Diversity of the Envelope Glycoprotein among Human Immunodeficiency Virus Type 1 Isolates of Clade E from Asia and Africa


Henry M. Jackson Foundation, Rockville, Maryland; Walter Reed Army Institute of Research, Washington, D.C.; The Johns Hopkins University School of Hygiene and Public Health, Baltimore, Maryland; Chiang Mai University, Chiang Mai, Thailand; and Transgene, Strasbourg, France

Received 31 October 1995/Accepted 20 February 1996

Human immunodeficiency virus type 1 isolates of clade E, known to be largely responsible for the fulminating epidemic in Southeast Asia, have been derived exclusively from Asia and Africa. Here we provide full or partial sequences of the envelope glycoprotein gene from 13 additional clade E isolates from Asia representing patients in both early and late stages of disease. More extensive comparison of isolates within clade E by geographic locale, stage of disease, and year of isolation is now possible. The genetic diversity of clade E isolates from Asia, particularly among those derived from early-stage patients, is restricted compared with African isolates (mean interisolate distances in gp120, 5.4 and 20.2%, respectively). However, patients hospitalized with AIDS-related illnesses in Thailand harbored clade E isolates exhibiting broader interisolate diversity and with highly heterogeneous third hypervariable loop sequences. An additional pair of cysteine residues, predicting a novel disulfide bridge and present in 80% of clade E isolates from Asia, was uniformly absent from six African isolates. Clade E isolates in Thailand from early-stage subjects continue to be genetically similar to potential vaccine prototype strains, providing a favorable environment for the evaluation of genotype E candidate vaccines. However, evidence of increasing interisolate diversity is appearing among late-stage patients in Asia. This diversification of the clade E virus, if sustained, may impact preventive vaccine development strategies.

The genetic variability of human immunodeficiency virus type 1 (HIV-1) is a major challenge in the development of a globally effective vaccine. A variety of HIV-1 genotypes or clades, designated A through H, and an outlier (O) group, have been identified in the epidemic (30). Both vaccine design and vaccine evaluation are complicated by the diversity within and among genotypes and by the geographic intermixing of genotypes (7, 23, 40). The envelope glycoprotein, a major target of antiviral immunity and a pivotal element of current vaccine strategies, differs by as much as 30% between genotypes, while within genotypes, envelope sequences may vary up to 15% (13, 21, 30). Active surveillance of the prevalent genotypes and of the diversity of viral proteins is a key element in the preparation for HIV-1 vaccine field trials.

The global distribution of HIV-1 genotypes is constantly evolving. HIV-1 isolates of clade E were discovered when an increase in HIV-1 seropositivity in northern Thailand in 1989 prompted molecular characterization of the prevalent strains (22, 34). Since its introduction into Thailand, the clade E virus has spread rapidly, mostly through heterosexual transmission (31, 38, 39). It is now estimated that up to 500,000 individuals are infected with HIV-1 clade E in Thailand, and evidence is accumulating for the wider spread of this genotype in Southeast Asia (2, 12, 35) and to the Western Hemisphere (2, 5). In Africa, HIV-1 isolates of clade E have been recovered from the Central African Republic and from Cameroon (29, 32). The available data indicate that in Africa, genotype E is less prevalent and less widely distributed geographically than genotypes A, C, and D, whereas in Southeast Asia, clade E viruses are the most prevalent strains (7, 38, 39). World Health Organization projections suggest that clade E HIV-1 strains, subsequent to their introduction into Asia, may account for a significant proportion of global HIV-1 infections in coming decades (41).

The initial clade E HIV-1 isolates from Thailand, collected between 1989 and 1990, were remarkable for their genetic homogeneity (22, 34). Isolates from different patients varied less than 5% in their envelope protein. Indeed, the low genetic diversity of clade E in Thailand, coupled with other favorable epidemiological attributes, contributed to the selection of this country as a candidate vaccine trial site by several groups (9, 10, 26). A prototype clade E virus, CM235, now forms the basis for a multicomponent subunit vaccine whose evaluation in Thailand, subject to successful completion of the appropriate review processes, could begin next year (26).

Recent studies have provided evidence that clade E viruses are not intrinsically more homogeneous than other clades. The African clade E isolates show greatly increased interisolate diversity (29, 30) compared with the initial Asian isolates. A study of isolates from patients hospitalized for AIDS-related illness in northern Thailand showed extensive interisolate diversity within a selected region of the clade E envelope (42). These findings are important for HIV-1 vaccine field trials, as they establish the potential for, if not the occurrence of, a widening gap between the vaccine prototype and the local strains in Thailand.

It is unclear whether the diversity of the clade E isolates from patients with AIDS in Thailand, evaluated within a small region of the envelope gene including the third hypervariable

* Corresponding author. Mailing address: Henry M. Jackson Foundation Research Laboratory, 1600 E. Gude Dr., Rockville, MD 20850. Phone: (301) 217-9410. Fax: (301) 762-6240.
were SI. The phenotypes of isolates 1110, 1018, UR2, UR4, UR6, and UR7 have
was used to determine that isolates CM238, CM242, CM243, CM244, and
TH006, TH011, and TH022 were NSI (8). A standardized microtiter assay (14)
exception of isolate CM010, were of the SI phenotype (29, 32, 42). Isolates
African patients and those from symptomatic patients in Thailand, with the
syncytium-inducing (SI) or as non-syncytium-inducing (NSI). The isolates from
majority of the virus isolates in this study have been characterized either as
tral African Republic (29), and isolate CA10 was from Cameroon (32). The
CARELO, CARMBA, CAR4017, CAR4039, and CAR4071 were from the Cen-
unpublished data]) or in 1992 (TH022, TH006, and TH011 [40]). Isolates
CM239, CM240, CM241, CM242, CM243, and CM244 [reference 22 and our
in Africa CARELO CA10

Materials and Methods

Virus isolates. The virus isolates that were sequenced as a part of this study
were acquired in 1993 from individuals infected with HIV-1 in Southeast Asia.
Six isolates (KH03, KH05, KH08, CMU02, CMU08, and CMU101) were from
patients hospitalized with AIDS-related illnesses in northern Thailand: the V3
region of these isolates has been reported previously (42). Isolates 1018 and 1110
were from asymptomatic patients in Thailand. Isolates UR2, UR4, UR6, and
UR7 were from United Nations peacekeepers who acquired HIV-1 infection
while deployed to Cambodia in 1993 (2). Isolate POC-30506 was from a U.S.
serviceman who acquired HIV-1 infection while deployed in Thailand (5). Initial
virus isolation was performed by Ficoll-Hypaque separation of peripheral blood
mononuclear cells and subsequent cocultivation with stimulated donor periph-
eral blood mononuclear cells as described previously (6). Isolates from the Central
African Republic were further propagated in SupT1 or CEM cells as described
previously (4). All other sequences had been reported previously; they
include those from asymptomatic patients in Thailand in 1990 (CM235, CM238,
CM240, CM241, CM242, CM243, and CM244 [reference 22 and our unpublished
data]) or in 1992 (TH022, TH006, and TH011 [40]). Isolates CARELO, CARMBA,
CAR4017, CAR4039, and CAR4071 were from the Central African Republic (29), and isolate CA10 was from Cameroon (32). The
majority of the virus isolates in this study have been characterized either as
syncytium inducing (SI) or as non-syncytium inducing (NSI). The isolates from
African patients and those from symptomatic patients in Thailand, with the
exception of isolate CM100, were of the SI phenotype (29, 32, 42). Isolates
TH006, TH011, and TH022 were NSI (8). A standardized microtiter assay (14)
was used to determine that isolates CM238, CM242, CM243, CM244, and
POC30506 were NSI, whereas isolates CM235, CM239, CM240, and CM241
were SI. The phenotypes of isolates 1110, 1018, UR2, UR4, UR6, and UR7 have
d not been determined.

PCR amplification and cloning. Cells from virus cultures were used to prepare
DNA for PCR amplification of HIV-1 envelope gene sequences. DNA was
isolated and purified by using Qiagen kits (Qiagen, Chatsworth, Calif.). A 660-bp
segment beginning in the second constant region (C2) and extending through the
V5 domain of gp120 was amplified from isolates KH05, CMU02, CMU08,
CMU101, 1018, 1110, POc30506, UR2, UR4, UR6, and UR7, using primers and
thermocycling routines as described previously (3). From isolates KH03 and
KH08, a 2,500-bp segment containing the full-length gp160 gene was amplified as
described previously (21). For isolates CARELO, CARMBA, CAR4017,
CAR4039, and CAR4071, a 1,500-bp fragment containing the complete gp120
sequence was PCR amplified and cloned. Amplicons were cloned by using the
TA cloning system (Invitrogen Corp., San Diego, Calif.) as indicated by the vendor.

DNA sequencing and data analysis. DNA sequencing included at least two
clones per isolate from the V2-V5 region and one clone per isolate from the
gp120 gene and the full-length gp160 gene. Clones were fully sequenced on both
strands, using fluorescent dye terminators and an Applied Biosystems 373A DNA
sequencer (Applied Biosystems Inc., Foster City, Calif.) according to the
manufacturer’s instructions. Data were assembled by using DNASTar (DNASTar,
Inc., Madison, Wis.) or Sequencher (Genecodes, Inc., Ann Arbor, Mich.) soft-
ware on Macintosh computers and analyzed by using the Phylip package (version
3.52c) on a Sun computer. Phylogenetic trees of the DNA sequence were con-
structed from optimized alignments by using maximum likelihood (11). Protein
distance matrices were constructed by using Protdist as implemented in Phylip
version 3.52c, and using the Dayhoff PAM250 matrix.

Nucleotide sequence accession numbers. GenBank accession numbers for the
isolates analyzed are as follows: POC-30506, U48272; 1110, U48273; 1018,
U48274; UR2, U48275; UR4, U48276; UR6, U48277; UR7, U48278; KH03,
U48264; KH05, U48265; KH08, U48266; CMU02, U48267; CMU08, U48268;
CMU100, U48269.

Results

Derivation of virus isolates. Full or partial new envelope sequences of 13 clade E isolates collected in 1993 (Table 1) are
reported here. Seven were from patients in the early, asympto-
matic stage of infection; six were from patients with AIDS. Other isolates from Asia and Africa that were included in

Table 1. Clade E HIV-1 isolates

<table>
<thead>
<tr>
<th>Geographic area</th>
<th>1990</th>
<th>1992</th>
<th>1993</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia</td>
<td>Asia</td>
<td>Asia</td>
<td>Asia</td>
</tr>
<tr>
<td>Asia</td>
<td>Asia</td>
<td>Asia</td>
<td>Asia</td>
</tr>
<tr>
<td>Asia</td>
<td>Asia</td>
<td>Asia</td>
<td>Asia</td>
</tr>
<tr>
<td>Asia</td>
<td>Asia</td>
<td>Asia</td>
<td>Asia</td>
</tr>
<tr>
<td>Africa</td>
<td>Africa</td>
<td>Africa</td>
<td>Africa</td>
</tr>
</tbody>
</table>

FIG. 1. Available envelope sequences from clade E HIV-1 isolates. The subunit structure of gp160, together with the locations of domains V1 to V5, is shown at the top. Sequenced isolates are listed, with an indication of the relative length of available sequences. The geographic locale, chronology, and stage of disease of the patients from whom these isolates were obtained are given in Table 1. The blocks indicate the segments of envelope that were analyzed with respect to interisolate diversity.
interisolate comparisons are also included in Table 1; these were selected because they are the most complete with respect to the region of envelope that was sequenced.

**Genetic data available for analysis of clade E.** The envelope protein of HIV-1 is a patchwork of conserved and variable domains (36). Interisolate comparisons based on different regions of the envelope gene necessarily include different proportions of conserved and variable regions and thus provide different estimates of interisolate diversity. We have analyzed three segments of the envelope on the basis of availability of sequence data as shown in Fig. 1. The gp160 analysis includes all of the conserved segments of the envelope, providing a minimum estimate of interisolate diversity. Comparisons in the gp120 domain and, accordingly, in a subsegment of it that
encompasses V3, V4, and V5, will emphasize highly variable portions of the envelope.

The availability of sequence data also placed limitations on the range of intersolate comparisons that could be performed (Fig. 1). The full gp160 sequence was compared among isolates from Asian patients at various stages of disease progression. African and Asian isolates were compared in the external glycoprotein subunit gp120. Intersolate distances between different geographic locales, clinical stages, and years of virus isolation could be fully evaluated only by using a 660-bp segment, beginning in C2, upstream of the V3 loop, and ending just downstream of V5.

**Deduced protein distances among clade E envelope glycoproteins.** The diversity of the envelope protein among clade E isolates has been analyzed by stage of disease, by year of isolation, or by geographic origin, using the C2-V5 region of gp120 for analysis. Pairwise comparisons among the derived amino acid sequences were performed, and the distribution of protein distances was considered (Fig. 2). Among the patients acquiring HIV-1 infection in Asia during the period 1990 to 1993, the intersolate diversity was lower in those who were in the early stage of disease progression. However, isolates from symptomatic Asian patients were not as diverse as those from African patients (Fig. 2A). The subset of isolates from patients acquiring HIV-1 infection in Asia in 1993, when analyzed separately, showed a similar result, with a trend toward greater diversity among the isolates from patients with symptomatic HIV-1 infections (Fig. 2B; compare horizontally and diagonally hatched bars). The isolates from asymptomatic patients showed very similar distributions of intersolate distances when those from 1990 to 1992 were compared with those from 1993 (Fig. 2B, stippled and horizontally hatched bars). It is concluded that, considering the C-V5 region of gp120, isolates from asymptomatic patients acquiring HIV-1 infection in Asia showed the lowest intersolate diversity, and this low diversity was maintained in the subset of the samples collected most recently. Symptomatic patients in Asia showed intersolate diversity intermediate between that of asymptomatic Asian patients and that of symptomatic African patients.

The available data permitted fewer comparisons using larger segments of the envelope glycoprotein, but the results were in general accord with the matrix of comparisons for the C2-V5 region. In Table 2, the mean intersolate distances and the ranges are recorded for C2-V5, gp120, and gp160 sequences. Among the asymptomatic patients infected in Asia, the mean intersolate distances were 6.6% over the C2-V5 region and decreased to 5.4 and 4.5% when more conserved segments of envelope were included. Among the sequences that were available from symptomatic Asian patients, the mean intersolate distances ranged from around 12% in C2-V5 and gp120 to 9.1% in gp160. The African clade E isolates were the most diverse group examined; the mean intersolate distance was more than 20% in both C2-V5 and in gp120. Thus, regardless of the segment of the envelope that was examined, the isolates from symptomatic Asian patients showed intermediate diversity between those from asymptomatic infections in Asia and those from symptomatic infections in Africa.

**The V3 loop.** The V3 domain of HIV-1 envelope is the target of a vigorous immune response and has been extensively investigated for its potential to invoke effective antiviral immunity (15, 17, 19). Amino acid substitutions in the V3 loop can have a profound effect on immune recognition as evaluated in vitro (17, 27, 37). A comparison of the 35-amino-acid V3 loop could be made among a large number of clade E isolates by including published sequences from several sources (Table 3). In all, 92 isolates could be evaluated. Fifty-seven were from African patients known or presumed to be in the early stages of disease progression and included patients acquiring HIV-1 infection by heterosexual and by intravenous (16) routes. From late-stage patients, 23 sequences from Asia and 12 sequences from Africa were available.

The effect of sequence variability at each position in the 35-amino-acid V3 loop on recognition of HIV-1 by the human immune system is not known with precision, but positions that are strictly conserved among isolates would provide a minimum estimate of epitope conservation. Table 3 shows those amino acid residues that are entirely conserved among the isolates in the various studies. The conservation of the V3 loop among the isolates from early-stage patients acquiring infection in Asia is strong, with 17 of 35 invariant residues among 57 sequences. Only 9 of the 35 amino acids were conserved among the late-stage patients from Asia, whereas 16 were invariant among the African isolates. Interestingly, the conserved positions among the African isolates were not always the same as those that were invariant among isolates from early-stage Asian patients; residues 11, 14, and 20 are examples. The conservation of the V3 loop among late-stage Asian isolates is, unexpectedly, less than was seen in the African isolates.

**V4 and the adjacent region involved in CD4 receptor binding.** HIV-1 infection is mediated principally, if not entirely, by an interaction between gp120 and the CD4 molecule on the surface of susceptible cells (20, 24). The CD4 binding site (CD4bs) of gp120 includes a portion of the cysteine-bonded loop adjacent to V4, although amino acid substitutions outside this region can also profoundly influence the receptor interaction (1, 18, 25, 33). The V4-CD4bs cysteine-bonded structure, like other cysteine loops in the envelope, is highly conserved among HIV-1 isolates regardless of clade (21, 30). Figure 3 shows an alignment of available clade E protein sequences in this region.

One of the unique features of the first clade E virus isolates from Thailand was an additional pair of cysteine residues in the V4 loop (22). Subsequent collections, including most of the isolates described here, also contained the extra C-C pair. This

### Table 2. Interpatient distances in the envelope glycoprotein

<table>
<thead>
<tr>
<th>Region of envelope</th>
<th>Asymptomatic patients</th>
<th>Symptomatic patients</th>
<th>Africa, symptomatic patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Years of isolation</td>
<td>n</td>
<td>% Difference</td>
</tr>
<tr>
<td>C2-V5</td>
<td>1990–1993</td>
<td>18</td>
<td>6.6</td>
</tr>
<tr>
<td>gp160</td>
<td>1990–1992</td>
<td>9</td>
<td>4.5</td>
</tr>
</tbody>
</table>

*Intersolate distances were calculated for all pairwise combinations by using derived amino acid sequences.*
The feature is highly uncommon among the available sequences from primate lentiviruses but is found in 80% of the Asian clade E HIV-1 isolates analyzed to date (Fig. 3 and reference 30). Remarkably, the African clade E isolates uniformly lack this feature; the cysteines are either replaced by other residues or removed by an apparent deletion (Fig. 3). Four of the six African isolates also possess an apparent insertion of up to four amino acids in V4.

Figure 4 depicts an interpretation of the two structural motifs. While most amino acids within the CD4bs and at the base of the V4 loop are conserved, the V4 loop is usually shorter and can be predicted to be cross-linked by a disulfide bridge in 19 of the 24 Asian isolates studied. The Asian isolates that are missing the C-C pair do not segregate by stage of disease or by year of isolation, and they do not seem to contain the length variability of V4 seen in the African isolates. The number and position of potential glycosylation sites appear to be conserved. Additional data will be required to determine whether this potential to form an additional disulfide bridge results in functional alterations in the envelope of the Asian clade E isolates.

Interclade comparisons in HIV-1 envelope. Nucleotide sequences provide a more comprehensive basis for interisolate comparisons because the synonymous substitutions that are inapparent in protein sequences contribute to the analysis. The interisolate nucleotide sequence distances in clade E can be compared with those in other clades, whose epidemics spread has apparently developed over a different course. Here we compare them with isolates of clade B, which, although apparently rare in Africa, comprise a slowly spreading epidemic of immense geographic proportions and account for the vast majority of HIV-1 infections in Europe and in the Western Hemisphere, and with clade D, which has been recovered principally from sub-Saharan Africa.

Figure 5 shows a phylogenetic tree constructed by using maximum likelihood and including clade E gp120 DNA sequences, together with reference sequences from clades B, C, and D. The striking conservation of clade E isolates comprising the Asian epidemic is apparent. Isolates KH03 and KH08, from patients with AIDS, only modestly broaden this cluster. Inclusion of the African clade E isolates expands the clade E interisolate distances to resemble those found in other HIV-1 clades. The apparent genetic diversity of HIV-1 clades may be influenced by the range of geographic locales from which isolates are drawn, by the elapsed time between the establishment of a new epidemic and the collection of samples, and by the stage of disease progression of the patients from whom the samples were obtained.

**TABLE 3. Conservation and diversity of the V3 loop among clade E isolates**

<table>
<thead>
<tr>
<th>Locale</th>
<th>Stage of disease</th>
<th>No. of isolates</th>
<th>No. of invariant amino acids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asia Early</td>
<td></td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>Asia Late</td>
<td></td>
<td>16</td>
<td>9</td>
</tr>
<tr>
<td>Africa Early</td>
<td></td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Africa Late</td>
<td></td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>Africa Late</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Africa Late</td>
<td></td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td>Africa Late</td>
<td></td>
<td>35</td>
<td>30</td>
</tr>
</tbody>
</table>

The symbol □ indicates 100% conservation of the consensus amino acid in the indicated data set.

**DISCUSSION**

The epidemic spread of HIV-1 isolates of clade E has been highly heterogeneous and unpredictable. In 1988 or 1989, they apparently emerged from a relatively slow and circumscribed epidemic course in Africa to establish a rapidly expanding epidemic first in Thailand (22, 31, 34, 38, 39) and then in other locales in Southeast Asia (2, 7, 12, 35). The reason for the low interisolate diversity among Thailand isolates is not known with certainty, but it has been ascribed to an epidemiological clonal effect and to a preponderance of transmission events from early-stage, asymptomatic patients (7). A broadening of interisolate diversity could be expected as the epidemic spread continues and as a greater proportion of transmissions involve late-stage patients who have experienced a breakdown of antiviral immunity and a broadening of the viral quasispecies. Also, a maturing epidemic with the accumulation of viral replication cycles, each generating new mutations, would generate...
the potential for a broadening of interisolate distances, even with a preponderance of transmission events from early-stage patients. Determining the validity of these inferences is important for the planning of antiviral interventions; maintenance of the low interisolate variability of clade E provides a favorable environment for vaccine evaluation and, on the population level, may affect the durability of any protection afforded by a genotype E vaccine. The genetic diversity of clade E in Southeast Asia requires continual surveillance in the framework of vaccine development.

The availability of 13 full or partial envelope sequences collected in 1993 provides a new opportunity to evaluate the interisolate distances in clade E, with more representation of virus isolates from patients with AIDS and from a more recent time frame. Sequences from African clade E isolates provide a separate measure of interisolate variability, perhaps predictive of the eventual course of clade E over many years of sustained, albeit geographically circumscribed, epidemic spread. The interisolate distance among Asian clade E virus isolates from patients in the early stage of disease progression, even those reflecting more recent collections, remains low compared with that in other clades; conservation of both protein and DNA sequences was observed even in the most variable segments of the viral envelope. The mean interisolate distance of 6.6% in C2-V5 among 18 early-stage patients in Southeast Asia, compared with 20.9% among six late-stage African patients, together with the observation that the distribution of interisolate protein distances in C2-V5 are virtually nonoverlapping (Fig. 2A), provides the main evidence for this conclusion. However, the isolates from symptomatic patients in Asia were more diverse, particularly in the V3 loop. The proportional contribution of these more diverse clade E isolates to incident infections in Asia is difficult to predict with precision but will probably be modest, as the majority of HIV-1-positive individuals are in the early, asymptomatic phase in this locale. The relative genetic homogeneity of clade E in Thailand may be sustained in the near term.

The divergence of gp120 among African clade E isolates establishes that substantial interisolate diversity in clade E can develop under some circumstances. Among the many potential contributing factors to this diversity, the duration of the epidemic in Africa may be the most significant. Under this paradigm, the emergence of the Asian clade E variant could be...
viewed as the chance selection of a single strain from a heterogeneous group of clade E isolates, developed over many years in Africa, and its rapid amplification in a new locale. Is this simply a genetic bottleneck, as has been described during primary HIV-1 infection (43, 44), or did some feature of the strain that gained prevalence in Asia, perhaps related to host genetic makeup or to behavioral factors, contribute to its rapid epidemic spread? The genetic data alone cannot inform us on this point, but one observation seems particularly suggestive: the unique structure of the V4 loop in the majority of the Asian clade E isolates is not found in any of the six African isolates studied to date. Conceivably, this structural motif could have some population-level influence on the epidemiology of the virus, perhaps by influencing the CD4-gp120 interaction. However, any direct link between the epidemiology of this strain and its structural peculiarity in V4 remains speculative.

The study of the development of genetic diversity in clade E during the course of the epidemic in Southeast Asia affords a rare opportunity to evaluate the accumulation of mutations in different regions of the viral envelope, which may be subject to different degrees of immune selection and functional constraint. The interisolate variability among the isolates from patients with AIDS in Thailand is lower, overall, than that of African patients with AIDS, except for the V3 loop; here the diversity is at least equivalent to and may exceed that of the African isolates. The propagation of the virus isolates from Africa in continuous T-cell lines may have restricted recovery of the full spectrum of V3 loop variants, or some other population-level variables may have contributed to the observed differences.

It is also clear that the increased diversity of isolates from patients with AIDS in Thailand is not restricted to the variable domains of gp120; the interisolate distance of two isolates in gp160 (9.1%) is twice that in isolates from the early-stage patients (4.5%) (Table 2). These broadly distributed alterations in the clade E envelope may be predictive of a gradual drift toward the high diversity seen in the African isolates. Correspondingly, the limited duration of the window of opportunity for facilitated vaccine evaluation in Thailand should be factored into the planning and preparation for global vaccine development.

ACKNOWLEDGMENTS

We are grateful to Kevin Porter, Naval Medical Research Institute, Bethesda, Md., for providing the sequence segment from isolate POC-30506 and to the U.S. and Thai staff at the Armed Forces Research Institute of Medical Science, Bangkok, Thailand, for invaluable assistance in obtaining some of the isolates studied here.

REFERENCES
