLysS and transformants were selected using ampicillin (50 μg/ml) and chloramphenicol (35 μg/ml). Expression of the 6Xhistidine-tagged protein was induced for 5 h with 1mM isopropyl β-D-thiogalactoside (IPTG) in 1X YT media. C/EBP β was then purified on nickel-chelating columns using imidazole elution (pRSET Xpress, Invitrogen). Protein purity was assessed by western blot analysis and silver staining (Bronstein *et al*, 1992; Burnette, 1981; Gillespie and Hudspeth, 1991; Sandhu *et al*, 1991).

Transient expression analyses
Transient expression analyses in U-937 cells were performed as described (McAllister *et al*, 2000). Cell extracts were harvested after 24 h and dual luciferase assays were performed as described (Promega). Firefly luminescence was normalized to the Renilla luminescence to control for variability in transfection efficiency. Firefly luminescence (pGL3 constructs) was presented with the parental activity set to 1.0 for each experiment and the relative activity of the mutagenized constructs calculated accordingly. Error bars indicate the standard deviation. Each value shown represents the average of three independent experiments performed with duplicate samples.

PCR amplification of brain tissue
For patients 2135, 2297, 2435, 2444, 2597, 2627, cha1, cha2, and cha4, brain tissue was collected at autopsy, frozen, and HIV-1 molecular clones were PCR amplified as described (Corboy and Garl, 1997). For patients 508, 510, 512, 513, 514, 525, 527, and 541, brain-derived molecular clones were PCR amplified as described (Liu *et al*, 2000). The outer primers used for PCR amplification were 5′-CGAAGGAATAGAAGAAGAAGGTG and 5′-CAGTG TGGAAAATCTCTAGC. The inner primers used were 5′-CCTAGAAGAATAAGACAGGGCTT and 5′-GGTCTGAGGGATTGTAGTTAC.

Patient information
In this report, only brains from which at least 3 LTR sequences were cloned per brain were analyzed. A total of 164 brain-derived LTR sequences were generated from 18 patients, while disease information was available for 7 of the patients. Patients 508, 510, and 513 exhibited no signs of clinical dementia while patients 512, 514, 525, and 527 exhibited signs of clinical dementia. Of these 7 patients, only patients 508 and 512 had AIDS. Clinical information was not available for the remaining 11 patients from whom brain-derived LTRs were isolated. The patients used in this analysis resided in Maryland, Minnesota, New York, or Hawaii.

Statistical methods
A one-way analysis of variance was conducted to establish the validity of LTR binding site sequence differences using a Bonferroni correction to adjust for multiple testing, using the SAS statistical program (Sokal and Rohlf, 1995).

Acknowledgements

These studies were supported by Public Health Service grants NS32092 (awarded to BW), MH56832 (awarded to JC) with technical support from Pamela Garl, NS35736 (awarded to SG) with technical support from Dr Yiling Liu, Dr Xiao Peitang, and Jason Scott, and NS26643 (awarded to JM).

References

Ait-Khaled M, McLaughlin JE, Johnson MA, Emery VC (1995). Distinct HIV-1 long terminal repeat quasispecies present in nervous tissues compared to that in lung, blood and lymphoid tissues of an AIDS patient. *Aids* **9:** 675–683.

Akira S, Isshiki H, Sugita T, Tanabe O, Kinoshita S, Nishio Y, Nakajima T, Hirano T, Kishimoto T (1990). A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. *EMBO J* **9:** 1897–1906.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA, Wheeler DL (2000). GenBank. *Nucleic Acids Res* **28:** 15–18.

Bronstein I, Voyta JC, Murphy OJ, Bresnick L, Kricka LJ (1992). Improved chemiluminescent Western blotting procedure. *Biotechniques* **12:** 748–753.

Burnette WN (1981). "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate–polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal Biochem* **112:** 195–203.

Choe H, Farzan M, Sun Y, Sullivan N, Rollins B, Ponath PD, Wu L, Mackay CR, LaRosa G, Newman W, Gerard N, Gerard C, Sodroski J (1996). The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* **85:** 1135–1148.

Corboy JR, Buzy JM, Zink MC, Clements JE (1992). Expression directed from HIV long terminal repeats in the central nervous system of transgenic mice. *Science* **258:** 1804–1808.

Corboy JR, Garl PJ (1997). HIV-1 LTR DNA sequence variation in brain-derived isolates. *J NeuroVirol* **3:** 331–341.

Delwart EL, Sheppard HW, Walker BD, Goudsmit J, Mullins JI (1994). Human immunodeficiency virus type 1 evolution in vivo tracked by DNA heteroduplex mobility assays. *J Virol* **68:** 6672–6683.

Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhart M, Di Marzio P, Marmon S, Sutton RE, Hill CM, Davis CB, Peiper SC, Schall TJ, Littman DR, Landau NR (1996). Identification of a major co-receptor for primary isolates of HIV-1 [see comments]. *Nature* **381:** 661–666.

Dignam JD, Lebovitz RM, Roeder RG (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* **11:** 1475–1489.

Doranz BJ, Rucker J, Yi Y, Smyth RJ, Samson M, Peiper SC, Parmentier M, Collman RG, Doms RW (1996). A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell* **85:** 1149–1158.

Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, Cayanan C, Maddon PJ, Koup RA, Moore JP, Paxton WA (1996). HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC- CKR-5 [see comments]. *Nature* **381:** 667–673.

Estable MC, Bell B, Merzouki A, Montaner JS, O'Shaughnessy MV, Sadowski IJ (1996). Human immunodeficiency virus type 1 long terminal repeat variants from 42 patients representing all stages of infection display a wide range of sequence polymorphism and transcription activity. *J Virol* **70:** 4053–4062.

Feng Y, Broder CC, Kennedy PE, Berger EA (1996). HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor [see comments]. *Science* **272:** 872–877.

Gallo P, Frei K, Rordorf C, Lazdins J, Tavolato B, Fontana A (1989). Human immunodeficiency virus type 1 (HIV-1) infection of the central nervous system: an evaluation of cytokines in cerebrospinal fluid. *J Neuroimmunol* **23:** 109–116.

Geleziunas R, Schipper HM, Wainberg MA (1992). Pathogenesis and therapy of HIV-1 infection of the central nervous system [editorial]. *Aids* **6:** 1411–1126.

Gillespie PG, Hudspeth AJ (1991). Chemiluminescence detection of proteins from single cells. *Proc Natl Acad Sci USA* **88:** 2563–2567.

Henderson AJ, Calame KL (1997). CCAAT/enhancer binding protein (C/EBP) sites are required for HIV-1 replication in primary macrophages but not CD4(+) T cells. *Proc Natl Acad Sci USA* **94:** 8714–8719.

Henderson AJ, Connor RI, Calame KL (1996). C/EBP activators are required for HIV-1 replication and proviral induction in monocytic cell lines. *Immunity* **5:** 91–101.

Henderson AJ, Zou X, Calame KL (1995). C/EBP proteins activate transcription from the human immunodeficiency virus type 1 long terminal repeat in macrophages/monocytes. *J Virol* **69:** 5337–5347.

Ho DD, Bredesen DE, Vinters HV, Daar ES (1989). The acquired immunodeficiency syndrome (AIDS) dementia complex [clinical conference]. *Ann Intern Med* **111:** 400–410.

Kim FM, Kolson DL, Balliet JW, Srinivasan A, Collman RG (1995). V3-independent determinants of macrophage tropism in a primary human immunodeficiency virus type 1 isolate. *J Virol* **69:** 1755–1761.

Kirchhoff F, Greenough TC, Hamacher M, Sullivan JL, Desrosiers RC (1997). Activity of human immunodeficiency virus type 1 promoter/TAR regions and tat1 genes derived from individuals with different rates of disease progression. *Virology* **232:** 319–331.

Korber BT, Kunstman KJ, Patterson BK, Furtado M, McEvilly MM, Levy R, Wolinsky SM (1994). Genetic differences between blood- and brain-derived viral sequences from human immunodeficiency virus type 1-infected patients: evidence of conserved elements in the V3 region of the envelope protein of brain-derived sequences. *J Virol* **68:** 7467–7481.

Lipton SA, Gendelman HE (1995). Seminars in medicine of the Beth Israel Hospital, Boston. Dementia associated with the acquired immunodeficiency syndrome [see comments]. *N Engl J Med* **332:** 934–940.

Liu Y, Tang XP, McArthur JC, Scott J, Gartner S (2000). Analysis of human immunodeficiency virus type 1 gp160 sequences from a patient with HIV dementia: evidence for monocyte trafficking into brain. *J NeuroVirol* **6 Suppl 1:** S70–S81.

Lukashov VV, Kuiken CL, Goudsmit J (1995). Intrahost human immunodeficiency virus type 1 evolution is related to length of the immunocompetent period. *J Virol* **69:** 6911–6916.

Maxam AM, Gilbert W (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol* **65:** 499–560.

McAllister JJ, Phillips D, Millhouse S, Conner J, Hogan T, Ross HL, Wigdahl B (2000). Analysis of the HIV-1 LTR

NF-kappaB-proximal Sp site III: evidence for cell type-specific gene regulation and viral replication. *Virology* **274:** 262–277.

McArthur JC, Hoover DR, Bacellar H, Miller EN, Cohen BA, Becker JT, Graham NM, McArthur JH, Selnes OA, Jacobson LP (1993). Dementia in AIDS patients: incidence and risk factors. Multicenter AIDS Cohort Study. *Neurology* **43:** 2245–2252.

Michael NL, D'Arcy L, Ehrenberg PK, Redfield RR (1994). Naturally occurring genotypes of the human immunodeficiency virus type 1 long terminal repeat display a wide range of basal and Tat-induced transcriptional activities. *J Virol* **68:** 3163–3174.

Pang S, Shlesinger Y, Daar ES, Moudgil T, Ho DD, Chen IS (1992). Rapid generation of sequence variation during primary HIV-1 infection [published erratum appears in AIDS 1992 Jun;6(6):following 606]. *Aids* **6:** 453–460.

Perelson AS, Essunger P, Cao Y, Vesanen M, Hurley A, Saksela K, Markowitz M, Ho DD (1997). Decay characteristics of HIV-1-infected compartments during combination therapy [see comments]. *Nature* **387:** 188–191.

Perrella O, Carrieri PB, Guarnaccia D, Soscia M (1992). Cerebrospinal fluid cytokines in AIDS dementia complex. *J Neurol* **239:** 387–388.

Sandhu GS, Eckloff BW, Kline BC (1991). Chemiluminescent substrates increase sensitivity of antigen detection in Western blots. *Biotechniques* **11:** 14–16.

Sokal RR, Rohlf FJ (1995). *Biometry* (3rd ed). New York: W. H. Freeman.

Tesmer VM, Rajadhyaksha A, Babin J, Bina M (1993). NF-IL6-mediated transcriptional activation of the long terminal repeat of the human immunodeficiency virus type 1. *Proc Natl Acad Sci USA* **90:** 7298–7302.