*Proc. Natl. Acad. Sci. USA*
Vol. 89, pp. 10247–10251, November 1992
Medical Sciences

# Relationship of human immunodeficiency virus type 1 sequence heterogeneity to stage of disease

(macrophage/tropism/quasi-species/AIDS)

TERRY MCNEARNEY*, ZUZANA HORNICKOVA*, RICHARD MARKHAM†, ANAHID BIRDWELL*, MAX ARENS‡, ALFRED SAAH†, AND LEE RATNER*

Departments of *Medicine and Molecular Microbiology, and ‡Pediatrics, Washington University School of Medicine, St. Louis, MO 63110; and †Department of Immunology and Infectious Diseases, School of Hygiene and Public Health, The Johns Hopkins University, Baltimore, MD 21205

*Communicated by Donald C. Shreffler, July 8, 1992*

ABSTRACT **V3 envelope sequences were determined from amplified human immunodeficiency virus type 1 (HIV-1) sequences of uncultivated leukocytes obtained sequentially from four infected adults over the course of infection. Lower levels of sequence heterogeneity were noted in samples obtained early in HIV-1 infection, prior to CD4 depletion, than in samples obtained at later times during disease. The pattern of amino acid sequence divergence included nonrandom changes, with evidence of sequence variants arising from HIV-1 quasi-species present earlier in infection. Consensus sequences for isolates obtained early after infection from different patients demonstrated a high level of conservation with one another and a consensus sequence for macrophage-tropic HIV-1 isolates. These findings suggest that a highly restricted subset of HIV-1 quasi-species can be transmitted and can establish infection.**

Sequence heterogeneity is a well-recognized characteristic of all RNA viruses, particularly human immunodeficiency virus type 1 (HIV-1), a member of the lentivirus family (1, 2). A wide range of sequence heterogeneity has been noted among HIV-1 isolates obtained from unrelated infected individuals (6–22% of nucleotides in *env*) (3–5), close contacts (0–9% of nucleotides in *env*) (6–8), or the same individual (0–2% of nucleotides in *env*) (7, 9, 10). The high level of HIV-1 diversity has been ascribed to an error-prone reverse transcriptase (11, 12), recombination during virus replication (13, 14), and selective pressures exerted by the host immune system. Sequence heterogeneity may alter cell-specific tropism (15–17), replication kinetics (18), cytopathic activity (19), and responses to neutralizing antibodies (20, 21) and cytotoxic T lymphocytes (22).

The *env* gene of HIV-1 manifests the greatest level of heterogeneity of any element in the viral genome (23). Distinct regions within *env* vary by as much as 75% of nucleotides among viral isolates (5). These domains encode portions of the envelope protein that are designated variable regions V1–V5 (24).

The third variable domain, or V3, comprises a loop structure in a $\beta$-turn configuration (25), which includes 34–37 amino acids bordered by a disulfide-linked cysteine bridge (26). Although residues at the tip of the loop are highly conserved, sequences on either side are highly variable (5, 25). The V3 loop includes a single potential N-linked glycosylation site and is flanked on either side by several additional potential sites for oligosaccharide modification of asparagine residues.

The V3 loop sequence is a major epitope for cytotoxic T-lymphocyte responses (27), and it is also the primary neutralizing domain for antibodies capable of blocking HIV-1 infection (28). Although the V3 loop is not directly involved in binding envelope protein to the first immunoglobulin-like loop in CD4 (29), it is involved in subsequent events of virus entry mediated by fusion of viral and cellular membranes (30, 31). In addition, sequences immediately C terminal to the V3 loop may provide structural support for the CD4 binding pocket in the envelope protein (29). The V3 loop has also been demonstrated to be critical in modulating HIV-1 tropism for macrophages, T-lymphoid cell lines, and brain-derived fibroblasts (15–17, 32). These activities may be mediated by V3 loop cleavage by a serine protease(s) (33) or by interactions with a second cell-surface molecule or a second binding domain in CD4 (34).

In the current study, we analyzed V3 loop sequences from naturally occurring HIV-1 isolates to determine whether sequence heterogeneity is related to the stage of HIV-1 infection. One interesting aspect of this study is the use of clinical samples from four individuals, including samples from early through late stages of infection, as measured by $CD4^+$ lymphocyte counts. In addition, these experiments utilized uncultivated mononuclear cells for this analysis to avoid bias introduced by *in vitro* propagation of HIV-1 (35).

## MATERIALS AND METHODS

**Subject Selection.** High-risk individuals were evaluated every 6 months over a 2.5- to 4.5-year period with clinical evaluations and CD4 determinations. Ficoll-purified peripheral blood leukocytes were stored from each clinic visit. Subjects were selected based on the availability of leukocyte samples through the full course of infection as reflected by CD4 levels. The four subjects are designated S1–S4. None received 3′-azido-3′-deoxythymidine (AZT) therapy during the study.

**Sequence Analysis.** Peripheral blood leukocytes were resuspended in lysis buffer (1% SDS/1% 2-mercaptoethanol/10 mM Tris·HCl, pH 8.0/150 mM NaCl/5 mM EDTA) and treated with proteinase K (300 $\mu$g/ml) (Sigma) overnight at 37°C. Negative controls were included in each group of amplification reaction mixtures. DNA was extracted with phenol/chloroform, precipitated with ethanol, and resuspended in water. Nucleotide sequences 6587–6920 (numbering according to ref. 7) were amplified using a nested primer protocol (36) with a sensitivity of detection of 10 copies of DNA and an error rate of <1/2000 nucleotides (37). Amplified sequences were cloned into pUC19. For each sample, 2–10 recombinant plasmids were sequenced with Sequenase version 2.0 (United States Biochemical) on both DNA strands according to the manufacturer's recommendations (38). Sequences were validated by independent analysis of separate PCRs performed with the same primers and an

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: HIV-1, human immunodeficiency virus type 1.

alternative primer pair combination. Nucleotide sequences are available upon request.§

## RESULTS

Sequential samples of primary mononuclear cells were obtained from four subjects over a 2.5- to 4.5-year period of HIV-1 infection. For two subjects (S2 and S4), the initial sample (sample A) was obtained when the CD4 count was >900 per mm³ (1225 and 943, respectively), a second sample (sample B) when the CD4 count was 500–800 per mm³ (756 and 575, respectively), and a third sample (sample C) when the CD4 count was 100–400 per mm³ (368 and 187, respectively). Sample A from S2 was obtained within 3 months of an influenza-like illness characteristic of acute HIV-1 infection. For S3, samples were obtained at only the latter two time points (CD4 counts, 537 and 142). Subject S1 had fluctuating numbers of CD4⁺ lymphocytes, and sequential samples from this individual are designated D, E, F, and G (470, 826, 273, and 515 CD4 lymphocytes per mm³, respectively). The A and D time point samples were both obtained at the first 6 monthly visit at which the subject seroconverted.

For S2 and S4, the early stage sample (sample A) showed a restricted population of nucleotide sequences compared to those obtained from late disease stage samples (samples B and C from S4 and sample C from S2; Fig. 1). For S4, the mean level of HIV-1 nucleotide sequence divergence increased steadily from 1.1% to 2.3%, while, for S2, mean divergence increased from 0.2% to 0.9% over the course of infection. For S3, the mean divergences were 1.2% and 1.9% for the two available samples. For S1, sequence divergence increased from 0.2% in sample D to 3.2% in sample E and to 2.5% in sample F, but in sample G it decreased to 0.2%.

The predicted amino acid sequences manifested nonrandom changes over time in the sequential samples (Fig. 2). Sequence differences compared to the first clone were generally present in more than one clone obtained at a given time point. Furthermore, some clones obtained from the later time points also exhibit the same sequence alterations identified at an earlier time point. Sequence alterations were also noted to be significantly more frequent downstream than upstream of the V3 loop, particularly in late time point samples (Fig. 2 legend).

Amino acid sequence comparisons demonstrated changes at three of nine, one of eight, three of eight, and three of nine potential N-linked glycosylation sites in samples from S1–S4, respectively (Figs. 2 and 3). Potential N-glycosylation sites at residues 229 and 243 were conserved among all sequences. Other sites were variably conserved, and in several cases distinct differences were noted between sequences obtained from early and late stage samples. For example, a potential N-glycosylation site was present at residue 256 in 100% of sequences of the early disease sample and 21% of sequences of late disease stage samples and at residue 305 in 0% of early stage sequences and 86% of late stage sequences from S1. A potential N-glycosylation site at residue 262 was noted in 0% of early and in 88% of late stage sequences from S4. Also of note is an insertion in the majority of late time point clones of S1, which shifts by one residue the predicted N-glycosylation site at position 321.

Residues 329, 330, 334, and 336 are critical for CD4 binding, although additional sequences required for receptor binding are found on both sides of this region (29, 41). Alterations are found at residue 329 in clone S2.A.6 and at residues 329 and 330 in most S4 clones. Cysteine residues flanking the V3 loop were completely conserved, although an additional cysteine is present in clone S3.C.27. The GPGR sequence at the center of the V3 domain was conserved in samples from three of the subjects, but changes are noted in this sequence in clones obtained from S2.

Predicted amino acid sequences obtained at different times from the same patient demonstrated up to 9.8% divergence in three cases (S2–S4) and up to 19.5% divergence for S1. Higher levels of sequence heterogeneity were seen in comparing consensus amino acid sequences from different patients, with divergences of 13.4–25.9% among these and other representative North American or European patients (JFL, WMJ2, and HXB2) and 37.5–43.8% for a representative African isolate (ELI) (5).

Fig. 2 portrays the pattern of amino acid divergence of sequences obtained from each patient. Sequences with the greatest similarity from clones obtained at different time points are indicated by boxes. A clear pattern of evolution of predominant amino acid sequence alterations among clones at different time points is noted. However, in at least one case, S1.F.7, a sequence is identified at a late point that more closely resembles those obtained at an earlier time point, S1.D.31, -32, -33, and -37, than sequences for the same time point.

V3 loop amino acid sequences obtained from the earliest time point samples also demonstrated a high level of sequence similarity to one another and to a consensus sequence generated for North American and European isolates (Fig. 4; ref. 25). Conservation with a recently reported consensus sequence for macrophage-tropic clones was also noted (40, 42). In contrast, more divergent V3 domain sequences were noted for late time point samples (Fig. 4) and sequences obtained from isolates tropic for T-lymphoid cell lines rather than macrophages (data not shown; ref. 42).

§The sequences reported in this paper have been deposited in the GenBank data base (accession nos. L03430–L03515).

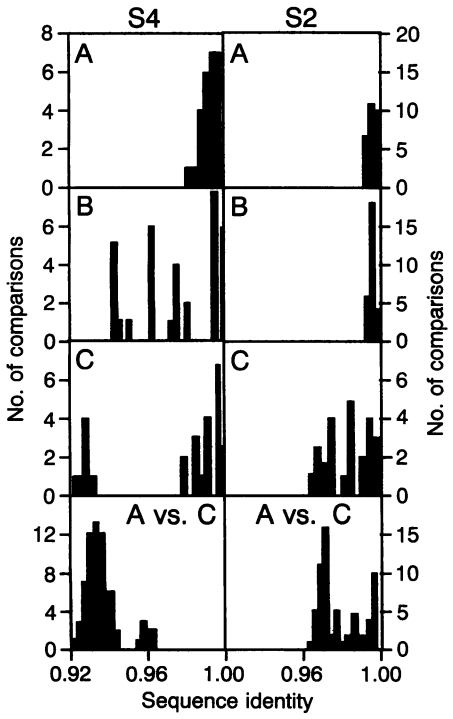



FIG. 1. Sequence heterogeneity during progressive HIV-1 infection. Sequence comparisons are shown for clones obtained from two subjects for each time point, A, B, and C (intraperiod comparisons), and a comparison is shown between clones obtained at the A and C time points (interperiod comparisons). All possible 2-fold comparisons among a group of distinct sequences were performed, and the number of comparisons for each determined level of sequence identity (ranging from 92% to 100% identity) is shown (39). Increasing nucleotide sequence divergence is represented by migration of bars to the left. Mean nucleotide divergences (see *Results*) were calculated as the average of all possible comparisons of two sequences.

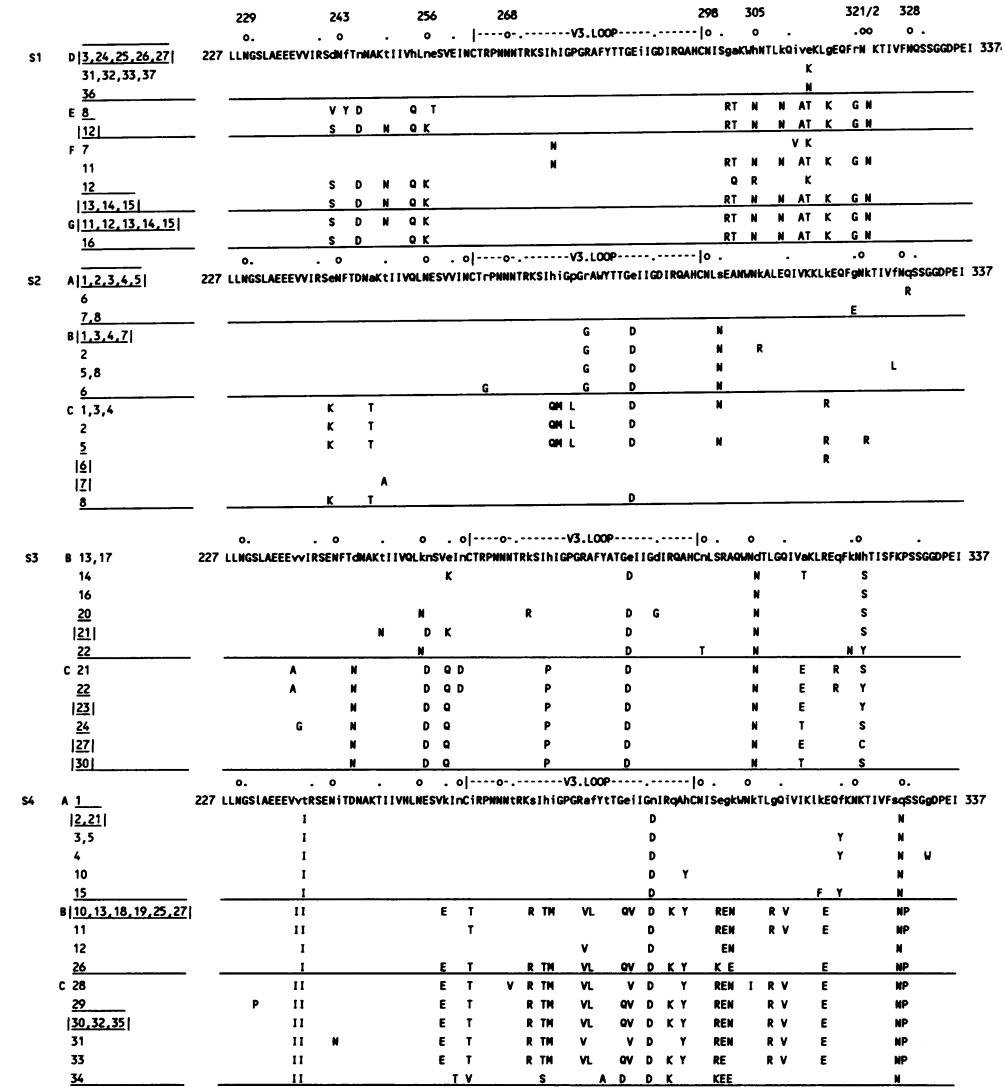


FIG. 2. Amino acid sequence alignments of HIV-1 sequences derived from infected leukocytes during progressive infection in four subjects. Amino acid sequence from a clone obtained during the first time point of each patient is shown, and differences in other clones are indicated. Positions at which differences are noted in other sequences are listed in lowercase letters. Clone numbers for each patient and time point are indicated on the left, and those for which identical sequences were determined are grouped together. Clone numbers in boxes indicate those that are most similar to clones at subsequent or preceding time points. Above each group of sequences the position of every 10th residue is shown by a dot numbered according to Westervelt *et al.* (40), potential N-glycosylation sites are shown by open circles, with the position number of each indicated above, and the position of the V3 domain is indicated. Sequence alterations were noted upstream and downstream of the V3 loop at 19.4% and 25.6% of positions, respectively, for S1.D–F, at 8.3% and 13.5% of positions, respectively, for S2.B–C, and at 16.6% and 23.0% of positions, respectively, for S4.B–C. Date and CD4 count per $mm^3$ (in parentheses) for each sample are as follows: S1.D, Nov. 1985 (470); S1.E, July 1987 (826); S1.F, Jan. 1988 (273); S1.G, May 1989 (515); S2.A, May 1985 (1225); S2.B, Apr. 1987 (756); S2.C, Oct. 1987 (368); S3.B, June 1987 (537); S3.C, Dec. 1987 (142); S4.A, Jan. 1985 (943); S4.B, Jan. 1989 (575); S4.C, June 1989 (187).

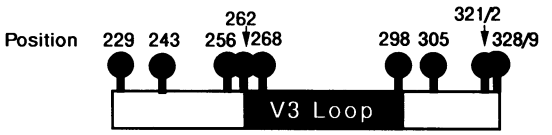
## DISCUSSION

The current findings demonstrate that sequence heterogeneity is related to disease stage, as measured by CD4 lymphocyte count. Samples obtained at the earliest stage of HIV-1 infection demonstrated limited nucleotide and predicted V3 loop amino acid sequence diversity. At later stages of disease, 4- to 14-fold more sequence diversity was noted than at earlier stages of infection. This is analogous to a recent report of restricted HIV-1 sequence heterogeneity in infected children compared to their infected mothers (39).

One exception to the correlation of sequence heterogeneity and disease progression is noted for S1. The highest levels of sequence diversity were noted in samples obtained at the second and third time points (E and F) compared to the initial time point (D), but heterogeneity was restricted in sequences at the fourth time point (G). Similar findings were also noted by Nowak *et al.* (43) in one of two patients. Predominance of a single virus type late in infection may be due to its ability to avoid an attenuated immune response.

Other sequential HIV-1 isolates have been examined from infected children, hemophiliacs, and homosexual adults (7, 9, 10, 35, 44). In each case, levels of sequence variation of 0–5% were noted in V3 loop amino acid sequences. However, these studies did not include samples from the earliest time points of infection.

Neither termination codons nor frameshifts were identified in these sequences (Fig. 2). Although deletions and insertions are commonly found in HIV-1 *env* sequences (5), only a single in-frame insertion was found in clones from S1. Studies of *tat* (35) and other areas of *env* (7) identified a high proportion of defective proviral genomes, which may contribute to HIV-1 pathogenesis. However, in several studies of V3 sequences, no inactivating mutations were detected among




Frequency (%) of Potential N-Glycosylation Sites at Each Position Among HIV-1 Clones

| Subject | Time Point | 229 | 243 | 256 | 262 | 268 | 298 | 305 | 321/2 | 328/9 |
|---|---|---|---|---|---|---|---|---|---|---|
| S1 | Early | 100 | 100 | 100 | 100 | 100 | 100 | 0 | 100* | 100 |
| | Late | 100 | 100 | 21 | 100 | 100 | 100 | 86 | 100* | 100 |
| S2 | Early | 100 | 100 | 100 | 100 | 100 | 100 | 0 | 100 | 100 |
| | Late | 100 | 100 | 100 | 100 | 100 | 25 | 0 | 100 | 100 |
| S3 | Late | 100 | 100 | 15 | 85 | 100 | 92 | 100 | 100 | 0 |
| S4 | Early | 100 | 100 | 100 | 0 | 100 | 100 | 100 | 100 | 87* |
| | Late | 100 | 100 | 100 | 88 | 94 | 100 | 100 | 100 | 12* |

FIG. 3. Substitutions in potential N-glycosylation sites. (*Upper*) Schematic of the portion of the envelope examined, location of the V3 loop, and positions of the nine potential N-glycosylation sites. (*Lower*) Frequency of potential N-glycosylation sites at each position is shown for early (A or D) and late (B, C, E, F, and G) time point clones for each of the four subjects. For S1, a shift in glycosylation sites from 321 to 322 is noted for late time point clones, and, for S2, a potential glycosylation site is at position 329 rather than 328 (indicated by *).

amplified products (8, 45, 46). These findings, together with the observation from the current study that 86% of nucleotide changes result in amino acid changes, suggest that envelope variants arose as a result of phenotypic selection.

N-glycosylation may modify HIV-1 infectivity or recognition by neutralizing antibodies (47–49). Potential N-glycosylation sites at residues 229 and 243 were completely conserved among all sequences in the current study, suggesting a potentially important role for these modifications. In other clones, differences in the frequency of specific potential

```
                          * * V * *
Consensus:   CtRPNNNTRKSIHIGPGRAfYtTGEIIGDIRQAHC

S1.D
S2.A                          W
S4.A         I
MTcon                           A

Consensus:   CIRPNNNTRKsIhiGpGrafYtTGeiIGDIRqAhC

S1.F                  N
S1.G
S2.B                        G     D
S2.C                  QM L        D
S3.B                              A D
S3.C                  P           A D
S4.B                R TM    VL    QV      K Y
S4.C                R TM    VL    QV      K Y
```

FIG. 4. Conserved V3 domain sequence early in HIV-1 infection. Consensus amino acid sequence derived for 245 North American and European isolates is shown. (*Upper*) Differences from consensus V3 loop amino acid sequence among sequences obtained from the earliest time points of three of the patients in the study and from the consensus sequence for 21 macrophage tropic isolates. (*Lower*) Differences from consensus V3 loop amino acid sequence of those obtained from the later time points of all four subjects. Sequences shown for different time points of subjects in this analysis represent those present in ≥50% of clones (100% of S1.D and S2.A, 71% of S4.A, 67% of S1.F, 83% of S1.G, 86% of S2.B, 62% of S2.C, 57% of S3.B, 100% of S3.C, 78% of S4.B, and 50% of S4.C clones). MTcon, consensus sequence for macrophage-tropic isolates (40, 42). *, Residues potentially important for HIV-1 tropism (40); V, potential cleavage site (33).

N-glycosylation sites were noted between early and late stages of infection, which may contribute to biological and immunological differences among such isolates.

The pattern of evolution of V3 loop sequences identified in the current study suggests that sequence heterogeneity is not completely random. The finding that certain variants may be progenitors of viruses found later in infection has been confirmed in studies of simian immunodeficiency virus-infected macaques (50) and HIV-1-infected children (46). However, not all mutations were preserved in the later time points, suggesting that mutations may be progressive, but not necessarily cumulative. These findings are consistent with the proposed selective influences of the immune system that determine which escape mutants will predominate.

Increased viral diversity may be a by-product of disease progression or it may be critical for disease progression. Nowak *et al.* (43) proposed that an antigenic threshold exists above which the number of viral variants exceeds the immune regulatory effectors. The threshold level may depend on the replication rates of the quasi-species and the strength of the immune response.

The functional role of sequence changes described in this study remains to be analyzed. Isolates obtained from asymptomatic infected individuals generally replicate slowly and to low titers *in vitro* and are unable to infect T-lymphoid cell lines or induce syncytia, whereas virus isolates from more advanced disease stages are characterized by high titers and rates of replication, T-lymphoid cell line infection, and syncytia induction (19, 51, 52). These studies have utilized cultured isolates, and, thus, it will be instructive to assess the activity of more representative HIV-1 clones, including those obtained early after infection and without the bias of *in vitro* cultivation. Of particular interest in this regard are amino acid alterations occurring at positions that may affect V3 loop cleavage (residue 280 changes in all S2.B clones and residue 281 changes in most S4.B and C clones), cell tropism (residue 278 alterations in most S2.C clones and residue 287 alterations in S4.B and S4.C clones), or CD4 binding activity (residue 329 changes in S2.A.6 and residues 329 and 330 change in S4 clones) (16, 32, 33, 40, 42). Furthermore, frequent sequence alterations, including a shift of an N-glycosylation site C terminal to the V3 loop may affect the conformation of the CD4 binding pocket (29).

In the current study, sequences were obtained exclusively from peripheral blood leukocytes. It is unclear whether these are representative of HIV-1 isolates present in lymphocytes or monocytes at each stage of disease. Although the predominant infected cell in blood at late stages of disease is the lymphocyte (53), comparable studies have not been performed during early stages of infection. In addition, it is likely that tissue reservoirs for HIV-1 may account for a significant portion of the virus load in an infected individual (54). Distinct viral variants may evolve in different tissue compartments (37, 45, 55). Studies of HIV-1 heterogeneity in lymph nodes or other tissue sources at different stages of infection may be particularly useful in defining the role of sequence heterogeneity in disease progression and disease manifestations.

In the current study, only sequences from the V3 region were examined. It is unclear whether a limited range of sequences in other portions of the genome occur in the early stages of infection. Nevertheless, this is a major epitope for both T- and B-lymphocyte responses to HIV-1, and, thus, V3 loop sequences are likely to be an important determinant of viral pathogenicity.

One striking and unexpected observation was that consensus V3 sequences from early infection samples of different individuals are remarkably similar. A similar consensus sequence is also present in an early isolate from an infected child (44) and an infected hemophiliac (56). Moreover, this

consensus sequence matches that determined to be necessary for HIV-1 infection of macrophages (40, 42). Thus, macrophage infection may be critical for virus transmission or establishment of infection. It is possible that HIV-1 may be transmitted primarily in macrophages, or that macrophage-tropic isolates may be most efficient in infection or replication at portals of infection within an immunologically intact host. Alternatively, macrophage-tropic isolates may be less cytopathic and, thus, a greater proportion of such isolates will be found in viable leukocyte samples. Perhaps the capacity to infect macrophages confers a survival advantage by allowing these isolates to evade the vigorous host immune response generated in early HIV-1 infection. However, it is also possible that macrophage infection is not directly involved in establishment of infection *in vivo* but that it merely serves as an indicator of another property critical for HIV-1 infection of the host.

These findings also have important implications for vaccine development. A vaccine targeted at the consensus V3 loop demonstrated in early infection may be particularly useful for protection of uninfected individuals.

This work was supported by grants from the National Institutes of Health and U.S. Army Medical Research Acquisition Activity. L.R. is an American Cancer Society Research Professor.

1. Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. & Van de Pol, S. (1982) *Science* **215**, 1577–1585.
2. Hu, W. S. & Temin, H. M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9253–9257.
3. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. J., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., Pearson, M. L., Lautenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C. & Wong-Staal, F. (1985) *Nature (London)* **313**, 277–284.
4. Alizon, M., Wain-Hobson, S., Montagnier, L. & Sonigo, P. (1986) *Cell* **46**, 63–74.
5. Myers, G., Korber, B., Berzofsky, J. & Smith, R. (1991) *Human Retroviruses and AIDS* (Los Alamos National Lab., Los Alamos, NM).
6. Burger, H., Weiser, B., Flaherty, K., Gulla, J., Nguyen, P.-N. & Gibbs, R. A. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 11236–11240.
7. McNearney, T., Westervelt, P., Thielan, B. J., Trowbridge, D. B., Garcia, J., Whittier, R. & Ratner, L. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1917–1921.
8. Balfe, P., Simmonds, P., Ludlam, C., Bishop, J. & Brown, A. J. L. (1990) *J. Virol.* **64**, 6221–6233.
9. Hahn, B. H., Shaw, G. M., Taylor, M. E., Redfield, R. R., Markham, P. D., Salahuddin, S. Z., Wong-Staal, F., Gallo, R. C., Parks, E. S. & Parks, W. P. (1986) *Science* **232**, 2548–2553.
10. Saag, M. S., Hahn, B. H., Gibbons, J., Li, Y., Parks, E. S., Parks, W. P. & Shaw, G. M. (1988) *Nature (London)* **334**, 440–444.
11. Preston, B. D., Poiesz, B. J. & Loeb, L. A. (1988) *Science* **242**, 1168–1171.
12. Roberts, J. D., Bebenek, K. & Kunkel, T. A. (1988) *Science* **242**, 1171–1173.
13. Coffin, J. M. (1979) *J. Gen. Virol.* **42**, 1–26.
14. Clavel, F., Hoggan, M. D., Willey, L. R., Strebel, K., Martin, M. A. & Repaske, R. (1989) *J. Virol.* **63**, 1455–1459.
15. Shioda, T., Levy, J. A. & Cheng-Mayer, C. (1991) *Nature (London)* **349**, 167–169.
16. Westervelt, P., Gendelman, H. E. & Ratner, L. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 3097–3101.
17. O'Brien, W. A., Koyanagi, Y., Namazie, A., Zhao, J.-Q., Diagne, A., Idler, K., Zack, J. A. & Chen, I. S. Y. (1990) *Nature (London)* **348**, 69–73.
18. Fenyo, E. M., Morfeldt-Manson, L., Chiodi, F., Lind, B., von Gegerfelt, A., Albert, J., Olausson, E. & Asjo, B. (1988) *J. Virol.* **62**, 4414–4419.
19. Tersmette, M., Gruters, R. A., de Wolf, F., de Goede, R. E. Y., Lange, J. M. A., Schellekens, P. T. A., Goudsmit, J., Huisman, H. G. & Miedema, F. (1989) *J. Virol.* **63**, 2118–2125.
20. Albert, J., Abrahamsson, B., Nagy, K., Aurelius, E., Gaines, H., Nystrom, G. & Fenyo, E. M. (1990) *AIDS* **4**, 107–112.
21. Nara, P. L., Smit, L., Dunlop, N., Hatch, W., Merges, M., Waters, D., Kelliner, J., Gallo, R. C., Fischinger, P. J. & Goudsmit, J. (1990) *J. Virol.* **64**, 3779–3791.
22. Phillips, R. E., Rowland-Jones, S., Nixon, D. F., Gotch, F. M., Edwards, J. P., Ogunlesi, A. O., Elvin, J. G., Rothbard, J. A., Bangham, C. R. M., Rizza, C. R. & McMichael, A. J. (1991) *Nature (London)* **354**, 453–459.
23. Starcich, B. R., Hahn, B. H., Shaw, G. M., McNeely, P. D., Modrow, S., Wolf, H., Parks, W. P., Josephs, S. F., Gallo, R. C. & Wong-Staal. (1986) *Cell* **45**, 637–648.
24. Modrow, S., Hahn, B. H., Shaw, G. M., Gallo, R. C., Wong-Staal, F. & Wolf, H. (1987) *J. Virol.* **61**, 570–578.
25. LaRosa, G. J., Davide, J. P., Weinhold, K., Waterbury, J. A., Profy, A. T., Lewis, J. A., Langlois, A. J., Dreesman, G. R., Boswell, B. N., Shadduck, P., Holley, L. H., Karplus, M., Bolognesi, D. P., Matthews, T. J., Emini, E. A. & Putney, S. D. (1990) *Science* **249**, 932–935.
26. Leonard, C. K., Spellman, M. W., Riddle, L., Harris, R. J., Thomas, J. N. & Gregory, T. J. (1990) *J. Biol. Chem.* **265**, 10373–10382.
27. Takahashi, H., Merli, S., Putney, S. D., Houghten, R., Moss, B., Germain, R. N. & Berzofsky, J. A. (1989) *Science* **246**, 118–121.
28. Jahaverian, K., Langlois, A., McDanal, C., Ross, K., Eckler, L., Jellis, C., Profy, A., Rusche, J., Bolognesi, D., Putney, S. & Matthews, M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6768–6772.
29. Olshevsky, U., Helseth, E., Furman, C., Li, J., Haseltine, W. & Sodroski, J. (1990) *J. Virol.* **64**, 5701–5707.
30. Freed, E. O., Myers, D. J. & Risser, R. (1991) *J. Virol.* **65**, 190–194.
31. Page, K. A., Stearns, S. M. & Littman, D. R. (1990) *J. Virol.* **64**, 5270–5276.
32. Takeuchi, Y., Akutsu, M., Murayama, K., Shimizu, N. & Hoshino, H. (1991) *J. Virol.* **65**, 1710–1718.
33. Clements, G. J., Price-Jones, M. J., Stephens, P. E., Sutton, C., Schulz, T. F., Clapham, P. R., McKeating, J. A., McClure, M. O., Thomson, S., Marsh, M., Kay, J. & Weiss, R. A. (1991) *AIDS Res. Hum. Retrovirus* **7**, 3–16.
34. Camerini, D. & Seed, B. (1990) *Cell* **60**, 747–757.
35. Meyerhans, A., Cheynier, A., Albert, J., Seth, M., Kwok, S., Sninsky, J., Morfeldt-Manson, L., Asjo, B. & Wain-Hobson, S. (1989) *Cell* **58**, 901–910.
36. Albert, J. & Fenyo, E. M. (1990) *J. Clin. Microbiol.* **28**, 1560–1564.
37. Pang, S., Vinters, H. V., Akashi, T., O'Brien, W. A. & Chen, I. S. Y. (1991) *J. AIDS* **4**, 1082–1092.
38. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
39. Wolinsky, S. M., Wike, C. M., Korber, B. T. M., Hutto, C., Parks, W. P., Rosenblum, L. L., Kuntsman, K. J., Furtado, M. R. & Munoz, J.-L. (1992) *Science* **255**, 1134–1137.
40. Westervelt, P., Trowbridge, D. B., Epstein, L., Blumberg, B., Li, Y., Hahn, B. H., Shaw, G. M., Price, R. & Ratner, L. (1992) *J. Virol.* **66**, 2577–2582.
41. Fung, M. S. C., Sun, C. R. Y., Gordon, W. L., Liou, R.-S., Chang, T. W., Sun, W. C., Daar, E. S. & Ho, D. D. (1992) *J. Virol.* **66**, 848–856.
42. Chesbro, B., Wehrly, K., Nishio, J. & Perryman, S. (1992) *J. Virol.*, in press.
43. Nowak, M. A., Anderson, R. M., McLean, A. R., Wolfs, T. F. W., Goudsmit, J. & May, R. M. (1991) *Science* **254**, 963–969.
44. Wolfs, T., deJong, J., van den Berg, H., Tunagel, J., Krone, W. & Goudsmit, J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9938–9942.
45. Brichmann, J. E., Albert, J. & Vartdal, F. (1991) *J. Virol.* **65**, 2019–2023.
46. Steuler, H., Storch-Hagenlocher, B. & Wildemann, B. (1992) *AIDS Res. Hum. Retrovirus* **8**, 53–59.
47. Ratner, L. (1992) *AIDS Res. Hum. Retrovirus* **8**, 165–173.
48. Dedera, D., Gu, R. & Ratner, L. (1992) *Virology* **187**, 377–382.
49. Lee, W.-R., Syu, W.-J., Du, B., Matsuda, M., Tan, S., Wolf, A., Essex, M. & Lee, T.-H. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 2213–2217.
50. Overbaugh, J., Rudensey, L. M., Pappenahusen, M. D., Beneviste, R. E. & Morton, W. R. (1991) *J. Virol.* **65**, 7025–7031.
51. Asjo, B., Morfeld-Manson, L., Albert, J., Biberfeld, G., Karlsson, A. & Fenyo, E. M. (1986) *Lancet* **ii**, 660–662.
52. Cheng-Mayer, C., Seto, D., Tateno, M. & Levy, J. A. (1988) *Science* **240**, 80–82.
53. Schnittman, S. M., Psallidopoulos, M. C., Lane, H. C., Thompson, L., Baseler, M., Massari, F., Fox, C. H., Salzman, N. P. & Fauci, A. S. (1989) *Science* **245**, 305–308.
54. Pantaleo, G., Graziosi, G., Butini, L., Pizzo, P. A., Schnittman, S. M., Kotler, D. P. & Fauci, A. S. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9838–9842.
55. Epstein, L. G., Kuiken, C., Blumberg, B. M., Hartman, S., Sharer, L. R., Clement, M. & Goutsmit, J. (1991) *Virology* **180**, 583–590.
56. Simmonds, P., Zhang, L. Q., McOmish, F., Balfe, P., Ludlam, C. A. & Leigh Brown, A. J. (1991) *J. Virol.* **65**, 6266–6276.