JOURNAL OF VIROLOGY, Apr. 1995, p. 2708–2715 Vol. 69, No. 4
0022-538X/95/$04.00+0

Sequence Diversity of V1 and V2 Domains of gp120 from Human Immunodeficiency Virus Type 1: Lack of Correlation with Viral Phenotype

NING WANG, TUOFU ZHU, AND DAVID D. HO*

The Aaron Diamond AIDS Research Center, New York University School of Medicine, New York, New York 10016

Received 20 October 1994/Accepted 10 January 1995

We analyzed by PCR and direct sequencing 57 viral sequences from 47 individuals infected with human immunodeficiency virus type 1, focussing on the V1 and V2 regions of gp120. There was extensive length polymorphism in the V1 region, which rendered sequence alignment difficult. The V2 hypervariable locus also displayed considerable length variations, whereas flanking regions were relatively conserved. Two-thirds of the amino acid residues in these flanking regions were highly conserved (>80%), presumably reflecting their critical contribution to V2 structure or function. We also characterized the syncytium-inducing properties of the isolates from which we derived sequence information. There was no correlation between V1 or V2 sequences and the viral phenotype, contrary to a previous report (M. Groenink, R. A. M. Fouchier, S. Broersen, C. H. Baker, M. Koot, A. B. van't Wout, H. G. Huisman, F. Miedema, M. Tersmette, and H. Schuitemaker, Science 260:1513–1516, 1993). The sequence heterogeneity described in this study provides information to suggest that it would be most difficult to exploit the V1 and V2 domains for vaccine development.

A high degree of genetic variability in human immunodeficiency virus type 1 (HIV-1) has been shown for independent viral isolates (1, 42), for sequential isolates from the same patient (15, 18, 41), and within a single patient isolate (30, 39). The *env* gene displays a particularly high level of sequence variation (14), which led to the description of five variable regions (V1 to V5) and five conserved regions (C1 to C5) in the surface glycoprotein gp120 (32, 49). The viral envelope genotype has been shown to influence the biological phenotype with respect to replication rate in different cells, cytotropism, and the capacity to induce syncytia in vitro (5, 13, 28, 35, 45, 48). Viral phenotype may also be associated with HIV pathogenesis in vivo. Viruses isolated during primary infection, or soon thereafter, are generally found to possess macrophage tropism and the non-syncytium-inducing (NSI) phenotype (38, 52). In contrast, viruses with rapid-high replication rates, T-cell line tropism, and syncytium-inducing (SI) properties are associated with later stages of disease (7, 44, 46). Longitudinal studies have shown that the evolution of viral phenotype, notably increased viral replication, NSI-to-SI conversion, and macrophage to T-cell line tropism change, are associated with disease progression (5, 6, 7, 22, 44–46).

HIV-1 isolated from infected individuals contains different sequence variants, the genetic mixture being commonly referred to as a population or swarm of quasispecies (30). The dominant genotype is called the major variant. Cocultivation with normal donor peripheral blood mononuclear cells (PBMC) is not deemed a satisfactory approach to recover the complete spectrum of HIV-1 variants, since each variant is not amplified equally (23, 24, 30). Hence, in vitro cultivation influences the distribution of quasispecies, because rapidly replicating variants might be selected and amplified despite only a short-term passage. It has been shown that both NSI and SI clones can be detected in patients with SI isolates, but only NSI clones are detected in patients from whom no SI variants are recovered (6, 13, 22). In these latter patients, rapidly replicating NSI variants are selected on culturing (6). Thus, the consensus sequence still represents NSI variants. However, SI variants generally replicate more rapidly and may be selectively amplified during cocultivation, but the consensus, unselected sequence of SI viral isolates may still be representative of the NSI phenotype.

Phenotypic differences between viruses may be evident in the viral genome. As a result of comprehensive viral sequence and functional analyses, substitutions at specific amino acid residues in the V3 region have been associated with viral phenotype and disease progression (8, 26, 31). Recent reports suggest that sequence variation in the V1 and V2 regions may also influence cytotropism and cytopathicity (2, 12, 21, 35, 43, 48). Furthermore, longitudinal studies from epidemiologically related patients have demonstrated that there is positive selection for independent genetic variation in the V1 and V2 domains in vivo (25). Interpretation of these genetic studies is difficult in the absence of detailed structure-function relationships in the HIV-1 envelope glycoprotein. Nonetheless, it may be presumed that sequence variation influences the ability of the viral envelope to fuse with the $CD4^+$ target cell. Amino acid substitutions in the V1 and V2 regions affect the function of the viral envelope (43) but do not affect the gp120-CD4 interaction (43, 50). The V2 region is also a target for neutralizing antibodies, most of which recognize conformationally sensitive structures centered on the middle portion of V2 (10, 11, 17, 20, 34, 43). The V1 region is also a possible neutralization site (37). Thus, a definition of the extent of sequence variation and the identification of conserved elements in these regions among different isolates may be useful for vaccine development (33).

In this study, we have characterized the prevalence of V1 and V2 sequences in a set of HIV-1-infected individuals. We analyzed by direct sequencing the extent of *env* V1 and V2 genetic diversity in 47 HIV-1-infected individuals, using uncultured PBMC. Sequences obtained by this method are considered to be the consensus sequence for each HIV-1 isolate,

* Corresponding author. Mailing address: The Aaron Diamond AIDS Research Center, New York University School of Medicine, 455 First Ave., 7th Floor, New York, NY 10016. Phone: (212) 725-0018. Fax: (212) 725-1126. Electronic mail address: Ho@adarc.nyu.edu.

TABLE 1. Sequence analysis of V1 and V2 regions

Sample	Patient's clinical status	No. of CD4 cells/mm^3	Length of aa sequence in V2[a]	Positively charged aa residues in V2	No. of potential NXT/S sites[b]	Phenotype[c]
X44	Asymptomatic	50	40	3	6	NSI
MR	AIDS	8	40	5	7	NSI
X46	Asymptomatic	218	40	5	6	NSI
Case C	AIDS	1,200	40	5	8	NSI
JC5010	NA[d]	NA	40	5	8	NSI
AD8	Acute Infection	>200	40	5	7	NSI
ACTG8	NA	NA	40	6	6	NSI
N70	Asymptomatic	>200	40	6	10	NSI
AD13	Acute Infection	407	40	9	8	NSI
ACTG6	NA	NA	41	5	8	NSI
X56	Asymptomatic	41	41	5	8	NSI
ACTG11	NA	NA	41	6	8	NSI
Case D2	Asymptomatic	750	41	6	7	NSI
RT	Asymptomatic	300	42	5	7	NSI
BABY A	Acute Infection	NA	42	5	6	NSI
Case B	AIDS	528	42	6	7	NSI
HU	Acute Infection	1,712	42	6	7	NSI
AD10	Acute Infection	>200	43	5	9	NSI
ACTG10	NA	NA	43	6	8	NSI
AD11	Acute Infection	266	44	5	8	NSI
X47	Asymptomatic	38	44	6	7	NSI
B27	Acute Infection	NA	45	5	8	NSI
AD6	Acute Infection	900	45	9	6	NSI
EJ	Asymptomatic	>200	46	4	7	NSI
76B	Asymptomatic	NA	46	8	7	NSI
Case D1	Asymptomatic	700	47	5	10	NSI
X48	Asymptomatic	26	40	6	7	NA
X52	Asymptomatic	127	40	5	7	NA
431	Asymptomatic	794	40	8	7	NA
250B	Asymptomatic	316	41	5	8	NA
AD12	Donor to AD13	360	42	6	8	NA
X45	Asymptomatic	34	43	5	7	NA
ABC	Asymptomatic	120	48	8	7	NA
ACTG3	NA	NA	40	6	6	SI
X42	Asymptomatic	38	40	6	6	SI
X49	Asymptomatic	6	40	7	10	SI
X57	Asymptomatic	7	41	4	7	SI
X50	Asymptomatic	36	41	4	8	SI
277B	Asymptomatic	<200	42	6	6	SI
Case A	AIDS	947	43	9	6	SI
ACTG1	NA	NA	45	9	6	SI
JSH	AIDS	<200	46	5	8	SI

[a] aa, amino acid.
[b] NXT/S, N-linked glycosylation.
[c] Tested in MT-2 cells.
[d] NA, not available.

reflecting the major variant in a quasispecies. Subsequent experiments to characterize the SI properties of these isolates were also performed. By cocultivating patient PBMC with normal donor PBMC, viral isolates were generated and used in a standard MT-2 cell syncytium assay. From these studies, we showed that the ability of a virus to induce syncytium formation does not correlate with specific patterns of sequence variability in V1 and V2.

This study involves 47 HIV-1-infected subjects, including case A and case D (7) from whom samples were taken at two different time points. Samples were collected from geographically distinct regions spanning North America, Australia, and Haiti and included HIV-1 patients in asymptomatic and symptomatic stages, seroconvertors with primary infection syndrome and their corresponding sexual partner (transmitters), rapid progressors, and long-term survivors. The patients' clinical statuses and $CD4^+$ cell counts are shown in Table 1. From the 49 HIV-1 isolates, 9 isolates were classified as SI by their ability to infect MT-2 cells; all were subjected to sequence analysis.

Proviral DNA from uncultured PBMC was amplified by nested PCR. Outer primers were P0 (nucleotides 6322 to 6345), 5′-GTG GGT ACC AGT CTA CTA TTA TGG GGT-3′, and P2 (nucleotides 7815 to 7786), 5′-GAC GCT GCG CCC ATA GTG CTT CCT GCT GC-3′, spanning V1 to V5 (52); inner primers were P1 (nucleotides 6529 to 6557), 5′-GAT GGT ACC GGA TAT AAT CAG TTT ATG GG-3′, and P10 (nucleotides 6970 to 6944), 5′-CCT AAT TCT AGA TGT ACA TTG TAC TGT-3′, from V1 to C2 (53), according to the pNL4-3 sequence in the Los Alamos HIV Database. The amplification products were visualized by ethidium bromide staining after agarose gel electrophoresis (data not shown),

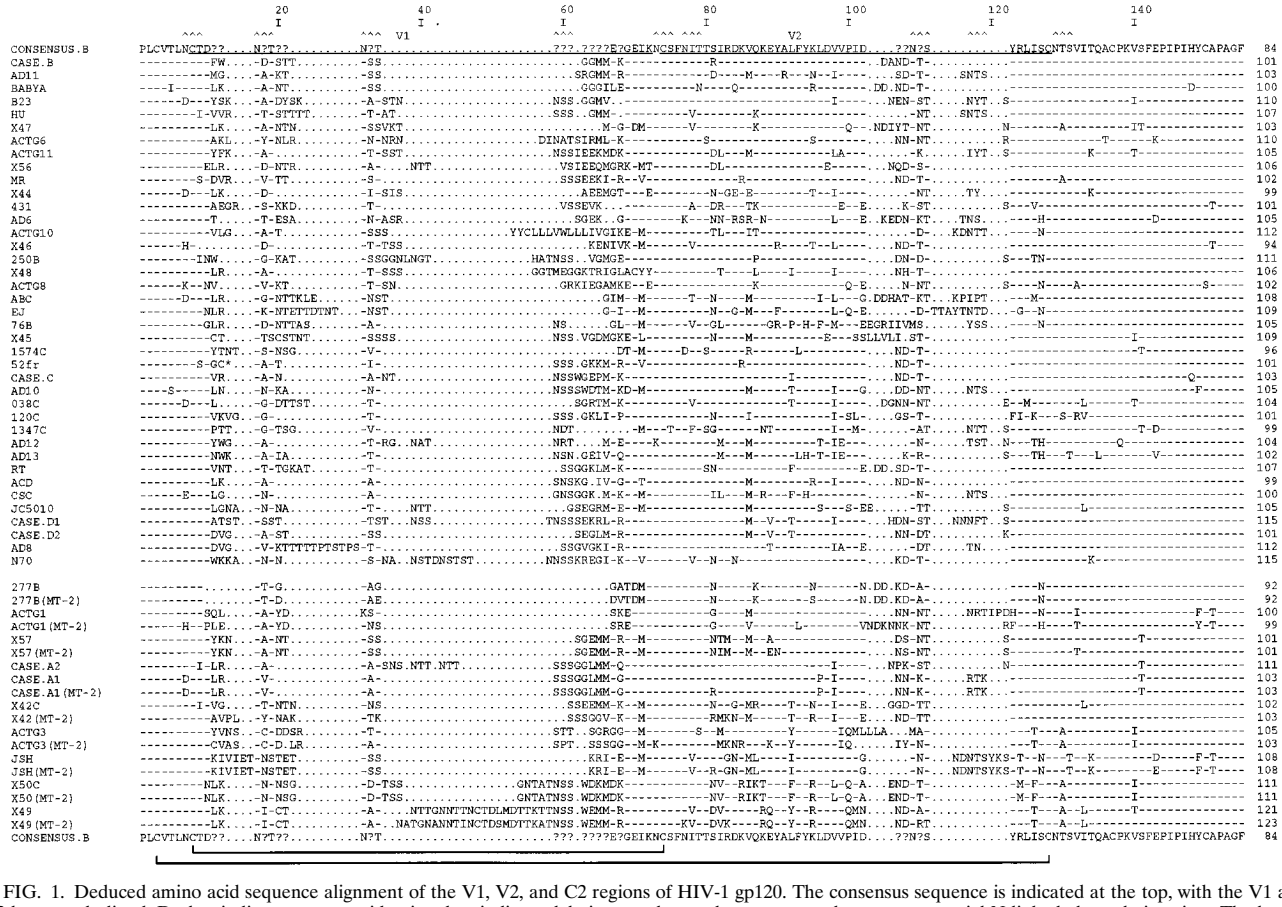

FIG. 1. Deduced amino acid sequence alignment of the V1, V2, and C2 regions of HIV-1 gp120. The consensus sequence is indicated at the top, with the V1 and V2 loops underlined. Dashes indicate sequence identity, dots indicate deletions, and stars denote stop codon. ^^^, potential N-linked glycosylation sites. The length of each sequence fragment is indicated on the right. The sequences below the spacing represent the sequences of SI isolates derived from uncultured PBMC and cultured in MT-2 cells. The brackets at the bottom indicate the V1 and V2 double-loop structures.

and their length was estimated with reference to a 440-bp fragment from the HXB2 clone of HIV-1; the average length was 460 ± 50 bp. To obtain the consensus nucleotide sequences, multiple copies of proviral DNA were subjected to limiting dilution. In serial dilutions, 6 to 10 copies of proviral DNA were subjected to PCR amplification (data not shown). This ensures that proviral copy numbers in samples are large enough to represent the provirus population in each individual yet small enough to allow detection of the major population in vivo. Although minor variants were apparent, the prevalent sequence type was discernible (data not shown).

Various deletions, insertions, substitutions, and duplications were present in the V1 and V2 regions, most of them being in frame, and only one alone was defective (sample X52) (Fig. 1, translated amino acid sequences). As noted above, length polymorphism was visible at the level of agarose gel electrophoresis. The length of the V1 segment ranged from 51 bp (sample 277B) to 144 bp (sample X49), with a mean of 87 bp. The V2 segment varied between 117 and 141 (sample ABC) bp, with a mean of 123 bp. Nine samples had a V2 length of 117 bp: X44, X46, X48, X49, X52, ACTG8, MR, 1574, and 431. In sum, the V1 plus V2 length was between 189 and 276 bp, with an average of 218 bp. Genetic distance was calculated by nucleotide alignments with removal of all gaps. Pairwise interisolate distances ranged from 2.0 up to 17.2%, with a mean distance of 10.5% between any two sequences across the V1-V2-C2 regions.

We compared the sequence data of 49 isolates in this study with a number of isolates from the Los Alamos database. Phylogenetic analysis was performed by the maximum parsimony method, both unweighted and weighted, on gap-stripped nucleotide alignments. Weighted parsimony analysis tries to produce a more accurate evolutionary picture by taking the actual frequency of base substitutions into account. All 49 isolates classified as subtype B. For clarity, only 34 sequences in our data set and 15 sequences from the Los Alamos database are represented in Fig. 2.

All four cysteine residues that define the V1 and V2 loops were uniformly conserved (residue 8 with 74 and residue 3 with 128 in our numbering system shown in Fig. 1) in our sequences. Alignment of the V1 sequences was not straightforward (Fig. 1), as comparison of the longest and shortest V1 sequences shows that more than one-third of the amino acid residues can be deleted or inserted. The V2 region is less variable than V1, especially in the segments from residues 157 to 185 and from residues 191 to 196 (numbered according to the HXB2 sequence in the Los Alamos database; 74 to 102 and 123 to 128, respectively [Fig. 1]). Length variation was not found in these V2 regions; only amino acid substitutions were found, with several positions being absolutely conserved. While there was considerable length variation between residues 185 to 191 (HXB2 numbering), most sequences contained a potential N-linked glycosylation site (Fig. 1). The frequency of amino acids at each position in flanking regions is shown in Fig. 3. There

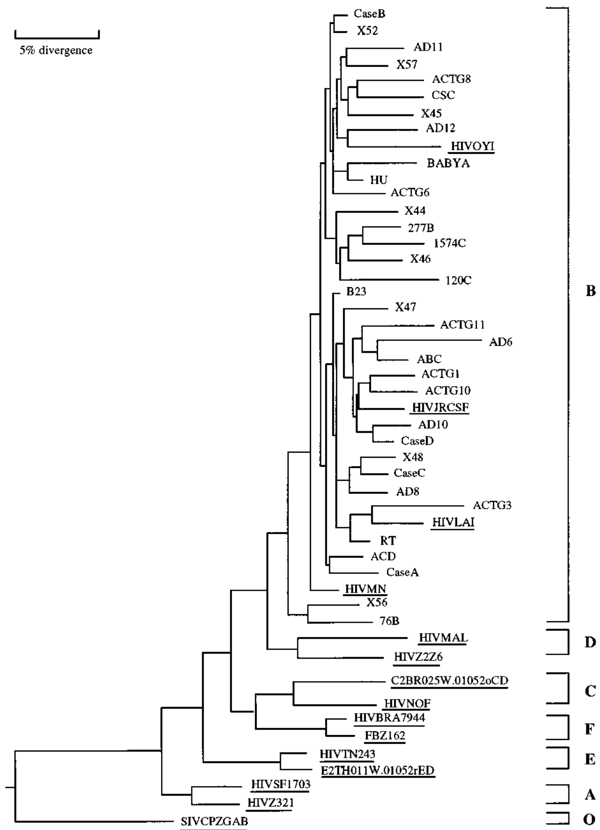

FIG. 2. Phylogenetic tree constructed on the basis of V1, V2, and C2 sequences by weighted maximum parsimony analysis. A total of 228 sites were analyzed, 157 of which show variation. Horizontal distances are drawn to scale, while the vertical separation is drawn only for clarity. Only 34 of the 50 HIV-1 isolates are shown, and underscored letter codes represent the published sequences derived from the Los Alamos database. Brackets on the right indicate sequence subtypes. SIVCPZGAB was used as the outgroup to root the tree. SIV, simian immunodeficiency virus.

was a tendency to conserve positively charged amino acids at five positions of V2: residues 166, 168, 171, 178, and 192 (HXB2 numbering), with 87 to 98% conservation (see below). Several continuous antibody epitopes have been mapped to a stretch of amino acids from residues 160 to 183 (HXB2) (34). In this region, over 80% of the residues were represented by a single amino acid at 16 of 24 positions. Important conformational or discontinuous antibody epitopes are influenced by residues at 183, 184, 191, 192, and 193 (HXB2 numbering) (33, 34, 43, 51); these residues showed greater than 80% conservation. These conserved features are presumed to contain structural elements critical to the function of gp120 V1 and V2. As shown in Fig. 3, the other residues were rather variable, with different kinds of amino acid substitutions, and there was no tendency to conserve amino acid character (nonpolar, polar, or charged) at particular positions.

Studies of simian immunodeficiency virus variation in macaques have demonstrated that common features in the V1 region include the serine- and threonine-rich motif TTTSTTT, as well as newly created sites for N-linked glycosylation, which are detected preferentially in variants isolated late in disease progression (36). The serine- and threonine-rich motifs have been proposed to create new sites for O-linked carbohydrate addition (3, 16). In this study, serine- and threonine-rich sequences in V1 were found, but the particular motif noted above was not observed. Several incomplete repeat sequences were observed as insertions in V1: NSSSSNSS starting from position 32 in subject X45, LLLVWLLLIV from position 56 in subject ACTG10, and NSTNANSTDNSTSTNNSS from position 39 in subject N70. There were also repeating amino acids from position 21 containing one to five threonine residues, from position 33 with one to four serine residues, and from position 63 with one to three glycine residues (numbering relates to Fig. 1). Despite the high variability of the V1 region, the two potential N-linked glycosylation sites were highly conserved, with 100 and 95% conservation at positions 17 and 32 (Fig. 1 numbering), respectively. The four N-linked glycosylation sites at positions 73, 77, 109, and 129 in V2 (Fig. 1 numbering) also showed 93 to 100% conservation. As shown in the consensus sequence with majority alignment, two potential N-linked glycosylation sites at positions 59 (V1) and 117 (V2) were not conserved (Fig. 1). Several V1 sequences contain additional N-linked glycosylation sites which could not be aligned with the consensus sequence.

PBMC used for direct sequencing were also cocultured with phytohemagglutin-stimulated PBMC from normal donors for a short period (5 to 7 days). Virus-containing supernatants obtained from these culture were titered, and 1,000 50% tissue culture infective doses were used to infect MT-2 cells (6, 7). Of 35 viral stocks obtained, only nine isolates were able to induce syncytium formation in MT-2 cells (Table 1). Because of selection during coculture, the phenotypes and genotypes of cultured and uncultured PBMC may not match. However, because all variants contained in NSI isolates presumably do not have SI variants that have a dorminant phenotype in the assay, the consensus sequence from uncultured PBMC still represents an NSI sequence. For those isolates with an SI phenotype, SI variants may be minor subtypes from the initial bulk patient isolate, so the consensus sequence from the bulk patient isolates may still represent NSI sequences.

To find the real sequences of SI variants, DNA was isolated from MT-2 cells infected with P1 isolates and subjected to PCR and sequence analysis. HIV-1-specific envelope se-

HXB2	157																												185		191					196
	C_{57}	S_{56}	F_{57}	N_{54}	I_{44}	T_{55}	T_{53}	S_{27}	I_{39}	R_{49}	D_{47}	K_{52}	V_{29}	Q_{44}	K_{51}	E_{47}	Y_{53}	A_{56}	L_{39}	F_{54}	Y_{54}	K_{44}	L_{50}	D_{55}	V_{32}	V_{50}	P_{47}	I_{52}	D_{35}	....	Y_{55}	R_{52}	L_{56}	I_{46}	S_{43}	C_{57}
		T_{1}		K_{2}	V_{10}	N_{1}	S_{4}	N_{14}	L_{5}	G_{3}	N_{5}	R_{4}	M_{20}	K_{7}	R_{3}	Q_{2}	R_{2}	R_{1}	T_{6}	L_{3}	H_{3}	R_{6}	P_{2}	E_{2}	I_{18}	E_{3}	Q_{8}	M_{4}	E_{10}		F_{1}	T_{2}	M_{1}	T_{7}	N_{10}	
				D_{1}	A_{1}	S_{1}		D_{6}	M_{5}	K_{3}	G_{2}	S_{1}	I_{3}	L_{3}	N_{3}	V_{2}	N_{1}		I_{4}			N_{3}	T_{2}		L_{6}	A_{2}	S_{2}	L_{1}	G_{4}		H_{1}	G_{1}		F_{1}	H_{4}	
					F_{1}			G_{4}	V_{3}	M_{2}	E_{1}		R_{2}	P_{1}	T_{1}	T_{2}			Y_{4}			S_{2}	F_{1}		M_{1}	I_{2}			N_{4}			I_{1}		K_{1}		
					T_{1}			R_{4}	T_{2}		Q_{1}		E_{1}	R_{1}		A_{1}			F_{3}			T_{2}	I_{1}						A_{1}			M_{1}		M_{1}		
								I_{1}	N_{1}		R_{1}		T_{1}	T_{1}		G_{1}			P_{1}				S_{1}						L_{1}					V_{1}		
								T_{1}	I_{1}				V_{1}			E_{1}													S_{1}							
									R_{1}							K_{1}													V_{1}							

FIG. 3. Frequency of occurrence of amino acid residues at each position in two regions (residues 157 to 185 and 191 to 196; numbering corresponds to amino acid sequence of HIV HXB2 envelope protein [HIV database, 1993]) that flank V2 hypervariable locus. The same two regions correspond to residues 74 to 102 and 123 to 128, respectively, in Fig. 1. The frequency of occurrence at each position is denoted by a subscript.

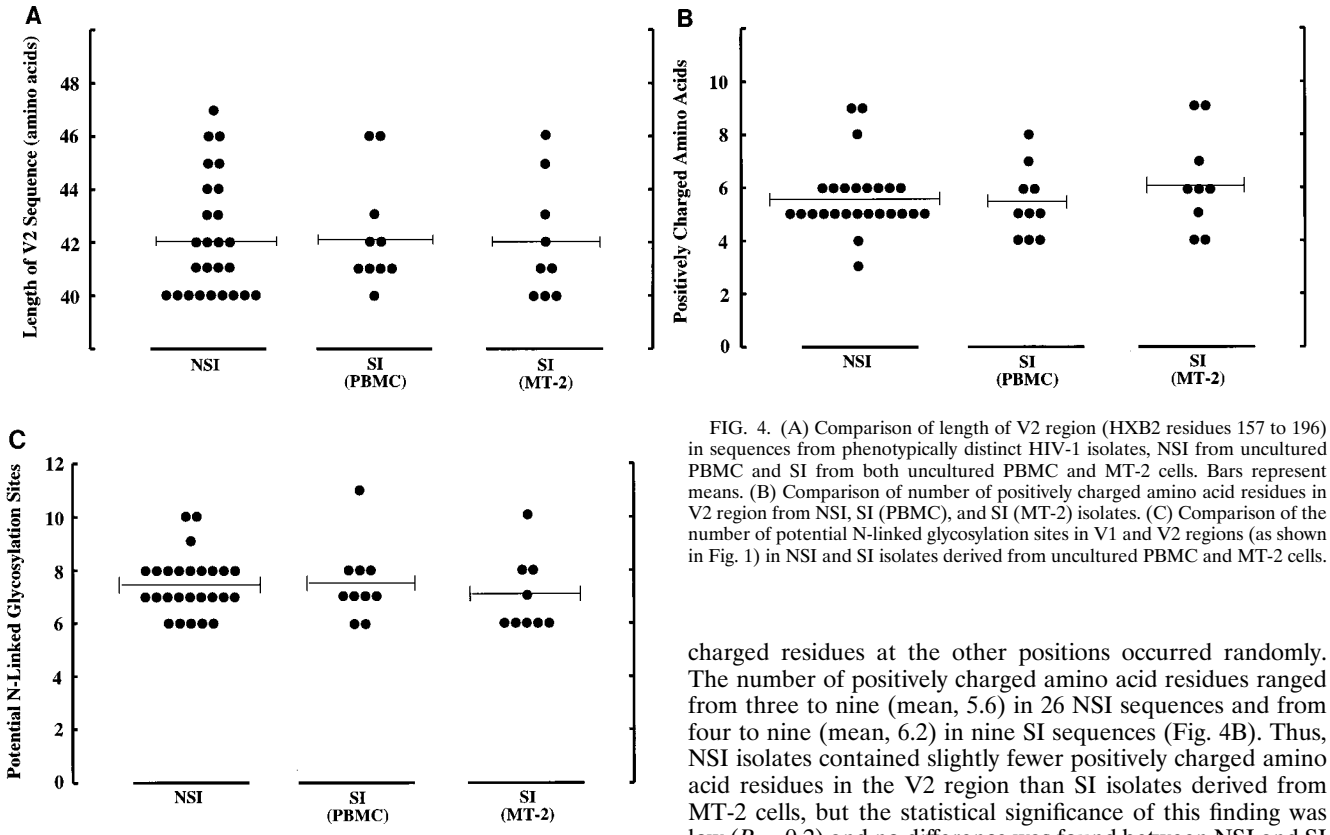

FIG. 4. (A) Comparison of length of V2 region (HXB2 residues 157 to 196) in sequences from phenotypically distinct HIV-1 isolates, NSI from uncultured PBMC and SI from both uncultured PBMC and MT-2 cells. Bars represent means. (B) Comparison of number of positively charged amino acid residues in V2 region from NSI, SI (PBMC), and SI (MT-2) isolates. (C) Comparison of the number of potential N-linked glycosylation sites in V1 and V2 regions (as shown in Fig. 1) in NSI and SI isolates derived from uncultured PBMC and MT-2 cells.

quences (V1-V2, C2-V5, and V1-V5 fragments of gp120) were not detected by PCR analysis in MT-2 cells infected with NSI viruses (data not shown), suggesting that the replication of NSI variants is blocked prior to reverse transcription in MT-2 cells (13). For the nine SI isolates grown in MT-2 cells, the V1-V2-C2 sequences were found to be more homogeneous than those present in uncultured PBMC by direct sequencing of the PCR product without diluting the proviral DNA. Comparison of the consensus sequences from uncultured PBMC with those from MT-2 cells showed that the two consensus sequences were identical in subject X50, whereas one and two amino acid residue changes were observed in cases A and JSH, respectively. Similar comparison of sequences in the other six SI isolates (277B, ACTG1, ACTG3, X42, X49, and X57) showed different degrees of divergence (Fig. 1). This result showed that despite the outgrowth of SI variants in culture, NSI variants were predominant in uncultured PBMC in most cases. Phylogenetic analysis showed that sequences from MT-2 cells and uncultured PBMC from any one individual clustered tightly together (data not shown). These viruses were, therefore, genetically related, whereas sequences from epidemiologically unrelated individuals were phylogenetically unrelated, and there was no evidence for a common origin of viruses in this study population.

We examined the amino acid sequences of V1 and V2 for distinguishing characteristics that might be associated with SI ability. Comparison of the length of V2 amino acid sequences did not reveal significant differences ($P > 0.5$) between NSI and SI (both from PBMC and MT-2 cells) isolates (Fig. 4A). For both NSI and SI isolates, five positively charged amino acid residues in the flanking regions of V2 hypervariable locus were rather conserved (see above), but substitutions of positively charged residues at the other positions occurred randomly. The number of positively charged amino acid residues ranged from three to nine (mean, 5.6) in 26 NSI sequences and from four to nine (mean, 6.2) in nine SI sequences (Fig. 4B). Thus, NSI isolates contained slightly fewer positively charged amino acid residues in the V2 region than SI isolates derived from MT-2 cells, but the statistical significance of this finding was low ($P = 0.2$) and no difference was found between NSI and SI isolates generated from uncultured PBMC (Fig. 4B). No differences in the number of positively charged amino acid residues in the V1 region or in its length were observed between NSI and SI isolates.

The creation of new N-linked glycosylation sites in V1 has been described for late variants of simian immunodeficiency virus (36), as well as in V2 (residues between 185 to 191 of HXB2 *env* protein) of HIV-1 (12). We therefore examined whether the changes in the N-linked glycosylation sites of the V1 and V2 regions correlated with viral SI capacity. There was no difference in the total numbers of N-linked glycosylation sites in the V1 and V2 regions between NSI and SI (derived from MT-2 cells) variants (Fig. 4C).

We also found no correlation between the length of V2, the number of positively charged amino acid residues in V2, or the number or location of potential N-linked glycosylation sites and the patients' $CD4^+$ cell counts (data not shown). Fourteen of 18 NSI isolates were derived from individuals with $CD4^+$ counts above 200/mm^3, whereas 6 of 7 SI isolates were derived from individuals with $CD4^+$ counts below 200/mm^3 (Table 1). This difference was statistically significant ($P < 0.0005$), and the finding was consistent with previous reports that NSI isolates predominate in individuals with asymptomatic infection (38, 52), while SI isolates were detected more often in patients with AIDS (44).

It has been known for almost a decade that the envelope gene of HIV-1 is highly heterogeneous, even for isolates obtained from a limited geographical area. Within *env*, the gp120-encoding region displays the greatest degree of variation, and within gp120, there are relatively variable (V) and relatively conserved (C) regions (27, 32, 42, 49). Much attention has been focussed on the V3 region, the principal neutralization determinant, because it contains strong neutralization epitopes, at least for T-cell line-adapted strains of HIV-1 (19).

However, the V1 and V2 domains also contain neutralization epitopes (10, 11, 17, 20, 21, 34, 37, 43, 47). As much less comparative sequence information is available for these regions of gp120, the influence of viral variation on HIV-1 neutralization via V1 and V2 has been difficult to assess. We have addressed this issue by sequencing the V1 and V2 regions of 47 clade B viruses from individual patients of American, Australian, and Haitian origin.

The V1 region was found to be hypervariable in our isolates, consistent with previous reports (25, 29). The only reasonably conserved features in V1 were two putative N-linked glycosylation sites (residues 17 and 32 in Fig. 1) and a tendency for preservation of a serine- and threonine-rich segment in the amino-terminal part of this region (Fig. 1). Serine- and threonine-rich motifs are potential sites for O-linked glycosylation, and as there is evidence for such posttranslational modification of gp120 (3, 16), it seems possible that some sequence conservation in V1 reflects the retention of O-linked glycosylation sites. A single amino-acid substitution (Ser to Asn) in the same region of V1 of HIV-1_{JRCSF} has been shown to expand viral tropism by an unknown mechanism (4). The altered residue is located between amino acids 35 and 62 in our sequences, the precise position being difficult to define because of hypervariability in the length of V1 (Fig. 1). Thus, we can make no simple generalization as to the effect of V1 sequence variation on HIV-1 tropism. The V1 loop of gp120 from HIV-1_{LAI} contains an immunodominant continuous epitope (34, 37) and is the target for neutralizing antibodies (37). This site is located within residues 142 to 153 of the HXB2 clone of HIV-1_{LAI} (residues 59 to 70 in our numbering system). Except for the last four residues, this stretch of V1 is hard to align, suggesting that neutralization targets in this region of V1 may be highly type specific.

Several conserved features could be identified in the V2 region. This hypervariable locus is flanked by two stretches in which about 70% of the residues were conserved by more than 80%. Mutagenesis studies of the HXB2 envelope show that alteration in these conserved regions impairs HIV replication, implying a role for some of the relatively invariant residues in preserving gp120 structure and/or function (43). Nonetheless, a reasonable degree of sequence variation can clearly be tolerated without compromising gp120 functions, for very few residues other than cysteines were absolutely conserved among all our isolates.

Neutralization sites have been identified in the V2 region, most of which are conformationally sensitive or discontinuous in nature (10, 17, 20, 33, 34, 43, 47, 51). A previous inspection of sequences of isolates from clades A to F that are or are not able to bind two murine anti-V2 monoclonal antibodies has shown that the binding of one monoclonal antibody to gp120 was largely predictable from primary sequences, but the other was not (33). For example, amino acid changes outside the V1 and V2 loops can affect the function or conformation of the V2 domain (9, 34, 43). Thus, it is not possible for generalized conclusions to be made as to the sensitivity of various viruses to neutralization via V2 on the basis of sequence data alone. Overall, given that the V1 and V2 loops both show extensive heterogeneity and that neutralization epitopes in the V2 cluster are generally weak, it seems unlikely that these segments of gp120 can be profitably exploited for vaccine development. This is consistent with recent studies in the simian immunodeficiency virus-rhesus macaque system (40).

We also investigated whether sequence variability in V1 and V2 was associated with HIV-1 phenotypic properties, in particular whether the length of V2 was correlated with the SI phenotype, as this has been reported by others (12). We were unable to find any correlation between V1 and V2 sequences and viral phenotype, except that there was a tendency for SI isolates to have a small increase in the number of positively charged residues in V2. The statistical relevance of the latter finding is questionable. What might account for the difference between our negative findings and the apparent V2 genotype-phenotype correlation reported by others (12)? It is unlikely that we have misclassified the phenotypes of our isolates. Firstly, like others (6, 7, 13), we found that only SI isolates could be detected in infected MT-2 cells. Thus, our sequence data from MT-2-grown isolates faithfully represent the genotype of SI isolates. This has been confirmed by DNA heteroduplex mobility assay, which showed that the predominant variants found in MT-2 cells matched only a minor population of the viruses presented in uncultured and cultured PBMC in most SI isolates (our unpublished results). Secondly, the evolution of phenotypically distinct variants during infection in vivo is a continuous process, in which slowly replicating NSI strains give rise first to rapidly replicating NSI viruses and then to SI variants (6, 45). Others have described intermediates in this pathway as stable NSI and switch NSI viruses, the latter being isolated after a phenotypic conversion (12). We have grouped together the switch NSI and stable NSI viruses in a single NSI category, on the grounds that in vitro studies demonstrate phenotypic homogeneity among this groups of viruses. Thus, we are confident that we are comparing two categories of virus with dissimilar phenotypes.

In summary, we can find no signature sequences in V2 that are predictive of viral phenotype, and to this extent we are unable to confirm the conclusions reached by others (12). Any relationship between the V2 sequence and phenotype may therefore be more subtle than has been appreciated or perhaps nonexistent. The former possibility would be consistent with the complexity of the V2 structure that has been revealed by studies with monoclonal antibodies and by mutational analysis. Perhaps, the conformation of the V2 domain in the context of the rest of the gp120-gp41 oligomeric complex does influence viral phenotype, but how it does so is not obvious from inspection of primary sequences.

Nucleotide sequence accession numbers. The nucleotide sequences have been deposited in GenBank under accession numbers U19621 to U19677.

We thank J. Moore for valuable suggestions and editorial assistance; Y. Cao and R. Connor for technical assistance; R. Connor and L. Zhang for helpful discussion; R. Koup, C. Cheng-Mayer, and E. Delwart for critical reading of the manuscript; and W. Chen for preparation of the figures. We are very grateful to K. MacInnes, B. Korber, and G. Myers at the Los Alamos National Laboratory for assistance with sequence analysis.

This work was supported by NIH grants AI24030, AI25541, AI27742 (Center for AIDS Research), AI32427, and AI27665, NIH contracts AI35168 and AI45218, the Ernst Jung Foundation, and the Aaron Diamond Foundation. N.W. was supported by a grant from the Pediatric AIDS Foundation.

REFERENCES

1. **Alizon, M., S. Wain-Hobson, L. Montagnier, and P. Sonigo.** 1986. Genetic variability of the AIDS virus: nucleotide sequence analysis of two isolates from African patients. Cell **46:**63–74.
2. **Andeweg, A. C., P. Leeflang, A. D. M. E. Osterhaus, and M. L. Bosch.** 1993. Both the V2 and V3 regions of the human immunodeficiency virus type 1 surface glycoprotein functionally interact with other envelope regions in syncytium formation. J. Virol. **67:**3232–3239.
3. **Bernstein, H. B., S. P. Tucker, E. Hunter, J. S. Schutzbach, and R. W. Compans.** 1994. Human immunodeficiency virus type 1 envelope glycoprotein is modified by O-linked oligosaccharides. J. Virol. **68:**463–468.
4. **Boyd, M. T., G. R. Simpson, A. J. Cann, M. A. Johnson, and R. A. Weiss.** 1993. A single amino acid substitution in the V1 loop of human immunodeficiency virus type 1 gp120 alters cellular tropism. J. Virol. **67:**3649–3652.

5. **Cheng-Mayer, C., D. Seto, M. Tateno, and J. A. Levy.** 1988. Biologic features of HIV-1 that correlate with virulence in the host. Science **240:**80–82.
6. **Connor, R. I., and D. D. Ho.** 1994. Human immunodeficiency virus type 1 variants with increased replicative capacity develop during the asymptomatic stage before disease progression. J. Virol. **68:**4400–4408.
7. **Connor, R. I., H. Mohri, Y. Cao, and D. D. Ho.** 1993. Increased viral burden and cytopathicity correlate temporally with CD4+ T-lymphocyte decline and clinical progression in human immunodeficiency virus type 1-infected individuals. J. Virol. **67:**1772–1777.
8. **Fouchier, R. A. M., M. Groenink, N. A. Kootstra, M. Tersmette, H. G. Huisman, F. Miedema, and H. Schuitemaker.** 1992. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type-1 gp120 molecule. J. Virol. **66:**3183–3187.
9. **Freed, E. O., and M. A. Martin.** 1994. Evidence for a functional interaction between the V1/V2 and C4 domains of human immunodeficiency virus type 1 envelope glycoprotein gp120. J. Virol. **65:**2503–2512.
10. **Fung, M. S. C., C. R. Y. Sun, W. L. Gordon, R.-S. Liou, T. W. Chang, W. N. C. Sun, E. S. Daar, and D. D. Ho.** 1992. Identification and characterization of a neutralization site within the second variable region of human immunodeficiency virus type 1 gp120. J. Virol. **66:**848–856.
11. **Gorny, M. K., J. P. Moore, A. J. Conley, S. Karwowska, J. Sodroski, C. Williams, S. Burda, L. J. Boots, and S. Zolla-Pazner.** 1994. Human anti-V2 monoclonal antibody that neutralizes primary but not laboratory isolates of human immunodeficiency virus type 1. J. Virol. **68:**8312–8320.
12. **Groenink, M., R. A. M. Fouchier, S. Broersen, C. H. Baker, M. Koot, A. B. van't Wout, H. G. Huisman, F. Miedema, M. Tersmette, and H. Schuitemaker.** 1993. Relation of phenotype evolution of HIV-1 to envelope V2 configuration. Science **260:**1513–1516.
13. **Groenink, M., R. A. M. Fouchier, R. E. Y. de Goede, F. de Wolf, R. A. Gruters, H. T. M. Cuypers, H. G. Huisman, and M. Tersmette.** 1991. Phenotypical heterogeneity in a panel of infectious molecular human immunodeficiency virus type 1 clones derived from a single individual. J. Virol. **65:**1968–1975.
14. **Hahn, B. H., M. A. Gonda, G. M. Shaw, M. Popovic, J. A. Hoxie, R. C. Gallo, and F. Wong-Staal.** 1985. Genetic diversity of the acquired immune deficiency syndrome virus HTLV-III: different viruses exhibit greatest divergence in their envelope genes. Proc. Natl. Acad. Sci. USA **82:**4813–4817.
15. **Hahn, B. H., G. M. Shaw, M. E. Taylor, R. R. Redfield, P. D. Markham, S. Z. Salahuddin, F. Wong-Staal, R. C. Gallo, E. S. Parks, and W. P. Parks.** 1986. Genetic variation in HTLV-III/LAV over time in patients with AIDS. Science **232:**1548–1553.
16. **Hansen, J. E. S., H. Clausen, S. L. Hu, J. O. Nielsen, and S. Olofsson.** 1992. An O-linked carbohydrate neutralization epitope of HIV-1 gp120 is expressed by HIV-1 *env* gene recombinant vaccinia virus. Arch. Virol. **126:**11–20.
17. **Ho, D. D., M. S. C. Fung, Y. Cao, C. Sun, T. W. Chang, and N. C. Sun.** 1991. Another discontinuous epitope on glycoprotein gp120 that is important in human immunodeficiency virus type 1 neutralization is identified by a monoclonal antibody. Proc. Natl. Acad. Sci. USA **88:**8949–8952.
18. **Holmes, E. C., L. Q. Zhang, P. Simmonds, C. A. Ludlam, and A. J. Leigh Brown.** 1992. Convergent and divergent sequence evolution in the surface envelope glycoprotein of human immunodeficiency virus type 1 within a single infected patient. Proc. Natl. Acad. Sci. USA **89:**4835–4839.
19. **Javaherian, K., A. J. Langlois, C. McDanal, K. L. Ross, L. I. Eckler, C. L. Jellis, A. T. Profy, J. R. Rusche, D. P. Bolognesi, S. D. Putney, and T. J. Matthews.** 1989. Principal neutralizing domain of the human immunodeficiency virus type-1 envelope protein. Proc. Natl. Acad. Sci. USA **86:**6768–6772.
20. **Kayman, S. C., Z. Wu, K. Revesz, H. C. Chen, R. Kopelman, and A. Pinter.** 1994. Presentation of native epitopes in the V1/V2 and V3 regions of human immunodeficiency virus type 1 gp120 by fusion glycoproteins containing isolated gp120 domains. J. Virol. **68:**400–410.
21. **Koito, A., G. Harrowe, J. A. Levy, and C. Cheng-Mayer.** 1994. Functional role of the V1/V2 region of human immunodeficiency virus type-1 envelope glycoprotein gp120 in infection of primary macrophages and soluble CD4 neutralization. J. Virol. **68:**2253–2259.
22. **Koot, M., A. H. V. Vos, R. P. M. Keet, R. E. Y. de Goede, M. W. Dercksen, F. G. Terpstra, R. A. Coutinho, F. Miedema, and M. Tersmette.** 1992. HIV-1 biological phenotype in long-term infected individuals evaluated with an MT-2 cocultivation assay. AIDS **6:**49–54.
23. **Kuiken, C. L., J. deJong, E. Baan, W. Keulen, M. Tersmette, and J. Goudsmit.** 1992. Evolution of the V3 envelope domain in proviral sequences and isolates of human immunodeficiency virus type 1 during transition of the viral phenotype. J. Virol. **66:**4622–4627.
24. **Kusumi, K., B. Conway, S. Cunningham, A. Berson, C. Evans, A. K. N. Iverson, D. Colvin, M. V. Gallo, S. Coutre, E. G. Shpaer, D. V. Faulkner, A. de Ronde, S. Volkman, C. Williams, M. S. Hirsch, and J. I. Mullins.** 1992. Human immunodeficiency virus type 1 envelope gene structure and diversity in vivo and after cocultivation in vitro. J. Virol. **66:**875–885.
25. **Lamers, S. L., J. W. Sleasman, J. X. She, K. A. Barrie, S. M. Pomeroy, D. J. Barrett, and M. M. Goodenow.** 1993. Independent variation and positive selection in *env* V1 and V2 domains within maternal-infant strains of human immunodeficiency virus type 1 in vivo. J. Virol. **67:**3951–3960.
26. **LaRosa, G. J., J. P. Davide, K. Weinhold, J. A. Waterbury, A. T. Profy, J. A. Lewis, A. J. Langlois, G. R. Dreesman, R. N. Boswell, P. Shadduck, L. Howard Holley, M. Karplus, D. P. Bolognesi, T. J. Matthews, E. A. Emini, and S. D. Putney.** 1990. Conserved sequence and structural elements in the HIV-1 principal neutralizing determinant. Science **249:**392–395.
27. **Leonard, C. K., M. W. Spellman, L. Riddle, R. J. Harris, J. N. Thomas, and T. J. Gregory.** 1990. Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. J. Biol. Chem. **265:**10373–10382.
28. **Lui, Z.-Q., C. Wood, J. A. Levy, and C. Cheng-Mayer.** 1990. The viral envelope gene is involved in macrophage tropism of the human immunodeficiency virus type 1 strain isolated from brain tissue. J. Virol. **64:**6148–6153.
29. **Martins, L. P., N. Chenciner, and S. Wain-Hobson.** 1992. Complex intrapatient sequence variation in the V1 and V2 hypervariable regions of the HIV-1 gp120 envelope sequence. Virology **191:**837–845.
30. **Meyerhans, A., R. Cheynier, J. Albert, M. Seth, S. Kwok, J. Sninsky, L. Morfeldt-Månson, B. Åsjö, and S. Wain-Hobson.** 1989. Temporal fluctuations in HIV quasispecies *in vivo* are not reflected by sequential HIV isolations. Cell **58:**901–910.
31. **Milich, L., B. Margolin, and R. Swanstrom.** 1993. V3 loop of the human immunodeficiency virus type 1 env protein: interpreting sequence variability. J. Virol. **67:**5623–5634.
32. **Modrow, S., B. H. Hahn, G. M. Shaw, R. C. Gallo, F. Wong-Staal, and H. de Wolf.** 1987. Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: predictions of antigenic epitopes in conserved and variable regions. J. Virol. **61:**570–578.
33. **Moore, J. P., F. E. McCutchan, S.-W. Poon, J. Mascola, J. Liu, Y. Cao, and D. D. Ho.** 1994. An exploration of antigenic variation in gp120 from clades A through F of human immunodeficiency virus type 1 using monoclonal antibodies. J. Virol. **68:**8350–8364.
34. **Moore, J. P., Q. Sattentau, H. Yoshiyama, M. Thali, M. Charles, N. Sullivan, S.-W. Poon, M. S. Fung, F. Traincard, J. E. Robinson, D. D. Ho, and J. Sodroski.** 1993. Probing the structure of the V2 domain of the human immunodeficiency virus type 1 surface glycoprotein gp120 with a panel of eight monoclonal antibodies: the human immune response to the V1 and V2 domains. J. Virol. **67:**6136–6151.
35. **O'Brien, W. A., Y. Koyanagi, A. Namazie, J. Zhao, A. Diagne, K. Idler, J. A. Zack, and I. S. Y. Chen.** 1990. HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside the CD4-binding domain. Nature (London) **348:**69–73.
36. **Overbaugh, J., and L. M. Rudensey.** 1992. Alterations in potential sites for glycosylation predominate during evolution of the simian immunodeficiency virus envelope gene in macaques. J. Virol. **66:**5937–5948.
37. **Pincus, S. H., K. G. Messer, P. L. Nara, W. A. Blattner, G. Colclough, and M. Reitz.** 1994. Temporal analysis of the antibody response to HIV envelope protein in HIV-infected laboratory workers. J. Clin. Invest. **93:**2505–2513.
38. **Roos, M. T. L., J. M. A. Lange, R. E. Y. de Goede, R. A. Coutinho, P. T. A. Schellekens, F. Miedema, and M. Tersmette.** 1992. Viral phenotype and immune response in primary human immunodeficiency virus type 1 infection. J. Infect. Dis. **165:**427–432.
39. **Saag, M. S., B. H. Hahn, J. Gibbons, Y. Li, E. S. Parks, W. P. Parks, and G. M. Shaw.** 1988. Extensive variation of human immunodeficiency virus type 1 *in vivo*. Nature (London) **334:**440–444.
40. **Schlienger, K., D. C. Montefiori, M. Mancini, Y. Rivière, P. Tiollais, and M.-L. Michel.** 1994. Vaccine-induced neutralizing antibodies directed in part to the simian immunodeficiency virus (SIV) V2 domain were unable to protect rhesus monkeys from SIV experimental challenge. J. Virol. **68:**6578–6588.
41. **Simmonds, P., L. Q. Zhang, F. McOmish, P. Balfe, C. A. Ludlam, and A. J. Leigh Brown.** 1991. Discontinuous sequence change of human immunodeficiency virus (HIV) type 1 *env* sequences in plasma viral and lymphocyte-associated proviral populations in vivo: implications for models of HIV pathogenesis. J. Virol. **65:**6266–6275.
42. **Starcich, B., B. H. Hahn, G. M. Shaw, P. D. McNeely, S. Modrow, H. Wolf, E. S. Parks, W. P. Parks, S. F. Josephs, R. C. Gallo, and F. Wong-Staal.** 1986. Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. Cell **45:**537–648.
43. **Sullivan, N., M. Thali, C. Furman, D. D. Ho, and J. Sodroski.** 1993. Effect of amino acid changes in the V1/V2 region of the human immunodeficiency virus type 1 gp120 glycoprotein on subunit association, syncytium formation, and recognition by a neutralizing antibody. J. Virol. **67:**3674–3679.
44. **Tersmette, M., R. E. Y. de Goede, B. J. M. Al, I. N. Winkel, R. A. Gruters, H. T. Cuypers, H. G. Huisman, and F. Miedema.** 1988. Differential syncytium-inducing capacity of human immunodeficiency virus isolates: frequent detection of syncytium-inducing isolates in patients with acquired immunodeficiency syndrome (AIDS) and AIDS-related complex. J. Virol. **62:**2026–2032.
45. **Tersmette, M., R. Gruters, F. de Wolf, R. E. Y. de Goede, J. M. A. Lange, P. T. A. Schellekens, J. Goudsmit, H. G. Huisman, and F. Miedema.** 1989. Evidence for a role of virulent human immunodeficiency virus (HIV) vari-

ants in the pathogenesis of acquired immunodeficiency syndrome: studies on sequential isolates. J. Virol. **63:**2118–2125.

46. **Tersmette, M., J. M. A. Lange, R. E. Y. de Goede, F. de Wolf, J. K. M. Eeftink Schattenkerk, P. T. A. Schellekens, R. A. Coutinho, J. G. Huisman, J. Goudsmit, and F. Miedema.** 1989. Association between biological properties of human immunodeficiency virus variants and risk for AIDS and AIDS mortality. Lancet **i:**983–985.
47. **Warrier, S. V., A. Pinter, W. J. Honnen, M. Girard, E. Muchmore, and S. A. Tilley.** 1994. A novel, glycan-dependent epitope in the V2 domain of human immunodeficiency virus type 1 gp120 is recognized by a highly potent, neutralizing chimpanzee monoclonal antibody. J. Virol. **68:**4636–4642.
48. **Westervelt, P., D. P. Trowbridge, L. G. Epstein, B. M. Blumberg, Y. Li, B. H. Hahn, G. M. Shaw, R. W. Price, and L. Ratner.** 1992. Macrophage-tropic determinants of human immunodeficiency virus type 1. J. Virol. **66:**2577–2582.
49. **Willey, R. L., R. A. Rutledge, S. Dias, T. Folks, T. Theodore, C. E. Buckler, and M. A. Martin.** 1986. Identification of conserved and divergent domains within the envelope gene of the acquired immunodeficiency syndrome virus. Proc. Natl. Acad. Sci. USA **83:**5033–5042.
50. **Wyatt, R., N. Sullivan, M. Thali, H. Repke, D. D. Ho, J. Robinson, M. Posner, and J. Sodroski.** 1993. Functional and immunologic characterization of human immunodeficiency virus type 1 envelope glycoproteins containing deletions of the major variable regions. J. Virol. **67:**4557–4565.
51. **Yoshiyama, H., H.-M. Mo, J. P. Moore, and D. D. Ho.** 1994. Characterization of mutants of human immunodeficiency virus type 1 that have escaped neutralization by monoclonal antibody G3-4 to the gp120 V2 loop. J. Virol. **68:**974–978.
52. **Zhu, T., H. Mo, N. Wang, D. S. Nam, Y. Cao, R. A. Koup, and D. D. Ho.** 1993. Genotypic and phenotypic characterization of HIV-1 in patients with primary infection. Science **261:**1179–1181.
53. **Zhu, T., N. Wang, A. Carr, S. Wolinsky, and D. D. Ho.** 1995. Evidence of coinfection by multiple strains of human immunodeficiency virus type 1 subtype B in an acute seroconvertor. J. Virol. **69:**1324–1327.