Sequence Note

Protease Mutations in HIV-1 Non-B Strains Infecting Drug-Naive Villagers of Cameroon

FRANK A. J. KONINGS,1 PING ZHONG,2,3 MOISE AGWARA,9 LUCY AGYINGI,4 LEOPOLD ZEKENG,5 JACQUELINE M. ACHKAR,6 LEONARD EWANE,5 DR. SAA,7 EMMANUEL AFANE ZE,7 THOMPSON KINGE,7 and PHILLIPE N. NYAMBI2,8

ABSTRACT

To describe the presence of protease inhibitor (PI) resistance-associated mutations and subtype distribution in drug-naive villagers of six provinces of Cameroon, we sequenced the protease (PR) gene (297 bp) of 128 viruses. Secondary PI resistance-associated mutations were identified at five sites: L10I/V (16%), K20R (8%), M36I (98%), L63P (13%), and V77I (6%). No primary mutation in the PR was identified. Of the 128 specimens analyzed, subtypes A (11%), C (2%), D (6%), F2 (3%), G (6%), H (0.8%), J (6%), and CRF02_AG (60%) were identified. The mutations identified were not characteristic to any particular subtype. The absence of primary mutations, in addition to the few secondary mutations, gives good perspectives for PI treatment interventions in these rural areas.

IN SPITE OF THE EMERGENCE OF VARIOUS DRUG RESISTANCE-ASSOCIATED MUTATIONS in the viral protease (PR) gene, protease inhibitors (PIs) contribute substantially to the control of the AIDS epidemic in Europe and North America. However, of the estimated 42 million people infected with human immunodeficiency virus type 1 (HIV-1) worldwide, the majority (29.4 million) live in sub-Saharan Africa,1 the region with the broadest HIV-1 diversity. Since PIs were designed and tested for subtype B viruses only, they could be less effective in countries, like Cameroon, in which a broad genetic diversity exists.2 PI resistance is usually established during treatment by an initial primary mutation followed by gradual selection of secondary mutations.3 Primary mutations, which are located in the conserved regions of the PR gene, cause PI drug resistance by themselves, while secondary mutations further reduce drug susceptibility by increasing the replicative fitness of the virus, with little or no influence on resistance. Naturally occurring PI resistance-associated secondary mutations have different prevalence patterns for non-B subtypes and subtype B viruses. The most common natural polymorphisms in a panel of 187 non-B strains from drug-naive individuals worldwide were reported to be M36I (83%), L63P (17%), L10I/V (13%), K20R (10%), and V77I (2%).4

In Cameroon, the majority of the population lives in rural areas. While our recent reports reveal the HIV-1 infection in the rural villages is characterized by a broad genetic diversity,5,6 information on the natural polymorphisms that could be relevant to antiretroviral drugs, such as PIs, is sparse. Here we provide detailed insight into the naturally occurring mutations that could lead to PI resistance in HIV-1-infected villagers in six provinces of Cameroon. This information is needed to establish a baseline for monitoring resistance to PIs and aid in the

Departments of 1Microbiology and 2Pathology, New York University School of Medicine, New York, New York 10010. 3Shanghai Municipal Center for Disease Control and Prevention, Peoples Republic of China. 4Faculty of Science, University of Dschang, Dschang, Cameroon. 5Laboratoire de Santé Hygiène Mobile, Yaoundé, Cameroon. 6Department of Medicine, New York University School of Medicine, New York, New York 10010. 7Ministry of Public Health, Yaoundé, Cameroon. 8Research Center for AIDS and HIV Infection, VA Medical Center, New York, New York 10010. 9Faculty of Science, University of Yaoundé 1, Yaoundé, Cameroon.
successful introduction of antiretrovirals in the rural regions of Cameroon and neighboring countries. This work was part of a study aimed at determining HIV infection in rural villages in Cameroon.\textsuperscript{5,7} From a total of 300 HIV-1-positive plasma samples, 128 were randomly selected for analysis of the PR gene. These samples were obtained from individuals living in six provinces of Cameroon, including Center (\(n = 26\)), East (\(n = 13\)), South (\(n = 30\)), Southwest (\(n = 20\)), West (\(n = 28\)), and Northwest (\(n = 11\)) between 2000 and 2002. Of the 128 specimens analyzed, 56 were from males and 72 from females. Their ages ranged from 20 to 70 years for males (mean 39 years) and from 17 to 66 years for females (mean 31 years). None of the individuals in this study had previously received any kind of antiretroviral treatment.

Viral RNA was extracted from plasma as previously described by Boom \textit{et al.}\textsuperscript{8} One-tube reverse transcriptase polymerase chain reaction (RT-PCR) (Access RT-PCR System; Promega, Madison, WI) was performed for amplification of 297 bp of the PR region, corresponding to HXB2 positions 2258–2552. Two forward primers NYUPOLE (5’-AGGGAAGGCCAGGAATT-3', HXB2 location 2114–2132) and NYUPOLE7 (5’-GGGAATTTTCTTCAGAGCAG-3', HXB2 location 2125–2144) in combination with the reverse primer NYUPOLE8 (5’-TCTTCTGTCAATGGCCATTGT-3', HXB2 location 2615–2635) were used for RT-PCR. The two forward primers (NYUPOLE6 and NYUPOLE7) together in the same mix, in combination with a single reverse primer (NYUPOLE8), allowed the amplification of more PR sequences. For nested PCR the primers were NYUPOLE9 (5’-TCCTTAACTTCCCTCCAATCCT-3', HXB2 location 2241–2264) and NYUPOLE10 (5’-CTGGCACGGTTTCAATAGGACT-3', HXB2 location 2556–2577). For both RT-PCR and nested PCR the annealing temperature was 48°C.

PCR products were separated by agarose gel electrophoresis, gel extracted using the QIAquick spin gel extraction kit (QIAGEN, Valencia, CA) and directly DNA sequenced using an automated DNA sequencer (373XL; Applied Biosystems, Forster City, CA). The sequences were aligned with HIV-1 PR reference sequences of various subtypes from the HIV databases at http://www.hiv-web.lanl.gov. Multiple alignments were performed automatically by CLUSTAL X with minor manual adjustments and the sequence was cut to the 297-bp PR region by removal of the 5’ and 3’ ends. Sequences are available from GenBank with accession numbers AY359683–AY359810.

Phylogenetic analysis of the aligned sequences was performed by the neighbor-joining method of TREECON (Treecon for Windows, version 1.3b)\textsuperscript{9} Distance calculation was performed by Kimura’s two-parameter method. The statistical robustness of the neighbor-joining tree and reliability of the branching patterns were confirmed by bootstrapping (1000 replicates). The aligned DNA sequences were translated to amino acids by GeneRunner (GeneRunner 3.05 for Windows, http://www.generunner.com).

Previous studies have indicated that the 297-bp PR gene allows subtyping of nonrecombinant HIV-1 group M subtypes A, B, C, D, F, G, H, J, and K.\textsuperscript{10,11} Our results are in accordance with these observations and show that most of the strains (95%) could be subtyped, while only 5% were unclassifiable (U). Phylogenetic analysis revealed seven group M subtypes (A, C, D, F2, G, H, J), and CRF02_AG, while no subtype B was identified among the sequences analyzed. The HIV-1 group M subtype distribution in the six provinces studied is shown in Table 1. CRF02_AG was the predominant subtype in all the provinces and accounted for 60% of all the isolates in this study. HIV-1 subtypes A and G and CRF02_AG viruses were identified in all of the six provinces, except for the Northwest province where no subtype G was found. As reported earlier,\textsuperscript{7} the South and West provinces both had the broadest HIV-1 diversity with six different group M subtypes and unclassifiable strains.

The 297-bp PR sequences were translated to the corresponding 99 amino acids and analyzed for primary and secondary mutations that are associated with \textit{in vivo} PI resistance as reported for subtype B viruses.\textsuperscript{12} We found five secondary amino acid substitutions in the PR, namely L10I/V, K20R, M36I, L63P, and V77I. The prevalence of secondary amino acid substitutions for 89% of all the isolates in this study

| Province     | Number tested | AG\textsuperscript{a} | CRF02_AG | C | D | F2 | G | H | J | U\textsuperscript{a} |
|--------------|---------------|------------------------|----------|---|---|----|----|---|---|---|---|
| Center       | 26            | 2                      | 18       |   | 2 |    | 1  | 1 |   | 2 |
| East         | 13            | 1                      | 9        |   |   |    |    | 1 | 2 |   |
| South        | 30            | 3                      | 14       |   | 4 | 2  | 3  |   | 2 | 1 |
| Southwest    | 20            | 4                      | 10       |   | 2 | 2  | 2  |   | 2 |   |
| West         | 28            | 3                      | 16       | 3 | 2 | 1  | 1  |   | 2 |   |
| Northwest    | 11            | 1                      | 10       |   |   |    |    |   |   |   |
| Total        | 128           | 14                     | 77       | 3 | 8 | 4  | 8  | 1 | 7 | 6 |

\textsuperscript{a}AG, CRF02_AG; U, unclassifiable.

\textsuperscript{b}Percentage of total number of samples analyzed for each subtype.
(data not shown). The most common mutation among the five codons was M36I, which was present in 125 (98%) of the 128 sequences examined, while V77I (6%) was rare.

Next, we examined these non-B subtype sequences to identify PI resistance-associated mutations and to determine whether certain mutations predominated among specific non-B subtypes. The incidence of these different mutations of the respective subtypes is shown in Table 2. We found that no mutation predominated any particular subtype. The only two subtypes that harbored all five mutations in some isolates were subtypes A and F2. The three subtype C sequences had only the M36I mutation. Though the majority of the sequences examined belonged to CRF02_AG (60%), this group of sequences revealed only three different mutations, including M36I (99%), L10I/V (8%), and L63P (8%), and lacked K20R and V77I mutations. Subtype G viruses had the same three mutations as the CRF02_AG viruses (M36I, L10I/V, and L63P). Both subtypes D and J viruses harbored four mutations (L10I/V, K20R, M36I, and V77I), though only one (14%) of the seven subtype J and one (13%) of the eight subtype D viruses harbored all four mutations. The single subtype H sequence identified contained only two mutations, K20R and M36I.

Additionally, we investigated the sequences for the presence of multiple secondary mutations. While only one sequence had no secondary mutations, 86 (67%) of the 128 sequences harbored a single amino acid mutation, followed by 34 (27%) sequences with two mutations. In two (2%) sequences, four mutations with identified and only one sequence had accumulated five mutations. Of the AG sequences, 84% (65/77) had single mutations and 13% (10/77) had dual mutations (data not shown).

Finally, we examined all 128 sequences for PR mutations that have not been reported to be involved in drug resistance and found that the following mutations were present in more than 40% of the sequences: I13V (81%), K14R (53%), K20I (67%), E35D (41%), R41K (89%), P63L (73%), H69K (85%), and L89M (85%) (Table 2). All eight mutations were found together in sequences of subtypes A, G, and CRF02_AG, while only three of them were found in subtypes C, D, and H. Each mutation was present in more than 80% of the CRF02_AG sequences, except for E35D, which was present in 30% of the sequences.

This study has examined the presence of PI resistance-associated mutations in PR sequences of drug-naive HIV-1-infected villagers living in six provinces of Cameroon. Our analysis revealed a broad HIV-1 diversity with seven group M subtypes including subtype A, C, D, F2, G, H, J, and CRF02_AG. CRF02_AG viruses and group M subtype A viruses were found in all the provinces and caused the majority of all infections with 60 and 11%, respectively. These data support our previous studies in rural villages that revealed a broad env and gag subtype diversity and dominance of CRF02_AG infections. With the hypothesis that a broad genetic diversity within a population suggests a longstanding presence of the viruses in that population, the results of our studies suggest that these viruses must have been in these populations for many decades as previously proposed.

In the context of this broad genetic diversity, we wanted to study whether genes relevant to drug treatment have evolved and whether these viruses have acquired characteristic natural polymorphisms that would evade drug treatment. Analysis of the PI resistance-associated mutations revealed that primary mutations were absent, although five secondary mutations were identified in all the 128 sequences examined. Since secondary mutations have little or no effect on resistance in the absence of primary mutations, the presence of these mutations should be considered when starting PI treatment. Analysis of multiple mutations further revealed that the majority of the sequences had harbored only single mutations (67%) and that dual mutations could be found in only 27% of the sequences. The CRF02_AG viruses accounted for the largest group of viruses identified in this study. However, it is striking that almost all (99%) viruses harbor the M36I mutation and only two additional secondary mutations (L10I/V and L63P) were identified in 8% of the sequences. The large number of M36I could indicate that this is a natural polymorphism in non-B subtypes analyzed in this study. It remains to be confirmed whether individuals harboring isolates containing multiple secondary mutations may be at greater risk of virological failure during PI therapy, as has been hypothesized. With the results of our study in which the majority of the sequences contained only single mutations suggest that the PR genes of these viruses have not evolved sufficiently to allow for evasion of PI treatment and as such these drugs could be as effective as they are on subtype B viruses. However, to get a complete picture of resistance to other antiretrovirals drugs, further studies are warranted to examine naturally existing resistance mutations to drugs directed at other regions of the viruses, like the reverse transcriptase.

A total of five secondary amino acid mutation sites (codons 10, 20, 36, 63, and 77) were found in the sequences examined. Of these five sites, mutations at codons 10, 20, and 36 occur in up to 2–6% of drug-naive persons infected with subtype B viruses. Mutations at these same sites occurred in a greater proportion (8–98%) of the non-B subtype sequences examined and also occur at higher rates in non-B subtype isolates ranging from 10 to 83% as reported by others. With the exception of the M36I mutation that predominates non-B subtypes (95–97%) as observed in this study and by others, other mutations occurred at codon 10 (16%) and 20 (8%) with a similar prevalence as observed in subtype B isolates (13% (L10I/V), 10% (K20R)). The frequency of mutations at codons 63 and 77 in our study was 13 and 6%, respectively. Codon 63 is the most polymorphic PR position, but only L63P is implicated in PI resistance. While studies of non-B subtypes indicate that about 45% of isolates in untreated patients carry the L63P mutation, we observed that this mutation was present in only 13% of the sequences analyzed. Taken together, these patterns of mutations suggest some additional differences in the natural polymorphisms of these viruses from the rural villages to viruses from other regions of the world and thus should be considered when designing new PIs for use worldwide.

Of interest was the identification of several amino acid mutations that have not been associated with PI resistance but were present in more than 40% of the sequences. These mutations and the proportion of sequences carrying them included I13V (81%), K14R (53%), K20I (67%), E35D (41%), R41K (89%), P63L (73%), H69K (85%), and L89M (85%). All these mutations are located in the variable regions of the PR. Whether any
<table>
<thead>
<tr>
<th>Subtype</th>
<th>Number tested</th>
<th>Secondary PI resistance-associated</th>
<th>Nonresistance-associated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L10IV</td>
<td>K20R</td>
<td>M36I</td>
</tr>
<tr>
<td>A</td>
<td>14</td>
<td>2 (14)</td>
<td>3 (21)</td>
</tr>
<tr>
<td>AGa</td>
<td>77</td>
<td>6 (8)</td>
<td>—</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>—</td>
<td>3 (100)</td>
</tr>
<tr>
<td>D</td>
<td>8</td>
<td>3 (38)</td>
<td>—</td>
</tr>
<tr>
<td>F2</td>
<td>4</td>
<td>3 (75)</td>
<td>3 (75)</td>
</tr>
<tr>
<td>G</td>
<td>8</td>
<td>2 (25)</td>
<td>—</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>—</td>
<td>1</td>
</tr>
<tr>
<td>J</td>
<td>7</td>
<td>1 (14)</td>
<td>1 (14)</td>
</tr>
<tr>
<td>Ua</td>
<td>6</td>
<td>3 (50)</td>
<td>2 (33)</td>
</tr>
<tr>
<td>Total</td>
<td>128</td>
<td>20 (16)</td>
<td>10 (8)</td>
</tr>
</tbody>
</table>

*aAG, CRF02_AG; U, unclassifiable.
of these mutations will become relevant to escape from new PIs remains to be seen. It is important that the influence of these mutations on PI resistance is investigated and considered when designing new PIs that could be used in these regions where the viruses prevail.

In conclusion, this study shows that a broad HIV-1 diversity circulates in the rural villages of Cameroon. These HIV-1-infected drug-naive villagers harbor viruses with single or dual secondary PI resistance-associated mutations. The absence of primary mutations and the few secondary mutations portends well for PI treatment interventions in these rural populations as they have been carried out for subtype B infections.

ACKNOWLEDGMENTS

We thank the several doctors, nurses, and social workers in the different regions of Cameroon for their assistance, and the individual volunteers who made this study possible. We also thank the Minister of Public Health, Mr. Urbain Olanguena Awono for his support in these studies.

Financial support: National Institute of Allergy and Infectious Diseases (grant AI 47053); National Institute of Health (grant TW01254); Fogarty International Center at the NIH (grant TW01409); Funds from the Department of Veterans Affairs (Merit Review Award).

REFERENCES


Address reprint requests to:
Phillipe N. Nyambi
Department of Pathology
New York University School of Medicine
c/o V.A. Medical Center
423 E. 23rd Street, Room 18124N
New York, New York 10010
E-mail: phillipe.nyambi@mednyu.edu