Polymorphisms and resistance mutations in the protease and reverse transcriptase genes of HIV-1 F subtype Romanian strains

Simona Paraschiv, Dan Otelea*, Magdalena Dinu, Daniela Maxim, Mihaela Tinischi

Molecular Diagnostics Laboratory, 'Prof. Dr. Matei Bals' Institute for Infectious Diseases, Str. Calistrat Grozovici, nr. 1, sector 2, 021105 Bucharest, Romania

Received 15 July 2005; received in revised form 2 November 2005; accepted 15 November 2005
Corresponding Editor: Salim S. Abdool Karim, Durban, South Africa

KEYWORDS
HIV-1 F subtype; Resistance mutations; Natural polymorphisms

Summary
Objective: To evaluate the prevalence of resistance mutations in the genome of HIV-1 F subtype strains isolated from Romanian antiretroviral (ARV) treatment-naive patients and to assess the phylogenetic relatedness of these strains with other HIV-1 strains.
Methods: Twenty-nine HIV-1 strains isolated from treatment-naive adolescents (n = 15) and adults (n = 14) were included in this study. Resistance genotyping was performed by using Big Dye Terminator chemistry provided by the ViroSeq Genotyping System. The sequences of the protease and reverse transcriptase genes were aligned (ClustalW) and a phylogenetic tree was built (MEGA 3 software). For subtyping purposes, all the nucleotide sequences were submitted to the Stanford database.
Results: All the studied strains were found to harbor accessory mutations in the protease gene. The most frequent mutation was M36I (29 of 29 strains), followed by L63T, K20R, and L10V. The number of polymorphisms associated with protease inhibitor resistance was different for the two age groups. Intraphylogenetic divergence was greater for adults than for adolescents infected in childhood. All the strains were found to belong to the F1 subtype. The phylogenetic analysis revealed that Romanian strains clustered together, but distinctly from F1 HIV-1 strains isolated in other parts of the world (Brazil, Finland, and Belgium).
Conclusion: Protease secondary mutations are present with high frequency in the HIV-1 F subtype strains isolated from Romanian ARV treatment-naive patients, but no major resistance mutations were found.

© 2006 International Society for Infectious Diseases. Published by Elsevier Ltd. All rights reserved.
Introduction

With the advent of highly active antiretroviral therapy (HAART) the life expectancy of HIV infected patients has been significantly prolonged. However, it is now clear that mutations accumulate in treated patients, limiting the efficacy of the therapy and requiring adjustments of the antiretroviral regimen. Resistance genotyping has become standard of care in HIV infection management. 1 Resistance mutations are reported with increasing frequencies in treatment-naive individuals. 2—5 The prevalence of major resistance mutations for any antiretroviral drug observed in recent sequencing studies can range from 2% to 26% in North America, 23% to 26% in Spain, 4% to 17% in France, and 5% to 11% in Switzerland. 2 Information has accumulated so far mainly on type M subtype F strains which are predominant in the Western hemisphere. An increasing number of communications have shown that resistance mutations can occur spontaneously in the genome of viruses belonging to subtypes other than B isolated from untreated patients. 6—8 Most of these mutations are polymorphisms and are considered secondary mutations that by themselves play a limited, if any, role in resistance. However, their pre-existence might favor a more rapid evolution towards resistance when additional mutations are selected under therapy.

Since 1992, the HIV-1 strains isolated from patients worldwide have been classified into three distinct genetic groups: M (major) group, O (outlier) group, and N (non-M, non-O) group. Within the M group, 9 subtypes have been identified so far (A, B, C, D, F, G, H, J, and K) and intersubtype circulating recombinant forms (CRFs) and unique recombinant forms (URFs) have been described. Since most of the therapeutic experience has been derived from HIV-1 subtype B infections, the management of patients infected with HIV-1 strains of subtypes other than B is mostly the result of extrapolation.

In this study we present data obtained from the analysis of strains isolated from Romanian treatment-naive adolescents and adults. Previous anecdotal information has suggested that subtype F strains are rather common in Romania. 9,10 Our findings have confirmed these observations (the vast majority of the strains tested so far belong to the F subtype), although the strains displayed several dissimilarities with subtype F strains originating from other geographical regions. A recent study 11 has evaluated the subtype impact on polymorphism and resistance mutation development under ARV treatment and it has concluded that the positions associated with drug resistance for the B subtype are not significantly different from those found in other HIV-1 subtypes; however, the number of polymorphisms was higher in non-B subtype HIV-1 strains. Preliminary data indicated that the HIV-1 strains from Romania clustered distinctly within the F subtype. 12 It was therefore our purpose to assess the frequency of resistance mutations spontaneously occurring in the genome of the Romanian subtype F strains and to compare the relatedness of several subtype F strains isolated from naïve patients in Romania with other strains of subtype F and other subtypes.

Methods

Study population

We studied 29 HIV-1 strains isolated from Romanian patients, adolescents (n = 15) and adults (n = 14), originating from different regions of the country. All of the adolescents were the same age (14—15 years old) and most of the available epidemiological data suggest that they were infected by contaminated blood products of improper sterilization of needles used for parenteral treatment. The adults were all infected by the heterosexual route. The studied patients were ARV treatment-naïve at the time when the blood samples were drawn.

Clinical specimens

Plasma samples were obtained from the 29 HIV-1 infected patients. For all the patients the viral loads were higher than 10 000 copies/mL. Five mL of blood were collected in EDTA and the plasma was separated by centrifugation at 850 g for 15 min. The samples were frozen at —80 °C for storage until the RNA extraction.

Sequencing

The ViroSeq HIV-1 genotyping system (Celera Diagnostics, Alameda, USA) was used according to the manufacturer’s recommendations, with the exception of the RNA extraction which was performed with the Abbott m1000™ automatic nucleic acid extraction instrument. Briefly, after the determination of the viral load by LxC HIV RNA quantitative assay (Abbott Laboratories, Abbott Park, USA), the extracted RNA was reverse transcribed and then amplified by PCR. The thermal cycling was performed on a GeneAmp System 9700 (Applied Biosystems, Foster City, USA) thermal cycler. The 1.8 kb RT-PCR product was purified using MicroCon YM-100 concentrators (spin columns). The purity and the amount of the DNA amplification product were evaluated by running an aliquot on a 1% agarose electrophoresis gel. The amplicon was diluted to a standard 20—40 ng/sequencing reaction.

Big Dye Terminator chemistry was used for DNA sequencing: seven sequencing reactions were set up for covering the entire HIV-1 protease (PR) gene and two-thirds of the reverse transcriptase (RT) gene.

Capillary electrophoresis of the samples was performed on an ABI Prism 3100-Avant Genetic Analyzer (Applied Biosystems). Prior to loading, the purified products of the sequencing reactions were resuspended in formamide, denatured at 95 °C/2 min and chilled on ice. The raw analysis of the sequences was made using Sequencing Analysis software version 3.7 and then they were assembled with ViroSeq 2.5 HIV-1 Genotyping System software. The correctness of each electropherogram interpretation was validated by the operator and the sequences were saved in Fastq format. The resistance report generated by ViroSeq provides a list of nucleotide mutations detected in the PR (codons 1—99) and the RT gene (codons 1—383) known to determine resistance to antiretroviral drugs and infers a resistance profile for each strain. The Fastq files were then used to interpret the results by the HIV Seq Program (HIV-1 RT and protease search engine of queries), available at the Stanford database (http://hivdb.stanford.edu). The same resistance mutations were identified by both Virosel 2.5 and Stanford software.

Phylogenetic and sequence analysis

The subtyping of all the Romanian strains was done by using the publicly available algorithm at the Stanford database
The nucleotide sequences were translated into protein sequences in order to identify amino acid changes. The alignment of our sequences and the reference sequences of HIV-1 group M, subtype B, F1, F2 available in the Los Alamos HIV-1 database (www.hiv.lanl.gov) was performed using ClustalW multiple-sequence alignment software. The reference sequences used for the phylogenetic tree were as follows: HXB2 (reference strain for the B subtype), three F1 Brazilian strains (BZ163 AN173958, BZ126 173957, and 93BR020 AF005494 which is representative for the F subtype), a further two F1 European strains (FIN9363 AF075703 and VI850 AF077336), and four F1 Romanian strains (97R101 AF204010, 96R96 AF204004, 93R26 AF203986, and 97R105 AF204014). We also included in this analysis four F2 strains isolated from Cameroon (95CM-MP257C AJ249237, 95CM-MP255C AJ249236, CM53657 AF377956, and 02CM006BBY AY131158). Because all the Romanian sequences identified from the Los Alamos HIV-1 database were shorter (the first five codons of the PR gene were missing) our nucleotide sequences were trimmed to equivalent lengths (1149 n), representing 94 of the 99 PR gene codons and the first 289 codons of the RT gene. The phylogenetic analysis of all the sequences was made by using the neighbor-joining method as performed by the MEGA 3 software. The distance matrix was generated by Kimura’s two parameter model for nucleotides. The statistical strength of the neighbor-joining method was assessed by bootstrap resampling (1000 replicates). We used as outgroup the sequence of HIV-1 group O (MVP5180 L20571).

Our sample names are according to the WHO/UNAIDS nomenclature, indicating the year of isolation, the country of origin (two letter codes) and the isolate number.

Results

In order to detect mutations that confer resistance to protease inhibitors (PI), we compared the entire amino acid protease sequences of each strain included in this study with the subtype B consensus sequence (Stanford HIV database). We could not detect any primary mutations (D30N, G48V, I50V, V82A/T/F, I84V, L90M). However, secondary mutations were found in all 29 strains and affected six positions: codons 10, 20, 36, 77, and 93 (Table 1). The most frequent mutation encountered was M36I present in all the PR sequences. It was followed by L63T (93.1%; 27 of 29), K20R (72.4%; 21 of 29), and L10V (44.8%; 13 of 29). Previous studies suggested that M36I represents a molecular signature for non-B strains, particularly for the F subtype, whereas L63P is a distinct mark for B strains. In our study, substitutions at positions 36 and 63 occurred with very high frequencies.

All the sequenced strains had accessory resistance mutations that had been previously described as polymorphisms in the protease gene. Two mutations were present in one of 29 strains (3.4%), three mutations in 14 of 29 strains (48.2%), four mutations were found in 12 of 29 (41.3%) sequences and two strains (6.8%) had five mutations. The polymorphisms were not evenly distributed in the adolescent and adult groups (Figure 1). In certain positions that are known to be associated with PI resistance we observed other mutations, different from those involved in resistance. For example, V82I (sample 04RO523c) or L63A (sample 04RO4983) are known to have little or no impact on PI susceptibility.

In contrast with the protease gene, the analysis of 289 codons of the RT revealed no major and very few accessory

### Table 1 Amino acid substitutions in the protease gene sequences of HIV-1 F subtype Romanian strains at positions associated with resistance to protease inhibitors

<table>
<thead>
<tr>
<th>Codon position</th>
<th>Subtype B amino acid</th>
<th>Amino acid substitution</th>
<th>No. of strains</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>L</td>
<td>V</td>
<td>13</td>
<td>44.8</td>
</tr>
<tr>
<td>20</td>
<td>K</td>
<td>R</td>
<td>21</td>
<td>72.4</td>
</tr>
<tr>
<td>36</td>
<td>M</td>
<td>I</td>
<td>29</td>
<td>100</td>
</tr>
<tr>
<td>63</td>
<td>L</td>
<td>T</td>
<td>27</td>
<td>93.1</td>
</tr>
<tr>
<td>77</td>
<td>V</td>
<td>I</td>
<td>1</td>
<td>3.4</td>
</tr>
<tr>
<td>82</td>
<td>V</td>
<td>I</td>
<td>1</td>
<td>3.4</td>
</tr>
<tr>
<td>93</td>
<td>I</td>
<td>L</td>
<td>4</td>
<td>13.8</td>
</tr>
</tbody>
</table>

(Rega HIV-1 subtyping tool version 1.0). The nucleotide sequences were translated into protein sequences in order to identify amino acid changes. The alignment of our sequences and the reference sequences of HIV-1 group M, subtype B, F1, F2 available in the Los Alamos HIV-1 database (www.hiv.lanl.gov) was performed using ClustalW multiple-sequence alignment software. The reference sequences used for the phylogenetic tree were as follows: HXB2 (reference strain for the B subtype), three F1 Brazilian strains (BZ163 AV173958, BZ126 AV173957, and 93BR020 AF005494 which is representative for the F subtype), a further two F1 European strains (FIN9363 AF075703 and VI850 AF077336), and four F1 Romanian strains (97R101 AF204010, 96R96 AF204004, 93R26 AF203986, and 97R105 AF204014). We also included in this analysis four F2 strains isolated from Cameroon (95CM-MP257C AJ249237, 95CM-MP255C AJ249236, CM53657 AF377956, and 02CM006BBY AY131158). Because all the Romanian sequences identified from the Los Alamos HIV-1 database were shorter (the first five codons of the PR gene were missing) our nucleotide sequences were trimmed to equivalent lengths (1149 n), representing 94 of the 99 PR gene codons and the first 289 codons of the RT gene. The phylogenetic analysis of all the sequences was made by using the neighbor-joining method as performed by the MEGA 3 software. The distance matrix was generated by Kimura’s two parameter model for nucleotides. The statistical strength of the neighbor-joining method was assessed by bootstrap resampling (1000 replicates). We used as outgroup the sequence of HIV-1 group O (MVP5180 L20571). Our sample names are according to the WHO/UNAIDS nomenclature, indicating the year of isolation, the country of origin (two letter codes) and the isolate number.

Results

In order to detect mutations that confer resistance to protease inhibitors (PI), we compared the entire amino acid protease sequences of each strain included in this study with the subtype B consensus sequence (Stanford HIV database). We could not detect any primary mutations (D30N, G48V, I50V, V82A/T/F, I84V, L90M). However, secondary mutations were found in all 29 strains and affected six positions: codons 10, 20, 36, 77, and 93 (Table 1). The most frequent mutation encountered was M36I present in all the PR sequences. It was followed by L63T (93.1%; 27 of 29), K20R (72.4%; 21 of 29), and L10V (44.8%; 13 of 29). Previous studies suggested that M36I represents a molecular signature for non-B strains, particularly for the F subtype, whereas L63P is a distinct mark for B strains. In our study, substitutions at positions 36 and 63 occurred with very high frequencies.

All the sequenced strains had accessory resistance mutations that had been previously described as polymorphisms in the protease gene. Two mutations were present in one of 29 strains (3.4%), three mutations in 14 of 29 strains (48.2%), four mutations were found in 12 of 29 (41.3%) sequences and two strains (6.8%) had five mutations. The polymorphisms were not evenly distributed in the adolescent and adult groups (Figure 1). In certain positions that are known to be associated with PI resistance we observed other mutations, different from those involved in resistance. For example, V82I (sample 04RO523c) or L63A (sample 04RO4983) are known to have little or no impact on PI susceptibility.

In contrast with the protease gene, the analysis of 289 codons of the RT revealed no major and very few accessory

Figure 1 The distribution of PI-resistance associated polymorphisms in HIV-1 Romanian strains isolated from adolescents (gray boxes) and adults (black boxes). The two series were found to be different when compared with Fisher’s exact test (p < 0.02).
mutations. The mutations associated with resistance to nucleoside reverse transcriptase inhibitors (NRTIs) and non-nucleoside reverse transcriptase inhibitors (NNRTIs) have been well characterized and differ between these two classes of RT inhibitors: resistance mutations to NRTIs are localized in ‘fingers’, ‘thumb’, and ‘palm’ domains of the reverse transcriptase while mutations responsible for resistance to NNRTIs are localized into the hydrophobic cavity (NNRTI binding site). The V179D mutation appeared in only two of the samples (04RO2165 and 04RO2301) isolated from adults. Low-level resistance to each NNRTI has been attributed to this mutation, but its effect is considered insignificant in the absence of other mutations. In the RT gene we identified some atypical mutations at key positions known to be linked with drug resistance, namely M41R (03RO155), T69AT (03RO1436), Y318I (03RO155), and K101RK (04RO4420), but these mutations have not been reported as important for resistance to NRTIs or NNRTIs.

**Phylogenetic analysis**

The subtype analysis of the PR and RT sequences indicated that all 29 HIV-1 strains belonged to the F1 subtype, with a bootstrap support ranging between 76% and 100%. The 1149 nucleotide sequences of the 29 Romanian strains representing the PR gene and half of the RT gene were used to build a phylogenetic tree. Our strains were compared with reference strains of different subtypes (B, F1, F2) and sequences of HIV-1 strains isolated between 1993 and 1997 from Romanian patients, available in the Los Alamos HIV-1 sequence database. We used as outgroup the MVP5180 sequence (HIV-1 group O). The result of the phylogenetic analysis is presented in Figure 1. As expected, the algorithm could efficiently discriminate among different subtypes. The studied strains clustered with the F1 reference strains and were distinct from the F2 strains. Moreover, all the Romanian strains clustered together with the other Romanian strains identified from the Los Alamos HIV-1 sequence database, but were distinct from F1 Brazilian strains or F1 European strains (Belgium and Finland), suggesting a different epidemiological context for the infection.

The phylogenetic analysis did not discriminate between the HIV-1 strains isolated from the two age groups. However, the average sequence divergence was higher in the adult group (6.4%) in comparison with the adolescent group (5.1%).

It has been shown that although the pol gene is very conserved, it can be used for subtyping purposes. Our data confirm this observation and, furthermore, show that the gene can discriminate between groups of strains belonging to the same subtype (Romanian and Brazilian F strains clustered distinctly).

**Discussion**

Recent studies have tried to assess the influence that different HIV-1 subtypes exert on the efficacy of ARV treatment and the progression to AIDS. The results have often been contradictory, but there is evidence that strains belonging to group O are naturally resistant to NNRTIs and some G subtype strains are less susceptible to PIs.
Phylogenetic analysis of PR and RT nucleotide sequences (1149 bp) from 29 Romanian strains and reference strains available in the Los Alamos HIV-1 database. The tree was generated as described in the Methods. The reference strains are presented in the same section and are distinctly marked (●). The numbers at the nodes indicate the percentage of the bootstrap value.
variants during ARV treatment in viruses where one or more secondary mutations are already present.

In this study direct sequencing of PCR products was performed. Therefore, each sequence represents the consensus genotype; it is possible that additional mutations may exist in these HIV isolates, but PCR cloning followed by sequencing would be necessary to detect them.

Kantar et al. have recently evaluated the frequency of polymorphisms and resistance mutations in protease and reverse transcriptase genes in non-B subtypes. Our findings show a higher frequency of protease polymorphisms that are secondary resistance mutations as well. We interpret this observation as due to the higher homogeneity of the strains we examined (only 18% of the subtype F strains analyzed by Kantar et al. were isolated from Romanian patients). Furthermore, this higher frequency correlates with the increased frequency of the same polymorphisms as detected in viruses isolated from Romanian treated patients (data not shown).

In conclusion, no major mutations associated with resistance to NRTIs, NNRTIs, and PIs were detected in subtype F strains isolated from Romanian patients, but many accessory substitutions were found as natural variants, mostly in positions associated with resistance to protease inhibitors. Their incidence is partly different from that reported for other subtype F HIV-1 strains. It is not yet clear if the high frequency of accessory protease mutations may contribute to a more rapid crossing of the genetic resistance barrier. Sequential phenotypic analyses and long-term follow-up of patients under antiretroviral treatment will be necessary to evaluate the clinical significance of these minor PR mutations.

**Acknowledgements**

The authors wish to thank Adriana Hristea, Gratiela Tardei, Sorin Petrea, and Mihai Rosca for technical support.

**Conflict of interest:** No conflict of interest to declare.

**References**