The HIV-1 protease substitution K55R: a protease-inhibitor-associated substitution involved in restoring viral replication

E. S. Margerison¹*, M. Maguire¹, D. Pillay², P. Cane² and R. C. Elston¹†

¹GlaxoSmithKline, Gunnels Wood Road, Stevenage SG1 2NY, UK; ²Centre for Infections, Health Protection Agency, 61 Colindale Avenue, London NW9 5HT, UK

Received 17 September 2007; returned 4 November 2007; revised 17 December 2007; accepted 18 December 2007

Objectives: The identification and in vitro characterization of novel protease mutations strongly associated with known protease resistance mutations.

Methods: The association between pairs of protease amino acid substitutions was identified using a database of protease sequences derived from protease inhibitor-experienced patients (n = 803). In vitro characterization included drug susceptibility and viral replication studies performed on recombinant viruses harbouring site-directed mutations.

Results: The K55R mutation, which is not a natural polymorphism, was identified to be strongly associated with protease mutations M46I/L and to a lesser extent L24I, I54V and V82A/T/S/F. In vitro characterization of the K55R substitution indicated a primary role for this substitution in increasing replicative capacity in the presence of specific protease mutations.

Conclusions: The K55R mutation is a secondary drug resistance mutation that can improve viral replication capacity in the presence of other primary protease mutations.

Keywords: HIV, fitness, resistance, mutations

Introduction

Highly active antiretroviral therapy (HAART) to combat HIV-1 was introduced in the mid-1990s and resulted in a significant decline in AIDS-related mortality.¹,² The initial protease inhibitor (PI)-based HAART regimens suffered from variable pharmacokinetics, high pill burden and drug-related side effects. More recently, the introduction of new PIs with improved potency, dosing and side effect profiles, together with the use of ‘boosted’ PI-based regimens, has led to significant improvements in efficacy.³ However, these improvements are threatened by the development of cross-resistance within the PI class and this has resulted in resistance testing being recommended both before and after virological failure of an antiretroviral regimen (DHHS Guidelines 2006).

A consequence of widespread resistance testing has been the accumulation of genotypic and phenotypic resistance information in large databases.⁴,⁵ Such databases have and continue to offer the opportunity to investigate the associations between amino acid substitutions within protease and reverse transcriptase. The first objective of this analysis was to investigate the association between non-consensus protease substitutions using a large database of HIV-1 protease genotypes derived from PI-experienced patients and to compare the findings with associations previously reported.⁶–⁸ The second objective was to investigate the role of previously uncharacterized protease substitutions that were associated with PI resistance mutations using recombinant virus.

Materials and methods

Patient databases and analysis

The HIV-1 protease amino acid sequence database used in this study was composed of 803 protease sequences derived from PI-experienced patients involved in clinical trials PRO30017 (n = 457) and APV30003 (n = 346).⁹,¹⁰ Eligibility criteria for screening in PRO30017 required that subjects had experienced virological failure on at least one PI-containing regimen. Eligibility...
criteria for APV30003 required that subjects had to have failed at least one PI-containing regimen but not more than two.

Genotyping of virus present in plasma samples was performed by Monogram Biosciences, Inc., CA, USA or Tibotec-Virco, Mechelen, Belgium, according to their standard procedures. Clinical trial participants were recruited between 2000 and 2002, and subjects were predominately infected with clade B HIV-1 (91%). For comparative purposes, a similar database of genotypes derived from 371 drug-naive subjects entering clinical studies APV30001 and APV30002 was used. The test HIV-1 protease sequences were imported from an in-house Oracle HIV-1 drug resistance database repository into Microsoft Excel 97 and processed by a Visual Basic for Applications program hosted by the Excel application. The program interrogated each residue in turn, and for each residue, it calculated both its consensus and non-consensus for all test sequences; the consensus sequence used was from the laboratory strain HIV-1HXB2. For each of the remaining 98 residues, it determined both their consensus and non-consensus totals, for instance, where the test residue was either consensus or non-consensus. Mixtures of consensus and mutant amino acids were reported as mutant. The resulting four aggregate counts for a given residue pair were used to create 2 by 2 contingency tables from which the program calculated both chi-squared values and associated P values. Positive associations were determined if the observed frequency was greater than the expected frequency and negative if the observed frequency was less than the expected frequency. Substitutions existing at low incidence (less than five) were excluded.

**Production of recombinant virus**

Site-directed mutagenesis was performed to introduce the K55R substitution into the HIV-1_HXB2 protease in combination with M46I/L±154V and V82A substitutions. The L24I substitution was not included due to lower incidence as a resistance mutation. In addition, a clinically derived protease gene encoding multiple substitutions including the K55R substitution was cloned (TOPO TA Cloning, Invitrogen) and mutated to remove the K55R substitution, but leave the PI resistance mutations intact (designated HPA1; for protease genotype, see Table 3 footnote). Mutant recombinant HIV-1 was obtained for all constructs generated, using the recombinant virus assay.15,16

To confirm the nucleotide sequences of the recombinant viruses, viral RNA was extracted from culture supernatants using guanidinium isothiocyanate lysis and amplified as described previously, omitting DMSO, and nucleotide sequences determined using ABI PRISM BigDye Terminator Reagent Kit, 3730 DNA Analyzer and SEQScape™ version 2.1 (Applied Biosystems). Mixtures of mutant and wild-type (WT) sequences with electropherogram peak size ratios >25% of the highest peak were recorded as mixed populations.

**Phenotyping of virus**

Drug susceptibility, phenotypic evaluation of the recombinant viruses, was performed with a panel of PIs using an MT4 cell viability-based assay.14,18 Fold change was defined as the change in the drug susceptibility between viruses with and without the amino acid substitution K55R.

**Viral replication studies**

An alternative role for minor PI-associated mutations is the restoration of viral replication capacity.19–21 Viral replicative capacity was therefore assessed using a competition-based approach between pairs of recombinant virus. Briefly, separate flasks of MT4 cells (1×10^5) were infected at a low multiplicity of infection (0.0001), incubated for 24 h and mixed in a range of input ratios. The cells were then co-cultured until peak cytopathic effect was reached.

Viral RNA was extracted from cell-free supernatant and the relative proportions of viruses were determined using Pyrosequencing™ Technology (Biotage AB, Uppsala, Sweden). RT–PCR was performed using a 5′ end biotinylated outer primer that facilitated purification of the 200 bp product using streptavidin beads. Sequencing primers were designed using Pyrosequencing™ software and terminated just prior to the single nucleotide polymorphism (SNP) to be detected. Pyrosequencing™ was performed on the PSQ 96 MA instrument (Biotage AG) and the results were analysed in the allele quotient mode to determine the percentage of each SNP present in the population.23

**Results**

Amino acid substitutions relative to HIV-1_HXB2 were observed at 63/99 amino acid positions in the HIV-1 protease in the drug-experienced database. Totally conserved positions included functionally important regions including those coding for the active site (codons 25–27), N- and C-terminus (codons 1–4 and 96–99).

Analysis of the protease database identified 102 pairwise associations that were considered highly significant (P < 0.05), adjusted for multiple testing using the Bonferroni correction. The 10 most significant positive associations are shown in Table 1 (positive associations column). These included associations between positions V32 and I47, and D30 and N88. Substitutions at V32I and I47V are associated with resistance to ampranavir,24 lopinavir25 and indinavir.26 Substitutions D30N and N88D are mutations associated with nelfinavir resistance.27 The 10 most significant negative associations (Table 1, negative associations column) included the pair of positions D30 and L90. Substitutions at these positions confer resistance to nelfinavir and are rarely found in the same viral genome.28

Two substitutions that were not previously considered to be common natural polymorphisms or positions associated with PI resistance with a known function were identified at positions 45 and 55 of the HIV-1 protease. Amino acid substitutions at position K45 were present in 4.9% (39/803) of sequences and were associated with substitutions at positions D30 (P = 2.6 × 10^-15) and N88 (P = 5.7 × 10^-15) of the HIV-1 protease gene. Mutations at position K45 have been found to be associated with PI treatment and with the protease mutations D30N and N88D.29

Substitutions at position K55 were observed in 6.1% (n = 49/803) of the sequences and were associated with substitutions at positions M46, L24, V82 and I54 (Table 1, K55 associations column). The K55R mutation was the most common amino acid change observed (46/49, 94%), which resulted from a single nucleotide change (AAG→AGA). Substitution at K55 was rarely detected in treatment-naive patients (6.1% experienced versus 0.3% naive). Furthermore, two of the three treatment-naive
patients harbouring virus with a K55R substitution also harboured PI-resistance-associated substitutions at codons 46, 54 and 82, indicating potential infection by drug-resistant virus.

To identify the amino acid substitution specifically associated with the protease mutation K55R, binomial correlation coefficients (phi) were determined. The K55R substitution was positively correlated with M46I (phi = 0.22) and M46L (phi = 0.19). Substitutions V82A (phi = 0.21) and V82S (phi = 0.15) were positively correlated with K55R, with V82F (phi = 3.1 \times 10^{-5}) and V82T (phi = 0.04) being more weakly associated. Of note, V82I, which is also a natural polymorphism, was negatively correlated (phi = -0.45) with the K55R substitution and the two were never observed together. Additionally, I54V (phi = 0.21) and L24I (phi = 0.27) were positively correlated with K55R.

The majority of sequences harbouring a K55R substitution also harboured an M46I (14%) or M46L (20%) substitution, with or without L24I, I54V and V82A/T/S (Table 2). The substitutions M46I/L and V82A/S/T are classified as major PI-associated mutations, whereas I54V and L24I are classified as minor PI-associated mutations.30 Virus from only two subjects harboured the K55R substitution in the absence of the amino acid changes L24I, M46I/L, I54V and V82A/T/S (genotypes L10I, L13V, K55R, L63P, H69K, A71V, V77I, L90M and K55R, L63P, A71T, V77I, L90M, I93L). Interestingly, the L24I, M46I/L, I54V and V82A/T/S substitutions were also observed in the absence of the K55R change (n = 20).

The K55R substitution was investigated further in order to establish the effect of this mutation on viral replication and/or phenotypic resistance. The K55R substitution conferred no or limited (2- to 3-fold) increases in phenotypic resistance to the PIs tested, as shown in Table 3. Although some reductions in susceptibility were observed in the co-presence of certain

<table>
<thead>
<tr>
<th>Positive associations (10 most significant)</th>
<th>Negative associations (10 most significant)</th>
<th>K55 associations (positive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pos 1</td>
<td>Pos 2</td>
<td>P value</td>
</tr>
<tr>
<td>V32</td>
<td>I47</td>
<td>1.0E-97</td>
</tr>
<tr>
<td>D30</td>
<td>N88</td>
<td>2.1E-82</td>
</tr>
<tr>
<td>I54</td>
<td>V82</td>
<td>7.9E-73</td>
</tr>
<tr>
<td>E35</td>
<td>M36</td>
<td>1.2E-57</td>
</tr>
<tr>
<td>L10</td>
<td>I54</td>
<td>1.6E-35</td>
</tr>
<tr>
<td>G73</td>
<td>L90</td>
<td>6.2E-32</td>
</tr>
<tr>
<td>H69</td>
<td>L89</td>
<td>3.7E-30</td>
</tr>
<tr>
<td>K20</td>
<td>M36</td>
<td>1.7E-27</td>
</tr>
<tr>
<td>A71</td>
<td>L90</td>
<td>8.7E-27</td>
</tr>
<tr>
<td>L10</td>
<td>M46</td>
<td>4.3E-26</td>
</tr>
</tbody>
</table>

Level of significance (P < 0.05) adjusted for multiple testing (Bonferroni correction).

Table 1. Summary of the most statistically significant associations for non-consensus pairs of HIV-1 positions in isolates from treated person

| Virus harbouring K55R mutation in the HIV-1 protease (5.7%; K55R n = 46) |
|------------------------|------------------------|------------------------|
| M46 (WT) none (n = 0/4) of the virus | M46I 14% (n = 30/213) of the virus | M46L 20% (n = 12/60) of the virus |
| M46 (WT) alone | L24I | V82A/F/S/T | IS4V | V82A, IS4V | V82A, IS4V, L24I |
| n = 2 | n = 0 | n = 0 | n = 0 | n = 2 | n = 0 |
| M46I alone | V82A | V82F | IS4V | V82A, IS4V | V82A, IS4V |
| n = 6 | n = 3 | n = 1 | n = 2 | n = 4 |
| M46L alone | L24I | V82A | IS4V | V82S, IS4V | V82A, IS4V |
| n = 0 | n = 4 | n = 0 | n = 1 | n = 1 |

Total presence of mutations at protease K55 in database 6.1% (K55R n = 46/49, K55N n = 3/49).

Amino acid not ascertained.
mutations, this was not observed consistently in the different protease backbones. Given the limited variability observed in the phenotypic assay using site-directed mutants, it is possible that K55R exerts a modest reduction in susceptibility in the presence of certain protease backgrounds.

When co-cultured, pairs of recombinant viruses harbouring PI mutations with or without the K55R mutation, a growth advantage for virus containing the K55R mutation was observed (Table 3). The largest percentage increases were observed when the K55R+M46L containing input virus was initiated at the lowest ratio (M46L+K55R, 82.4% increase and M46L+I54V+K55R+V82A, 83.7% increase). In contrast, no increase in viral replication was observed in the presence of K55R with the M46I mutation. However, when the M46I mutation was present with the additional protease mutations I54V and V82A either as site-directed mutations or included as part of protease mutations present in the clinical isolate, HPA1, a modest increase in replication (37.7% and 29.4%, respectively) was observed in addition to K55R.

### Discussion

This investigation into the association between pairs of non-consensus protease substitutions has identified the K55R mutation as a substitution significantly associated with recognized protease resistance mutations. Previously, Wu et al. identified 22 substitutions that were newly described treatment-associated mutations including K55R. In this analysis, only the K55R substitutions were identified. Wu et al. did not identify an association between M46I and K55R, but did not report the association of
K55R with L24I, M46L, I54V and V82A/S. Interestingly, Wu et al. also investigated clusters of residues and identified six clusters including positions L10, L24, M46, I54 and V82.

More recently, Svirch et al. also reported an association between K55R and L24I, M46L, I54V, L24I and V82A. The relative order of significance (the first being the most significant) of these associations, M46L > I54V > L24I > V82A, differed from the analysis reported here, M46L > L24I > V82A > I54 (see Table 1, K55 substitutions column). Additionally, Svirch et al. reported an association between K55R and substitutions L10I and R57K, albeit statistically less significant than the previous associations. Notably, the association between K55R and M46L was not described. The absence of detection of the M46L association and the differing order of significance most likely relates to subtle differences in the approach and composition of the database.

The binomial coefficients showed that several mutations were positively correlated with K55R. These included the protease substitutions M46L, V82A and I54V, which had similar phi values to those determined by Svirch et al.

The association of the K55R substitution with known protease substitutions prompted us to investigate the biological role of the K55R mutation. Current experiments indicate a role for the K55R substitution in enhancing viral replication in the presence of specific protease mutations and, in particular, the M46L mutation. The K55R substitution is positioned in the protease flap, close to M46L/I and I54V, but further away from codons 82 and 24.

Finally, the importance of the K55R substitution is highlighted by data from two recent clinical studies (RESIST 1 and 2) with the ritonavir-boosted HIV-1 PI tipranavir (Aptivus®). In this heavily treatment-experienced population, virological failure of tipranavir was associated with the development of the K55R substitution in at least 10% of the patients (Tipranavir US Package Insert).

In conclusion, three database analyses have now demonstrated an association between the K55R substitution and protease mutations associated with PI resistance. This study is the first to investigate the biological role of the K55R substitution and has demonstrated a role for this substitution in restoring viral replication capacity. This finding is consistent with the hypothesis that mutations resulting in decreased susceptibility also decrease protease activity, and that this activity is restored through the development of additional mutations outside the protease active site.

Acknowledgements

We thank N. Richards, P. Griffin, M. Tisdale, J. P. Klein and the GSK Clinical Virology Resistance Database Team for assistance during the project.

Funding

Funding was provided by an MRC CASE studentship award to E. S. M. and GSK.

Transparency declarations

None to declare.

References

Characterization of the HIV protease substitution K55R