

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Selected Amino Acid Mutations in HIV-1 B Subtype gp41 are Associated with Specific gp120v3 Signatures in the Regulation of Co-Receptor Usage

Retrovirology 2011, **8**:33 doi:10.1186/1742-4690-8-33

Salvatore Dimonte (salvatore.dimonte@uniroma2.it)

Fabio Mercurio (mercurio.fabio@gmail.com)

Valentina Svicher (valentina.svicher@uniroma2.it)

Roberta D'Arrigo (darrigo@inmi.it)

Carlo-Federico Perno (cf.perno@uniroma2.it)

Francesca Ceccherini-Silberstein (ceccherini@med.uniroma2.it)

ISSN 1742-4690

Article type Research

Submission date 18 October 2010

Acceptance date 12 May 2011

Publication date 12 May 2011

Article URL <http://www.retrovirology.com/content/8/1/33>

This peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in *Retrovirology* are listed in PubMed and archived at PubMed Central.

For information about publishing your research in *Retrovirology* or any BioMed Central journal, go to

<http://www.retrovirology.com/info/instructions/>

For information about other BioMed Central publications go to

<http://www.biomedcentral.com/>

Selected Amino Acid Mutations in HIV-1 B Subtype gp41 are Associated with Specific gp120_{v3} Signatures in the Regulation of Co-Receptor Usage

Salvatore Dimonte¹, Fabio Mercurio¹, Valentina Svicher¹, Roberta D'Arrigo², Carlo-Federico Perno^{1,2}, Francesca Ceccherini-Silberstein^{1,*}

¹ University of Rome Tor Vergata, Via Orazio Raimondo, Rome, Italy

² National Institute of Infectious Diseases (INMI) L. Spallanzani, Rome, Italy

* Corresponding author

Email addresses:

SD: salvatore.dimonte@uniroma2.it

FM: mercurio.fabio@gmail.com

VS: valentina.svicher@uniroma2.it

RD: darrigo@inmi.it

CFP: cf.perno@uniroma2.it

FCS: ceccherini@med.uniroma2.it

Abstract

Background: The third variable loop (V3) of the HIV-1 gp120 surface protein is a major determinant of cellular co-receptor binding. However, HIV-1 can also modulate its tropism through other regions in gp120, such as V1, V2 and C4 regions, as well as in the gp41 protein. Moreover, specific changes in gp41 are likely to be responsible for of damage in gp120-CCR5 interactions, resulting in potential resistance to CCR5 inhibitors.

In order to genetically characterize the two envelope viral proteins in terms of co-receptor usage, we have analyzed 526 full-length *env* sequences derived from HIV-1 subtype-B infected individuals, from our and public (Los Alamos) databases. The co-receptor usage was predicted by the analysis of V3 sequences using Geno2Pheno (G2P) algorithm. The binomial correlation phi coefficient was used to assess covariation among gp120_{V3} and gp41 mutations; subsequently the average linkage hierarchical agglomerative clustering was performed.

Results: According to G2P false positive rate (FPR) values, among 526 *env*-sequences analyzed, we further characterized 196 sequences: 105 with FPR <5% and 91 with FPR >70%, for X4-using and R5-using viruses, respectively.

Beyond the classical signatures at 11/25 V3 positions (S11S and E25D, R5-tropic viruses; S11KR and E25KRQ, X4-tropic viruses), other specific V3 and gp41 mutations were found statistically associated with the co-receptor usage. Almost all of these specific gp41 positions are exposed on the surface of the glycoprotein. By the covariation analysis, we found several statistically significant associations between V3 and gp41 mutations, especially in the context of CXCR4 viruses. The topology of the dendrogram showed the existence of a cluster associated with R5-usage involving E25D_{V3}, S11S_{V3}, T22A_{V3}, S129DQ_{gp41} and A96N_{gp41} signatures (bootstrap=0.88). Conversely, a large cluster was found associated with X4-usage involving T8I_{V3}, S11KR_{V3}, F20IVY_{V3}, G24EKR_{V3}, E25KR_{V3}, Q32KR_{V3}, A30T_{gp41}, A189S_{gp41}, N195K_{gp41} and L210P_{gp41} mutations (bootstrap=0.84).

Conclusions: Our results show that gp120_{v3} and several specific amino acid changes in gp41 are associated together with CXCR4 and/or CCR5 usage. These findings implement previous observations that determinants of tropism may reside outside the V3-loop, even in the gp41. Further studies will be needed to confirm the degree to which these gp41 mutations contribute directly to co-receptor use.

Background

Human immunodeficiency virus type 1 (HIV-1) entry into the host cell is mediated by the viral mature envelope (*env*) glycoproteins, gp120 and gp41, that constitute a trimeric complex anchored on the virion surface by the membrane-spanning segments of gp41 [1-4]. The gp120 exterior glycoprotein is retained on the trimer via labile, noncovalent interactions with the gp41 ectodomain [5], and it must be flexible to allow correct conformational modifications. The initial binding of gp120 to the cellular CD4 receptor indeed triggers conformational changes in gp120 that promote its following interaction with one of the chemokine co-receptors, usually CCR5 or CXCR4 [6-13]. This binding also induces the arrest of the transmembrane gp41 transitions at a prehairpin intermediate stage that leads to the insertion of the fusion peptide into the target cell membrane and ultimately to virus-cell fusion activity [14,15]. Multiple intermolecular contacts are required to maintain trimer integrity in gp120: the C1 and C5 region in gp120 are thought to be a provider to the gp120/gp41 interface and to the disulfide bond loop region of gp41, respectively [5,16-18].

HIV-1 strains can be phenotypically classified according to the virus' ability to use the CCR5 and/or CXCR4 co-receptor. Pure R5-tropic and pure X4-tropic viruses can use only the CCR5 and CXCR4 co-receptors to enter the target cell respectively, while the dual-tropic virus can use both co-receptors [19-23]. The binding to the chemokine receptor is based upon the presence of selected amino acids in

gp120 (specifically within the V3 loop, but also in other regions), providing greater affinity to CCR5 or CXCR4, and therefore the viral tropism [24-32].

It has been shown that R5-tropic viruses are generally responsible for the establishment of the initial infection, and they predominate in the majority of drug-naïve patients (prevalence, > 80%) [33-36]. However, in roughly 50% of all infected individuals, the virus changes its chemokine receptor usage during the progression of HIV-1 infection, due to the appearance of dual/mixed viruses [37-44]. Conversely, pure X4-tropic viruses are rare and occur in less than 1% of treatment-naïve patients and less than 5% of treated individuals, even at very late stages of the disease [33-36,45].

Based on the V3 location of the main genetic co-receptor usage determinants, the genotypic approaches for the tropism determination are so far based on sequencing and analyzing the V3 loop of gp120 with different algorithms available online [46,47].

However, emerging data clearly indicate the involvement of other gp120 regions in co-receptor binding, beyond the V3 loop (as V1, V2, and C4), and even that of the gp41 transmembrane protein [48-55]. Interestingly, recent studies have also shown that several mutations in gp41 were found to be significantly associated with co-receptor usage [48,54,56,57].

Therefore, due to the above mentioned reasons, the present investigation aims to genetically characterize HIV-1 B-subtype *env* sequences in terms of co-receptor usage and to define the association of mutations within the gp120 V3-region and the gp41 protein according to CCR5 and/or CXCR4 usage. For this purpose, we analyzed 526 HIV-1 subtype-B *env* sequences, only viral isolates from single patient, mostly retrieved from the Los Alamos database.

Methods

Sequence analysis

The analysis included 526 HIV-1 subtype-B *env* full-length sequences, partially retrieved from our

database (from 33 HIV-positive patients receiving highly active antiretroviral therapy), and the majority from the Los Alamos database [58] from 493 infected individuals at all stages of infection, with one isolate per single patient [58]. Sequences available with pure phenotype and/or co-receptor determinations have been considered, while molecular clone and dual-mix viruses have not been used. Published *env* consensus sequences of pure HIV-1 (A, B, C, D, F1, F2, G, H, J, and K) were used as reference for each subtypes [58], and multiple sequence alignments of V3 and gp41 segments were performed by using ClustalX [59] and were manually edited with the Bioedit software [60].

V3 and gp41 sequencing.

The sequencing of the V3 gp120 region and the entire gp41 was performed on 33 plasma samples, as described elsewhere [61,62]. In brief, for gp41 sequencing, RNA was extracted, retrotranscribed, and amplified by use of 2 different sequence-specific primers. Gp41-amplified products were full-length sequenced in sense and antisense orientations by use of 8 different overlapping sequence specific primers for an automated sequencer (ABI 3100; Applied Biosystems). Sequences with a mixture of wild-type and mutant residues at single positions were determined to have the mutant(s) at that position. Nucleotide sequences were previously submitted to Genbank [63].

For the sequencing of gp120 V3-domain, HIV-1 RNA was extracted, the V3-containing region of the *env* gene was then reverse-transcribed and amplified using the forward primer V3S2 5' CAGCACAGTACAATGTACACA 3' (nucleotide [nt]: 630-650 of HIV-1 HxB2 gp120 *env* gene) and the reverse primer V3AS5 5' CTTCTCCAATTGTCCCTCA 3' (nt: 1292-1310). The conditions for reverse transcription and amplification were: one cycle at 50°C for 30 min, one cycle 94°C for 2 min, 40 cycles (94°C 30s, 52°C 30s, 72°C 40s), and a final step at 72°C for 10 min, using the following reaction mix: 25 µl of RNA template, 8 µl of 5 mM Mg⁺⁺, 3 µl of Dnase Rnase free water, 0.75 µl of each primer at a concentration of 10 µM, 1 µl of Rnase out (40 U/µl), 1.5 µl of RT/Taq, 1 µl of dNTPs at a concentration of 10 mM for a total of 40 µl.

PCR-products were then sequenced by using the BigDye terminator v.3.1 cycle sequencing kit (Applied-Biosystems), and an automated sequencer (ABI-3100). Four different overlapping sequence-specific primers were used to ensure the coverage of the V3-sequence by at least two sequence segments. The sequencing conditions were: one cycle 96°C 3 min, 25 cycles (96°C 30s, 50°C 10s, 60°C 4 min) and the following primers were used: V3S6 5' CTGTTAAATGGCAGTCTAGC 3', V3S5 5' GTTAAATGGCAGTCTAGCAG 3', V3AS1 5' GAAAAATTCCCCTCCACAATT 3' and V3AS3bis 5' CAATTTCTGGGTCCCCTC 3'.

Subtypes were assessed by the construction of phylogenetic trees generated with the Kimura 2-parameter model. The statistical robustness within each phylogenetic tree was confirmed with a bootstrap analysis using 1000 replicates.

Tropism prediction

Within all 526 gp160-sequences, the V3 region was extrapolated and submitted for tropism prediction to Geno2Pheno algorithm. Geno2Pheno [46] is a bioinformatics tool based on support vector machines. Beyond tropism prediction, it assigns to each V3 sequence a score, called false positive rate (FPR), ranging from 0% to 100%, which represents the probability for a sequence to belong to an R5-virus. According to FPR values, we selected sequences with FPR < 5% (indicating a strong X4 prediction) and sequences with FPR > 70% (indicating a strong R5 prediction) for X4-tropic and R5-tropic viruses, respectively. These sequences, together with the related gp41 sequences, were then used for the entire study.

Statistical analysis.

To analyze gp41 and V3 mutations, we calculated the frequency of all mutations in the 345 gp41 amino acids and 35 V3 amino acids, using the *env* selected sequences. Fisher exact tests were used to

determine whether the differences in frequency between the 2 groups of patients were statistically significant (sequences with strong R5 and X4 prediction, respectively).

The Benjamini-Hochberg method has been used to identify results that were statistically significant in the presence of multiple-hypothesis testing [64]. A false discovery rate of 0.05 was used to determine statistical significance.

To identify significant patterns of pairwise associations between V3 and gp41 mutations, we calculated the ϕ coefficient and its statistical significance for each pair of mutations. A positive and statistically significant correlation between mutations at two specific positions ($0 < \phi < 1$; $P \leq 0.05$) indicates that the latter mutates in a correlated manner in order to confer an advantage in terms of co-receptor selection and that the co-occurrence of these mutations is not due to chance. Moreover, to analyze the covariation structure of mutations in more detail, we performed average linkage hierarchical agglomerative clustering, as described elsewhere [63,65]. Mann-Whitney U tests have been used to assess statistically significant differences among all the pairwise mutations associated. Statistical tests have been corrected for multiple-hypothesis testing by using the Benjamini-Hochberg method at a false discovery rate of 0.05 [64].

Results and Discussion

Prevalence of mutations

The study included 526 HIV-1 subtype-B *env* sequences, with the majority retrieved from the Los Alamos database. The V3 region was extrapolated from these gp160-sequences and submitted to the Geno2Pheno algorithm for tropism prediction.

Based on the FPR values, we selected 105 V3 sequences with FPR < 5% and 91 sequences with FPR > 70%, for their X4-using and R5-using co-receptor, respectively. These 196 sequences, together with the related gp41 sequences, were then used for the rest of the study.

As a first analysis, we confirmed in our dataset that the classical V3 positions 11 and 25 (consistent with previous observations [66-68]), wild-type amino acid at position 11, S11S, and E25D mutation were significantly associated with R5-tropic viruses, while mutations S11KR and E25KRQ were significantly associated with CXCR4 co-receptor usage (Figure 1a).

Since networks of V3 mutations are variable and complex, positions 11 and 25 are not sufficient to provide a full understanding of the mechanisms underlying different co-receptor usage. For example, it has been demonstrated that CCR5 interacts with the conserved V3 region encompassing the residues 4 to 7 (P4-N5-N6-N7) and the binding of this co-receptor is blocked when N7 is replaced by charged amino acid [30]. In our dataset, the mutation N7K has been found only in X4-predicted viruses (prevalence 9.5%; $P = 0.002$) (Figure 1a).

By evaluating the V3 loop sequence, we have identified 9 V3 mutations whose prevalence was significantly higher in the R5-predicted viruses than in the X4-predicted viruses ($P < 0.05$) (Figure 1a). Seven of them had a prevalence > 10% in R5-predicted viruses (the known E25D, and H13P, G15A, R18Q, F20L, Y21F and T22A). We also identified 33 mutations whose prevalence was significantly higher in X4- than in R5-viruses, suggesting their association with CXCR4-usage ($P < 0.05$). Among them, 17 had a prevalence > 10% in X4-predicted viruses (the known S11KR and E25KRQ, and I12V, H13ST, A19V, F20VY, Y21H, I27TV, Q32KR, H34Y), suggesting that within the V3 region, many more mutations are associated with CXCR4 usage (Figure 1a).

Interestingly, the majority of these V3 mutations found associated with the co-receptor usage were also recently found by our group as being involved in mechanisms underlying different co-receptor usage, using a completely different approach and dataset of isolates [68].

In addition, it is important to note that the selected dataset of sequences used in this study is small compared to the total number of sequences available in the Los Alamos database ; we also analyzed a different dataset of sequences with known phenotyping determination, composed by 326 and 91 V3-sequences (one HIV-1 B-subtype sequence/patient), with non-syncytium-inducing (NSI)- and syncytium-inducing (SI)-information, respectively.

Almost all statistically significant associations among V3 mutations and tropism found previously in the study were confirmed with this new analysis. The classical R5-tropic determinants S11S and E25D were found with high prevalence in NSI-sequences (73.6% and 64.1%, respectively, versus 34% and 11%, respectively, in SI-sequences; $P < 0.05$), while the classical X4-tropic mutations S11KR and E25KRQ were found with high prevalence in SI-sequences (40.6% and 51.6%, respectively, versus 2% and 11%, respectively, in NSI-sequences; $P < 0.05$). Moreover, the novel identified V3 mutations T22A in the R5-predicted viruses, and I12V, A19V, Y21H and H34Y in the X4-predicted viruses were also confirmed ($P < 0.05$).

The high variability of the V3 loop found in our study should not be surprising, since positive selection has been implicated in the maintenance of such diversity, in individuals as well as at the population level and in co-receptor selection [68-72]. It is likely that the principal driving force in the evolution of the V3 region of HIV-1 is the cell receptor usage, the escape from host immune response, or a combination of the two [73,74].

By analyzing the gp41 sequences, we found 35 out of 345 gp41 positions significantly associated with different co-receptor usage ($P < 0.05$) (Figure 1b). In particular, we identified 13 gp41 mutations whose prevalence was significantly higher in R5-using than in X4-using viruses: 7 of them had a prevalence $> 10\%$ in R5-predicted viruses (A69N, E110K, S129D, R209L, F241L, V267A, and I270T). Beyond these mutations, the wild type amino acid at 13 gp41 positions were also significantly associated with the R5-prediction (Q52Q, N126N, L134L, Q142Q, D153D, L181L, V190V, F206F, A212A, R250R, E280E, N287N and G314G) (Figure 1b).

Conversely, we identified 13 mutations whose prevalence was significantly higher in X4- than in R5- viruses, suggesting their association with the CXCR4-usage. Among them, 5 mutations had a prevalence > 10% in X4-predicted viruses (V69I, A96T, S129N, D163N and A189S) (Figure 1b).

Several gp41 residues associated with different co-receptor-usage reside within the Heptad Repeat 1 and 2 (HR1 and HR2) (A30, L34, Q52, D125, N126, S129, L134, N140, N141 and Q142), in the cluster I epitope transiently exposed during fusion (V69), and in the tryptophan-rich membrane-proximal external region (MPER) (D153 and D163). All these positions are localized in gp41 ectodomain known to be immunodominant and to induce high-titer antibodies in the majority of HIV-1-infected individuals [75-81]. The fact that all these mutations are localized in the extracellular domain of gp41 is consistent with the idea that gp41 may act as a scaffold in order to maintain the stability of the gp120/gp41 complex, and therefore finally influencing the viral tropism as well, directly or indirectly.

Association among mutations

By the analysis of associations between mutations, for the first time we found specific and statistically-significant correlations between V3 and gp41 mutations. In particular, several associations among mutations were associated with the CXCR4 prediction. An exception was represented by the A96N_{gp41} mutation that was positively correlated with T22A_{V3} ($\phi = 0.22$; $P = 0.030$; both associated with CCR5-usage) and negatively correlated with the known S11KR mutations ($\phi = -0.17$; $P = 0.018$). The A96N_{gp41} mutation is specifically localized in gp41 ectodomain and in particular within the cluster-I, that is a gp41 immunodominant loop involved in the interactions with gp120 [16,18,82-85].

Similarly, S129DQ_{gp41}, associated with CCR5-usage and localized in the gp41 HR2 domain, established negative correlation with the S11KR, strongly associated with CXCR4-usage, ($\phi = -0.21$;

$P = 0.041$) (Table 1). Notably, antibodies directed to the HR1/HR2 complex exist in the sera of HIV-1-infected individuals and this highlights the immunogenic character of the complex [75,86,87].

Regarding the positive correlations between V3 and gp41 mutations associated with CXCR4-usage, several were localized in the gp41 ectodomain (Table 1). In particular, a strong correlation was observed for A30T_{gp41} with either F20IVY_{V3} ($\phi = 0.38$; $P = 0.001$) or E25KRQ_{V3} ($\phi = 0.29$; $P = 0.006$) (Table 1). Of note, F20IVY_{V3} and E25KRQ_{V3} were found in 80% and 90% of patients with A30T_{gp41} respectively, thus further supporting that these mutations are highly correlated with each other. Another positive correlation was observed for L34M_{gp41} with N7KTY_{V3} (Table 1).

Interestingly, both A30T_{gp41} and L34M_{gp41} were also found recently associated phenotypically with CXCR4 usage [54,56,57]. Specifically, evaluating the available gp41 sequence data from samples submitted for co-receptor tropism testing by Trofile™, a CLIA-validated cell-based recombinant virus assay, Stawiski et colleagues have observed 26 gp41 mutations associated with CXCR4-use (Dual Mix/CXCR4), with the majority being on the extracellular region [56].

A30T_{gp41} and L34M_{gp41} are located in a specific region of HR1 involved in a direct interaction with gp120 [88]. In addition, the presence of A30T_{gp41} and L34M_{gp41} was observed in CXCR4-using isolates characterized by a high infectivity and/or replication capacity in CXCR4-expressing cells, thus supporting their involvement in the mechanism underlying CXCR4 usage [56,89,90]. Overall, this supports the role of these two mutations in the stabilization of non-covalently complex gp120/gp41, and/or in viral receptor attachment and membrane fusion.

Of note, we also found positive correlations between V3 mutations and gp41 mutations localized in the transmembrane domain or in the cytoplasmic tail of gp41. This is the case of A189S_{gp41}, localized in gp41 transmembrane domain, which correlated with Q32KR_{V3} ($\phi = 0.27$; $P = 0.021$). Both mutations were found positively associated with the CXCR4 prediction. Moreover, it has already been noted that Q32KR_{V3} could determine a reduction of gp120 binding affinity for the CCR5 N-

terminus, and this reduction is even stronger than that observed when positive charges are present at the classical V3 positions 11 and 25 [68].

Similarly, L210P_{gp41}, localized before the Kennedy sequence (that is a loop of the C-terminal tail of gp41 which is supposed to be exposed on the viral surface [91]), showed a strong correlation with G24EKR_{V3} ($\phi = 0.31$; $P = 0.019$).

The correlation between V3 and gp41 mutations was also confirmed by hierarchical clustering analysis. In particular, the topology of the dendrogram suggests the existence of a cluster associated with R5-usage and involving S11S, E25D, and T22A in the V3 and A96N and S129DQ in gp41 (bootstrap = 0.88) (Figure 2). Conversely, a large cluster was found associated with X4-usage. This involves the V3 mutations T8I, S11KR, F20IVY, G24EKR, E25KQR, Q32KR along with the gp41 mutations A30T, A189S, N195K, L210P (bootstrap = 0.84) (Figure 2).

Overall, our results suggest that specific additional gp41 mutations could be taken into account in order to implement the genotypic prediction algorithms currently in common use, as already demonstrated by Thielen and colleagues, who observed an improvement (albeit marginal) of CXCR4 co-receptor usage prediction [57]. In this work, it has been shown that mutations at N-terminus of gp41, such as A30T and L34M, are strongly associated with co-receptor phenotype in two independent datasets (444 and 1916 patients screened, respectively). The authors affirm that this region could theoretically be used to predict co-receptor use, alone or in combination with the V3 region. In our study, these 2 mutations, A30T and L34M, were both 100% associated to CXCR4-tropic viruses (Table 1).

It is conceivable that even mutations in gp41 may modulate co-receptor specificity and facilitate efficient CXCR4-mediated entry. This is consistent with other observations that showed that determinants of CXCR4 use in a set of dual-tropic *env* sequences, with V3 sequences identical to those of R5-tropic clones, mapped to the gp41 glycoprotein. Indeed, Huang et colleagues have shown that mutations in the fusion peptide and cytoplasmic tail of gp41 contribute to CXCR4 use by a dual-

tropic clone, while a single G515V mutation (according to HXB2 gp140 numbering) in gp41-fusion-peptide of another dual-tropic clone was sufficient to confer CXCR4 use to the R5-tropic original clone [48]. Similarly, the same authors reported previously that for HIV-1 subtype-D the V3 loop sequence of dual-tropic clones was identical to those of co-circulating R5-tropic clones, indicating the presence of CXCR4 tropism determinants also in domains different from V3 [41]. Interestingly, the threonine in position 96 that we find mutated in 72.5% of our viral X4-tropic B-subtype sequences (A96T_{gp41}) and negatively correlated with the R5-determinant T22A_{V3}, is the wild-type amino acid of gp41 in HIV-1 consensus sequence of subtype D viruses.

Based on crystal structures of HIV-1 gp41 so far available [92-95], the positions A30, L34, A96, S129 and N140 are all exposed on the surface of the glycoprotein (in HR1 or HR2 domains). Similarly, position L210 too, being near the epitopes for neutralizing antibodies, is presumably exposed on the surface glycoprotein [91]. Differently, the position of gp41 N195 seems to be located at the end of the classical single membrane spanning domain (172-198 amino acids), recently proposed to shuttle between two different conformations during the fusion process [96]. The same residue, based on another work [91], is part of an external loop of gp41 in an alternative membrane-spanning model, suggesting its alternating intra- and extra-membrane localization.

Consequently, we could speculate that gp41 A30T, L34M, A96NT, S129DQN, N140IT, N195K and L210P mutations may act together (directly or indirectly) with specific V3 signatures, via allosteric effects on the gp120/gp41 complex. This may allow the best conformational structural plasticity of gp41 and gp120 for their appropriate and specific binding to the cellular receptors and co-receptors. To support this hypothesis, the x-ray crystal structures of CD4-bound HIV-1 gp120 have revealed that the gp120 “core” consists of a gp41-interactive inner domain, a surface-exposed and heavily glycosylated outer domain and a conformationally flexible bridging sheet [14,30,97]. In addition, recent studies showed that in CD4-bound state two potentially flexible topological layers in the gp120 inner domain apparently contribute to the noncovalent association of gp120 with gp41 [98]

and insertions in V3 or polar substitutions in a conserved hydrophobic patch near the V3 of gp120 resulting in decreased gp120/gp41 association and decreased chemokine receptor binding [99].

With regard to the gp120-CD4 binding, it was found that the resulting conformational modifications protrude the V3 flexible loop to interact with the cellular co-receptor [29,97]. Interestingly, monoclonal antibodies directed against the D19 epitope within the V3 region had a neutralizing function only for the X4-tropic viruses, regardless of the presence of sCD4, while for R5 isolates only upon addition of sCD4 [100]. Consequently, the inaccessibility of this antibody to R5-tropic viruses in the absence of sCD4 might indicate that there are significant V3 loop conformational differences between these two viral variants [101], but also that specific interactions occurring in the gp120/gp41 complex may participate in the HIV-1 co-receptor usage and neutralization sensitivity.

Finally, we should mention that Anastassopoulou et colleagues have shown that viruses resistant to the small molecule CCR5 inhibitor, vicriviroc, can be caused by 3 conservative changes in the fusion peptide of HIV-1 gp41 [102], and similarly Pfaff *et al.*, very recently, found the involvement of gp120 and gp41 mutations in modulating the magnitude of drug resistance to another small CCR5 antagonist, aplaviroc [103]. Overall, these studies, which focus on changes toward resistances without assessing the issue of tropism-switch, are complementary to our results.

Conclusions

In this study, we found that specific gp41 mutations are significantly associated with different co-receptor usage and with specific V3 mutations, thus providing new information that could be taken into account for improving co-receptor usage prediction. These findings implement previous observations that determinants of tropism may reside outside the V3 loop, even in the gp41 transmembrane protein. It is possible that the gp120/gp41 complex may become structurally or functionally involved at different stages during virus-cell entry and fusion. Probably, the associations

among V3 and gp41 mutations may also have an impact on the HIV pathogenesis, it is known that CXCR4 phenotype has been associated with progression and increased severity of HIV disease, and several gp41 mutations are associated with viral fitness and cytopathic effects. Additional studies are needed to confirm the degree to which these gp41 mutations contribute directly to co-receptor use and to establish the specific and precise utility of this information.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SD and FM participated in the design of the study and performed the tropism prediction and statistical analysis. SD drafted the manuscript. RD was responsible for HIV-1 sequencing. VS, FCS and CFP participated in the study design and coordination and helped on writing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This work was financially supported by grants from the Italian Ministry of Instruction University & Research (MIUR), “Progetto FILAS”, and by the European Commission Framework 7 Programme (CHAIN, the Collaborative HIV and Anti-HIV Drug Resistance Network, Integrated Project no. 223131).

We are thankful for Amalia Mastrofrancesco, Marzia Romani and Laura Scipioni for their excellent technical assistance.

References

1. Chan DC, Fass D, Berger JM, Kim PS: **Core structure of gp41 from the HIV envelope glycoprotein.** *Cell* 1997, **89**:263-273.
2. Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, Wiley DC: **Atomic structure of the ectodomain from HIV-1 gp41.** *Nature* 1997, **387**:426-430.
3. Farzan M, Choe H, Desjardins E, Sun Y, Kuhn J, Cao J, Archambault D, Kolchinsky P, Koch M, Wyatt R, Sodroski J: **Stabilization of human immunodeficiency virus type 1 envelope glycoprotein trimers by disulfide bonds introduced into the gp41 glycoprotein ectodomain.** *J Virol* 1998, **72**:7620-7625.
4. Zhu P, Chertova E, Bess J Jr, Lifson JD, Arthur LO, Liu J, Taylor KA, Roux KH: **Electron tomography analysis of envelope glycoprotein trimers on HIV and simian immunodeficiency virus virions.** *Proc Natl Acad Sci USA* 2003, **100**:15812-15817.
5. Helseth E, Olshevsky U, Furman C, Sodroski J: **Human immunodeficiency virus type 1 gp120 envelope glycoprotein regions important for association with the gp41 transmembrane glycoprotein.** *J Virol* 1991, **65**:2119-2123.
6. Alkhatib G, Combadiere C, Broder CC, Feng Y, Kennedy PE, Murphy PM, Berger EA: **CCR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1.** *Science* 1996, **272**:1955-1958.
7. Choe H, Farzan M, Sun Y, Sullivan N, Rollins B, Ponath PD, Wu L, Mackay CR, LaRosa G, Newman W, Gerard N, Gerard C, Sodroski J: **The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates.** *Cell* 1996, **85**:1135-1148.
8. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhart M, Di Marzio P, Marmon S, Sutton RE, Hill CM, Davis CB, Peiper SC, Schall TJ, Littman DR, Landau NR: **Identification of a major co-receptor for primary isolates of HIV-1.** *Nature* 1996, **381**:661-666.

9. Doranz BJ, Rucker J, Yi Y, Smyth RJ, Samson M, Peiper SC, Parmentier M, Collman RG, Doms RW: **A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors.** *Cell* 1996, **85**:1149-1158.
10. Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, Cayanan C, Maddon PJ, Koup RA, Moore JP, Paxton WA: **HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5.** *Nature* 1996, **381**:667-673.
11. Feng Y, Broder CC, Kennedy PE, Berger EA: **HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor.** *Science* 1996, **272**:872-877.
12. Trkola A, Dragic T, Arthos J, Binley JM, Olson WC, Allaway GP, Cheng-Mayer C, Robinson J, Maddon PJ, Moore JP: **CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5.** *Nature* 1996, **384**:184-187.
13. Wu L, Gerard NP, Wyatt R, Choe H, Parolin C, Ruffing N, Borsetti A, Cardoso AA, Desjardin E, Newman W, Gerard C, Sodroski J: **CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5.** *Nature* 1996, **384**:179-183.
14. Wyatt R, Sodroski J: **The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens.** *Science* 1998, **280**:1884-1888.
15. Eckert DM, Kim PS: **Mechanisms of viral membrane fusion and its inhibition.** *Annu Rev Biochem* 2001, **70**:777-810.
16. York J, Nunberg JH: **Role of hydrophobic residues in the central ectodomain of gp41 in maintaining the association between human immunodeficiency virus type 1 envelope glycoprotein subunits gp120 and gp41.** *J Virol* 2004, **78**:4921-4926.
17. Jacobs A, Sen J, Rong L, Caffrey M: **Alanine Scanning Mutants of the HIV gp41 Loop.** *J Biol Chem* 2004, **280**:27284-27288.
18. Cao J, Bergeron L, Helseth E, Thali M, Repke H, Sodroski J: **Effects of amino acid changes**

- in the extracellular domain of the human immunodeficiency virus type 1 gp41 envelope glycoprotein.** *J Virol* 1993, **67**:2747-2755.
19. Berger EA: **HIV entry and tropism. When one receptor is not enough.** *Adv Exp Med Biol* 1998, **452**:151-157.
20. Loftin LM, Kienzle MF, Yi Y, Lee B, Lee FH, Gray L, Gorry PR, Collman RG: **Constrained use of CCR5 on CD4+ lymphocytes by R5X4 HIV-1: efficiency of Env-CCR5 interactions and low CCR5 expression determine a range of restricted CCR5-mediated entry.** *Virology* 2010, **402**:135-148.
21. Scarlatti G, Tresoldi E, Björndal A, Fredriksson R, Colognesi C, Deng HK, Malnati MS, Plebani A, Siccardi AG, Littman DR, Fenyö EM, Lusso P: **In vivo evolution of HIV-1 co-receptor usage and sensitivity to chemokine-mediated suppression.** *Nat Med* 1997, **3**:1259-1265.
22. Yi Y, Isaacs SN, Williams DA, Frank I, Schols D, De Clercq E, Kolson DL, Collman RG: **Role of CXCR4 in cell-cell fusion and infection of monocyte-derived macrophages by primary human immunodeficiency virus type 1 (HIV-1) strains: two distinct mechanisms of HIV-1 dual tropism.** *J Virol* 1999, **73**:7117-7125.
23. Yi Y, Shaheen F, Collman RG: **Preferential use of CXCR4 by R5X4 human immunodeficiency virus type 1 isolates for infection of primary lymphocytes.** *J Virol* 2005, **79**:1480-1486.
24. Cocchi F, DeVico AL, Garzino-Demo A, Cara A, Gallo RC, Lusso P: **The V3 domain of the HIV-1 gp120 envelope glycoprotein is critical for chemokine-mediated blockade of infection.** *Nat Med* 1996, **2**:1244-1247.
25. Hwang SS, Boyle TJ, Lyerly HK, Cullen BR: **Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1.** *Science* 1991, **253**:71-74.
26. Westervelt P, Gendelman HE, Ratner L: **Identification of a determinant within the human**

- immunodeficiency virus 1 surface envelope glycoprotein critical for productive infection of primary monocytes.** *Proc Natl Acad Sci USA* 1991, **88**:3097-3101.
27. O'Brien WA, Koyanagi Y, Namazie A, Zhao JQ, Diagne A, Idler K, Zack JA, Chen IS: **HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 outside the CD4-binding domain.** *Nature* 1990, **348**:69-73.
28. Hoffman TL, Doms RW: **HIV-1 envelope determinants for cell tropism and chemokine receptor use.** *Mol Membr Biol* 1999, **16**:57-65.
29. Hung CS, Vander Heyden N, Ratner L: **Analysis of the critical domain in the V3 loop of human immunodeficiency virus type 1 gp120 involved in CCR5 utilization.** *J Virol* 1999, **73**:8216-8226.
30. Huang CC, Lam SN, Acharya P, Tang M, Xiang SH, Hussan SS, Stanfield RL, Robinson J, Sodroski J, Wilson IA, Wyatt R, Bewley CA, Kwong PD: **Structures of the CCR5 N Terminus and of a Tyrosine-Sulfated Antibody with HIV-1 gp120 and CD4.** *Science* 2007, **317**:1930-1934.
31. Nabatov AA, Pollakis G, Linnemann T, Kliphuis A, Chalaby MI, Paxton WA: **Intrapatent alterations in the human immunodeficiency virus type 1 gp120 V1V2 and V3 regions differentially modulate co-receptor usage, virus inhibition by CC/CXC chemokines, soluble CD4, and the b12 and 2G12 monoclonal antibodies.** *J Virol* 2004, **78**:524-530.
32. Speck RF, Wehrly K, Platt EJ, Atchison RE, Charo IF, Kabat D, Chesebro B, Goldsmith MA: **Selective employment of chemokine receptors as human immunodeficiency virus type 1 coreceptors determined by individual amino acids within the envelope V3 loop.** *J Virol* 1997, **71**:7136-7139.
33. Brumme ZL, Goodrich J, Mayer HB, Brumme CJ, Henrick BM, Wynhoven B, Asselin JJ, Cheung PK, Hogg RS, Montaner JS, Harrigan PR: **Molecular and clinical epidemiology of**

- CXCR4-using HIV-1 in a large population of antiretroviral-naive individuals.** *J Infect Dis* 2005, **192**:466-474.
34. Melby T, Despirito M, Demasi R, Heilek-Snyder G, Greenberg ML, Graham N: **HIV-1 co-receptor use in triple-class treatment-experienced patients: baseline prevalence, correlates, and relationship to enfuvirtide response.** *J Infect Dis* 2006, **194**:238-246.
35. Moyle GJ, Wildfire A, Mandalia S, Mayer H, Goodrich J, Whitcomb J, Gazzard BG: **Epidemiology and predictive factors for chemokine receptor use in HIV-1 infection.** *J Infect Dis* 2005, **191**:866–872.
36. Wilkin TJ, Su Z, Kuritzkes DR, Hughes M, Flexner C, Gross R, Coakley E, Greaves W, Godfrey C, Skolnik PR, Timpone J, Rodriguez B, Gulick RM: **HIV type 1 chemokine co-receptor use among antiretroviral-experienced patients screened for a clinical trial of a CCR5 inhibitor: AIDS Clinical Trial Group A5211.** *Clin Infect Dis* 2007, **44**:591-595.
37. Boyd MT, Simpson GR, Cann AJ, Johnson MA, Weiss RA: **A single amino acid substitution in the V1 loop of human immunodeficiency virus type 1 gp120 alters cellular tropism.** *J Virol* 1993, **67**:3649-3652.
38. Schuitemaker H, Koot M, Kootstra NA, Dercksen MW, de Goede RE, van Steenwijk RP, Lange JM, Schattenkerk JK, Miedema F, Tersmette M: **Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: progression of disease is associated with a shift from monocytotropic to T-cell-tropic virus population.** *J Virol* 1992, **66**:1354-1360.
39. Regoes RR, Bonhoeffer S: **The HIV co-receptor switch: a population dynamical perspective.** *Trends Microbiol* 2005, **13**:269-277.
40. Church JD, Huang W, Mwatha A, Toma J, Stawiski E, Donnell D, Guay LA, Mmiro F, Musoke P, Jackson JB, Parkin N, Eshleman SH: **HIV-1 tropism and survival in vertically infected Ugandan infants.** *J Infect Dis* 2008, **197**:1382-1388.

41. Huang W, Eshleman SH, Toma J, Fransen S, Stawiski E, Paxinos EE, Whitcomb JM, Young AM, Donnell D, Mmiro F, Musoke P, Guay LA, Jackson JB, Parkin NT, Petropoulos CJ: **Coreceptor tropism in human immunodeficiency virus type 1 subtype D: high prevalence of CXCR4 tropism and heterogeneous composition of viral populations.** *J Virol* 2007, **81**:7885-7893.
42. Lihana RW, Khamadi SA, Lwembe RM, Kinyua JG, Muriuki JK, Lagat NJ, Okoth FA, Makokha EP, Songok EM: **HIV-1 subtype and viral tropism determination for evaluating antiretroviral therapy options: an analysis of archived Kenyan blood samples.** *BMC Infect Dis* 2009, **9**:215.
43. Moreno, Clotet, Sarría, Ortega, Leal, Rodriguez-Arrondo, Sánchez-de la Rosa; Allegro Study Group: **Prevalence of CCR5-tropic HIV-1 among treatment-experienced individuals in Spain.** *HIV Clin Trials* 2009, **10**:394-402.
44. Shepherd JC, Jacobson LP, Qiao W, Jamieson BD, Phair JP, Piazza P, Quinn TC, Margolick JB: **Emergence and persistence of CXCR4-tropic HIV-1 in a population of men from the multicenter AIDS cohort study.** *J Infect Dis* 2008, **198**:1104-1112.
45. Simon B, Grabmeier-Pfistershammer K, Rieger A, Sarcelletti, B. Schmied M, Puchhammer-Stöckl E: **HIV co-receptor tropism in antiretroviral treatment-naïve patients newly diagnosed at a late stage of HIV infection.** *AIDS* 2010, **24**:2051-2058.
46. **Genotypic prediction of coreceptor usage** [<http://coreceptor.bioinf.mpi-inf.mpg.de/index.php>].
47. **Genotypic prediction of coreceptor usage** [<http://fortinbras.us/cgi-bin/fssm/fssm.pl>].
48. Huang W, Toma J, Fransen S, Stawiski E, Reeves JD, Whitcomb JM, Parkin N, Petropoulos CJ: **Co-receptor Tropism Can Be Influenced by Amino Acid Substitutions in the gp41 Transmembrane Subunit of Human Immunodeficiency Virus Type 1 Envelope Protein.** *J Virol* 2008, **82**:5584-5593.

49. Suphaphiphat P, Essex M, Lee TH: **Mutations in the V3 stem versus the V3 crown and C4 region have different effects on the binding and fusion steps of human immunodeficiency virus type 1 gp120 interaction with the CCR5 coreceptor.** *Virology* 2007, **360**:182-190.
50. Koito A, Stamatatos L, Cheng-Mayer C: **Small amino acid sequence changes within the V2 domain can affect the function of a T-cell line-tropic human immunodeficiency virus type 1 envelope gp120.** *Virology* 1995, **206**:878-884.
51. Carrillo A, Ratner L: **Cooperative effects of the human immunodeficiency virus type 1 envelope variable loops V1 and V3 in mediating infectivity for T cells.** *J Virol* 1996, **70**:1310-1316.
52. Labrosse B, Treboute C, BreLOT A, Alizon M: **Cooperation of the V1/V2 and V3 domains of human immunodeficiency virus type 1 gp120 for interaction with the CXCR4 receptor.** *J Virol* 2001, **75**:5457-5464.
53. Pastore C, Nedellec R, Ramos A, Pontow S, Ratner L, Mosier DE: **Human immunodeficiency virus type 1 co-receptor switching: V1/V2 gain-of-fitness mutations compensate for V3 loss-of-fitness mutations.** *J Virol* 2006, **80**:750-758.
54. Thielen A, Altmann A, Bogojeska J, Kaiser R, Lengauer T: **Estimating evolutionary pathways to CXCR4 usage from cross-sectional data [abstract].** *Antivir Ther* 2009, **14**(Suppl 1):A16.
55. Thielen A, Sichtig N, Kaiser R, Lam J, Harrigan PR, Lengauer T : **Improved prediction of HIV-1 coreceptor usage with sequence information from the second hypervariable loop of gp120.** *J Infect Dis* 2010, **202**:1435-1443.
56. Stawiski E, Huang W, Whitcomb J, Petropoulos C, Coakley E: **Amino Acid Changes in gp41 of HIV-1 Associated with Co-receptor Tropism [abstract].** *Antivir Ther* 2009,

14(Suppl 1):A133.

57. Thielen A, Lengauer T, Harrigan PR, Swenson L, Dong W, McGovern RA, Lewis M, Heera J, Valdez H: **Mutation Within GP41 are Correlated With Co-Receptor Tropism but Do Not Substantially Improve Co-Receptor Usage Prediction.** In: *8th European HIV Drug Resistance Workshop, From basic science to clinical decision making*: 17-19 March 2010; Italy.
58. **Los Alamos HIV Sequence Database**
[<http://www.hiv.lanl.gov/components/sequence/HIV/search/search.html>].
59. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.** *Nucleic Acids Res* 1997, **25**:4876-4882.
60. Hall TA: **BioEdit: A user-friendly biological sequence alignment, editor and analysis program for Windows 95/98 NT.** In: *Nucl Acids Symp Ser* 1999, pp. 95-98 Vol 41.
61. Aquaro S, D'Arrigo R, Svicher V, Perri GD, Caputo SL, Visco-Comandini U, Santoro M, Bertoli A, Mazzotta F, Bonora S, Tozzi V, Bellagamba R, Zaccarelli M, Narciso P, Antinori A, Perno CF: **Specific mutations in HIV-1 gp41 are associated with immunological success in HIV-1-infected patients receiving enfuvirtide treatment.** *J Antimicrob Chemother* 2006, **58**:714-722.
62. Svicher V, Balestra E, Vandenbroucke I, Sarmati L, D'Arrigo R, Van Marck H, Pollicita M, Saccomandi P, Scopelliti F, Cammarota R, Di Santo F, Aquaro S, Stuyver L, Ceccherini-Silberstein F, Andreoni M, Perno C: **Ultradeep pyrosequencing and phenotypic analysis to characterize the V3 genetic diversity among HIV-1 primary isolates and their responses to maraviroc [abstract 77].** In: *7th European HIV Drug Resistance Workshop*: 25-27 March 2009; Sweden.
63. Svicher V, Aquaro S, D'Arrigo R, Artese A, Dimonte S, Alcaro S, Santoro MM, Di Perri G,

- Caputo SL, Bellagamba R, Zaccarelli M, Visco-Comandini U, Antinori A, Narciso P, Ceccherini-Silberstein F, Perno CF: **Specific enfuvirtide-associated mutational pathways in HIV-1 Gp41 are significantly correlated with an increase in CD4(+) cell count, despite virological failure.** *J Infect Dis* 2008, **197**:1408-1418.
64. Benjamini Y, Hochberg Y: *Controlling The False Discovery Rate: A Practical and Useful Approach to Multiple Testing.* Journal of Royal Statistical Society. Series B. Vol. 57, No. 1 1995.
65. Svicher V, Alteri C, D'Arrigo R, Laganà A, Trignetti M, Lo Caputo S, Callegaro AP, Maggiolo F, Mazzotta F, Ferro A, Dimonte S, Aquaro S, di Perri G, Bonora S, Tommasi C, Trotta MP, Narciso P, Antinori A, Perno CF, Ceccherini-Silberstein F: **Treatment with the fusion inhibitor enfuvirtide influences the appearance of mutations in the human immunodeficiency virus type 1 regulatory protein rev.** *Antimicrob Agents Chemother* 2009, **53**:2816-2823.
66. Fouchier RA, Groenink M, Kootstra NA, Tersmette M, Huisman HG, Miedema F, Schuitemaker H: **Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule.** *J Virol* 1992, **66**:3183-3187.
67. De Jong JJ, De Ronde A, Keulen W, Tersmette M, Goudsmit J: **Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution.** *J Virol* 1992, **66**:6777-6780.
68. Svicher V, Cammarota R, Artese A, D'Arrigo R, Parisi S, Zazzi M, Antinori A, Angarano G, Nozza S, Perno CF, Oscar Study Group: **New V3-genetic Signatures Modulate Co-receptor Usage in vivo and the Interaction with CCR5 N-terminus.** In: *Program and Abstracts of the Seventeenth Conference on Retroviruses and Opportunistic Infections San*

Francisco, CA: 16-19 February 2010; Foundation for Retrovirology and Human Health, Alexandria, VA, USA.

69. Leal E, Janini M, Diaz RS: **Selective pressures of human immunodeficiency virus type 1 (HIV-1) during pediatric infection.** *Infect Genet Evol* 2007, **7**:694-707.
70. Lemey P, Kosakovsky Pond SL, Drummond AJ, Pybus OG, Shapiro B, Barroso H, Taveira N, Rambaut A: **Synonymous substitution rates predict HIV disease progression as a result of underlying replication dynamics.** *PLoS Comput Biol* 2007, **3**:e29.
71. Shankarappa R, Margolick JB, Gange SJ, Rodrigo AG, Upchurch D, Farzadegan H, Gupta P, Rinaldo CR, Learn GH, He X, Huang XL, Mullins JI: **Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection.** *J Virol* 1999, **73**:10489-10502.
72. Yang W, Bielawski JP, Yang Z: **Widespread adaptive evolution in the human immunodeficiency virus type 1 genome.** *J Mol Evol* 2003, **57**:212-221.
73. Ross HA, Rodrigo AG: **Immune-mediated positive selection drives human immunodeficiency virus type 1 molecular variation and predicts disease duration.** *J Virol* 2002, **76**:11715-11720.
74. Williamson S: **Adaptation in the env gene of HIV-1 and evolutionary theories of disease progression.** *Mol Biol Evol* 2003, **20**:1318-1325.
75. Xu JY, Gorny MK, Palker T, Karwowska S, Zolla-Pazner S: **Epitope mapping of two immunodominant domains of gp41, the transmembrane protein of human immunodeficiency virus type 1, using ten human monoclonal antibodies.** *J Virol* 1991, **65**:4832-4838.
76. Dimmock NJ: **The complex antigenicity of a small external region of the C-terminal tail of the HIV-1 gp41 envelope protein: a lesson in epitope analysis.** *Rev Med Virol* 2005, **15**:365-381.

77. Montero M, van Houten NE, Wang X, Scott JK: **The membrane-proximal external region of the human immunodeficiency virus type 1 envelope: dominant site of antibody neutralization and target for vaccine design.** *Microbiol Mol Biol Rev* 2008, **72**:54-84.
78. Hrin R, Montgomery DL, Wang F, Condra JH, An Z, Strohl WR, Bianchi E, Pessi A, Joyce JG, Wang YJ: **Short communication: In vitro synergy between peptides or neutralizing antibodies targeting the N- and C-terminal heptad repeats of HIV Type 1 gp41.** *AIDS Res Hum Retroviruses* 2008, **24**:1537-1544.
79. Cleveland SM, McLain L, Cheung L, Jones TD, Hollier M, Dimmock NJ: **A region of the C-terminal tail of the gp41 envelope glycoprotein of human immunodeficiency virus type 1 contains a neutralizing epitope: evidence for its exposure on the surface of the virion.** *J Gen Virol* 2003, **84(Pt 3)**:591-602.
80. Cheung L, McLain L, Hollier MJ, Reading SA, Dimmock NJ: **Part of the C-terminal tail of the envelope gp41 transmembrane glycoprotein of human immunodeficiency virus type 1 is exposed on the surface of infected cells and is involved in virus-mediated cell fusion.** *J Gen Virol* 2005, **86(Pt 1)**:131-138.
81. Prabakaran P, Dimitrov AS, Fouts TR, Dimitrov DS: **Structure and function of the HIV envelope glycoprotein as entry mediator, vaccine immunogen, and target for inhibitors.** *Adv Pharmacol* 2007, **55**:33-97.
82. Merat R, Raoul H, Leste-Lasserre T, Sonigo P, Pancino G: **Variable constraints on the principal immunodominant domain of the transmembrane glycoprotein of human immunodeficiency virus type 1.** *J Virol* 1999, **73**:5698-5706.
83. Binley JM, Sanders RW, Clas B, Schuelke N, Master A, Guo Y, Kajumo F, Anselma DJ, Maddon PJ, Olson WC, Moore JP: **A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated**

- structure.** *J Virol* 2000, **74**:627-643.
84. Maerz AL, Drummer HE, Wilson KA, Pombourios P: **Functional analysis of the disulfide-bonded loop/chain reversal region of human immunodeficiency virus type 1 gp41 reveals a critical role in gp120-gp41 association.** *J Virol* 2001, **75**:6635-6644.
85. Quintana FJ, Gerber D, Kent SC, Cohen IR, Shai Y: **HIV-1 fusion peptide targets the TCR and inhibits antigen-specific T cell activation.** *J Clin Invest* 2005, **115**:2149-2158.
86. Chen CH, Greenberg ML, Bolognesi DP, Matthews TJ: **Monoclonal antibodies that bind to the core of fusion-competent gp41.** *AIDS Res Hum Retroviruses* 2000, **16**:2037-2041.
87. Vincent N, Kone A, Chanut B, Lucht F, Genin C, Malvoisin E: **Antibodies purified from sera of HIV-1-infected patients by affinity on the heptad repeat region 1/heptad repeat region 2 complex of gp41 neutralize HIV-1 primary isolates.** *AIDS* 2008, **22**:2075-2085.
88. Park EJ, Quinnan GV Jr: **Both neutralization resistance and high infectivity phenotypes are caused by mutations of interacting residues in the human immunodeficiency virus type 1 gp41 leucine zipper and the gp120 receptor- and coreceptor-binding domains.** *J Virol* 1999, **73**:5707-5713.
89. Rodrigo AG: **Dynamics of syncytium-inducing and non-syncytium-inducing type 1 human immunodeficiency viruses during primary infection.** *AIDS Res Hum Retroviruses* 1997, **13**:1447-1451.
90. Leavitt M, Park EJ, Sidorov IA, Dimitrov DS, Quinnan GV Jr: **Concordant modulation of neutralization resistance and high infectivity of the primary human immunodeficiency virus type 1 MN strain and definition of a potential gp41 binding site in gp120.** *J Virol* 2003, **77**:560-570.
91. Hollier MJ, Dimmock NJ: **The C-terminal tail of the gp41 transmembrane envelope glycoprotein of HIV-1 clades A, B, C, and D may exist in two conformations: an analysis of sequence, structure, and function.** *Virology* 2005, **337**:284-296.

92. Tan K, Liu J, Wang J, Shen S, Lu M: **Atomic structure of a thermostable subdomain of HIV-1 gp41.** *Proc Natl Acad Sci U S A* 1997, **94**:12303-12308.
93. Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, Wiley DC: **Atomic structure of the ectodomain from HIV-1 gp41.** *Nature* 1997, **387**:426-430.
94. Chan DC, Fass D, Berger JM, Kim PS: **Core structure of gp41 from the HIV envelope glycoprotein.** *Cell* 1997, **8**:263-273.
95. Shi W, Bohon J, Han DP, Habte H, Qin Y, Cho MW, Chance MR: **Structural characterization of HIV gp41 with the membrane-proximal external region.** *J Biol Chem* 2010, **285**:24290-24298.
96. Gangupomu VK, Abrams CF: **All-atom models of the membrane-spanning domain of HIV-1 gp41 from metadynamics.** *Biophys J* 2010, **99**:3438-3444.
97. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA: **Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody.** *Nature* 1998, **393**:648-659.
98. Finzi A, Xiang SH, Pacheco B, Wang L, Haight J, Kassa A, Danek B, Pancera M, Kwong PD, Sodroski J: **Topological layers in the HIV-1 gp120 inner domain regulate gp41 interaction and CD4-triggered conformational transitions.** *Mol Cell* 2010, **37**:656-667.
99. Xiang SH, Finzi A, Pacheco B, Alexander K, Yuan W, Rizzuto C, Huang CC, Kwong PD, Sodroski J: **A V3 loop-dependent gp120 element disrupted by CD4 binding stabilizes the human immunodeficiency virus envelope glycoprotein trimer.** *J Virol* 2010, **84**:3147-3161.
100. Lusso P, Earl PL, Sironi F, Santoro F, Ripamonti C, Scarlatti G, Longhi R, Berger EA, Burastero SE: **Cryptic nature of a conserved, CD4-inducible V3 loop neutralization epitope in the native envelope glycoprotein oligomer of CCR5-restricted, but not CXCR4-using, primary human immunodeficiency virus type 1 strains.** *J Virol* 2005,

79:6957-6968.

101. Sander O, Sing T, Sommer I, Low AJ, Cheung PK, Harrigan PR, Lengauer T, Domingues FS: **Structural descriptors of gp120 V3 loop for the prediction of HIV-1 co-receptor usage.** *PLoS Comput Biol* 2007, **3:e58.**
102. Anastassopoulou CG, Ketas TJ, Klasse PJ, Moore JP: **Resistance to CCR5 inhibitors caused by sequence changes in the fusion peptide of HIV-1 gp41.** *Proc Natl Acad Sci USA* 2009, **106:5318-5323.**
103. Pfaff JM, Wilen CB, Harrison JE, Demarest JF, Lee B, Doms RW, Tilton JC: **HIV-1 resistance to CCR5 antagonists associated with highly efficient use of CCR5 and altered tropism on primary CD4+ T cells.** *J Virol* 2010, **84:6505-6514.**

Figure legends

Figure 1

Frequencies of HIV-1 gp120_{V3} and gp41 mutations.

Frequencies of gp120_{V3} (panel “a”) and gp41 (panel “b”) mutations in HIV-1 R5-tropic isolates with FPR > 70% by Geno2Pheno-algorithm prediction (dark grey) and HIV-1 X4-tropic isolates with FPR < 5% by Geno2Pheno-algorithm prediction (light grey). Statistically significant differences were assessed by chi-square tests of independence. *P* values were significant at a false-discovery rate of 0.05 following correction for multiple tests. *, *P* < 0.05; **, *P* ≤ 0.01; ***, *P* ≤ 0.001.

Figure 2

Clusters of correlated mutations.

Dendrogram obtained from average linkage hierarchical agglomerative clustering, showing significant clusters involving V3 and gp41 (gray box) mutations. The length of branches reflects distances between mutations in the original distance matrix. Bootstrap values, indicating the significance of clusters, are reported in the boxes. The analysis was performed in sequences derived from 196 patients, 91 reported as R5-tropic and 105 reported as X4-tropic at genotypic test.

Table

Table 1. Novel gp41 mutations significantly associated with gp120_{v3} mutations.

gp41 mutations	Frequency no. (%) of isolates ^a	Frequency % in X4-tropic viruses ^b	Correlated mutations	Frequency no. (%) of isolates ^a	Covariation frequency no. (%) of isolates ^c	ϕ ^d	P ^e
A30T _{gp41}	10 (5.1)	100	F20IVY _{v3}	34 (17.3)	8 (80.0)	0.38	0.001
			E25KQR _{v3}	62 (31.6)	9 (90.0)	0.29	0.006
			S11S _{v3}	109 (55.6)	0 (0)	-0.26	0.009
L34M _{gp41}	5 (2.5)	100	N7KTY _{v3}	15 (7.6)	3 (60.0)	0.32	0.055
			T22A _{v3}	113 (57.6)	23 (85.2)	0.22	0.030
A96N _{gp41}	27 (13.8)	29.6	S11KR _{v3}	51 (26.0)	3 (11.1)	-0.17	0.018
			N140IT _{gp41}	22 (11.2)	12 (23.5)	0.23	0.054
A96T _{gp41}	51 (26.0)	72.5	T22A _{v3}	113 (57.6)	19 (37.2)	-0.24	0.022

S129DQ_{gp41}	43 (20.9)	26.8	S11KR_{v3}	51 (26.0)	4 (9.8)	-0.21	0.041
S129N_{gp41}	24 (12.2)	75.0	I12MV_{v3}	19 (9.7)	7 (29.2)	0.26	0.041
N140IT_{gp41}			N7KTY_{v3}	15 (7.6)	6 (27.3)	0.26	0.046
	22 (11.2)	86.4	A96T_{gp41}	51 (26.0)	12 (54.5)	0.23	0.054
			S11S_{v3}	109 (55.6)	5 (22.7)	-0.24	0.028
A189S_{gp41}	13 (6.6)	92.3	Q32KR_{v3}	50 (25.5)	9 (69.2)	0.27	0.021
N195K_{gp41}			T8I_{v3}	8 (4.1)	3 (50.0)	0.41	0.022
	6 (3.1)	100	S11KR_{v3}	51 (26.0)	6 (100)	0.16	0.041
L210P_{gp41}	12 (6.1)	83.3	G24EKR_{v3}	23 (11.7)	6 (50.0)	0.31	0.019

^a Frequency was determined in 196 isolates from HIV-1 infected patients having FPR < 5% and FPR > 70%, using the Geno2Pheno algorithm.

^b Frequency was determined in 105 HIV-1 isolates reported as X4-tropic at genotypic test (FPR < 5%).

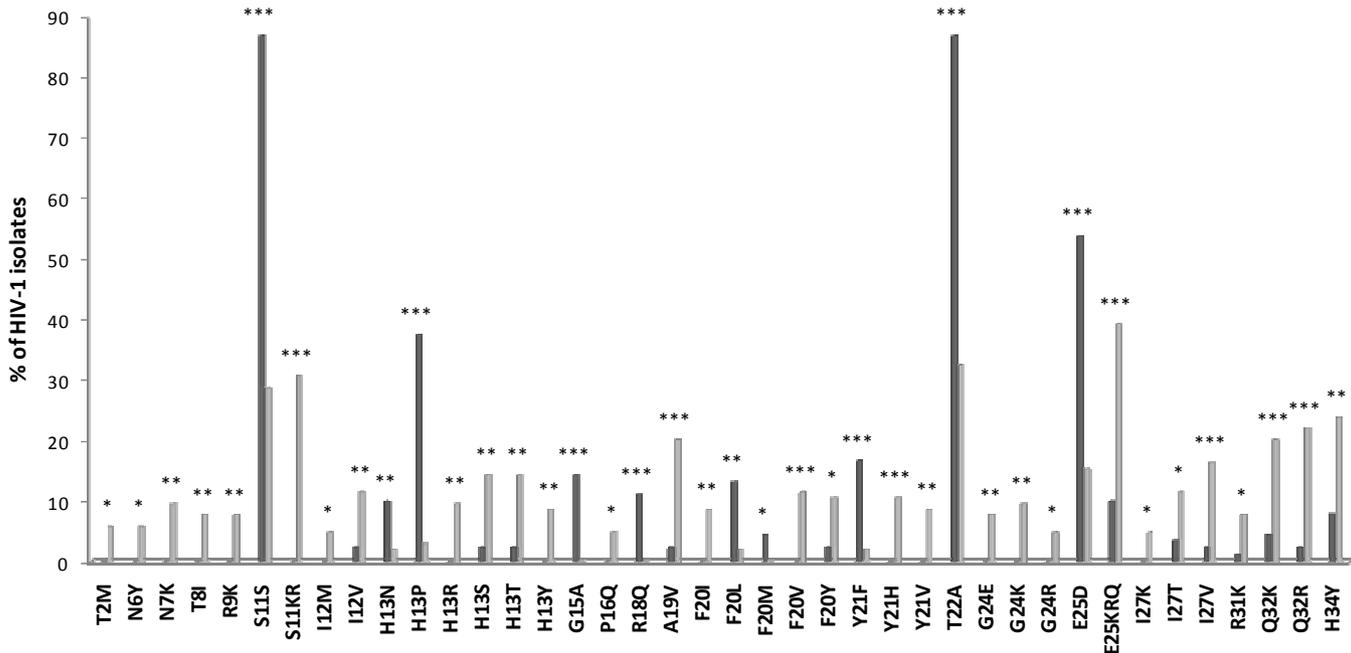
^c Percentages were calculated in patients with each specific gp41 mutation.

^d Positive and negative correlations with $\varphi > 0.15$ and $\varphi < -0.15$, respectively, are shown.

^e *P* values significant ($P \leq 0.05$) after correction for multiple hypothesis testing [65].

■ R5-using viruses, N=91 ■ X4-using viruses, N=105

a



■ R5-using viruses, N=91 ■ X4-using viruses, N=105

b

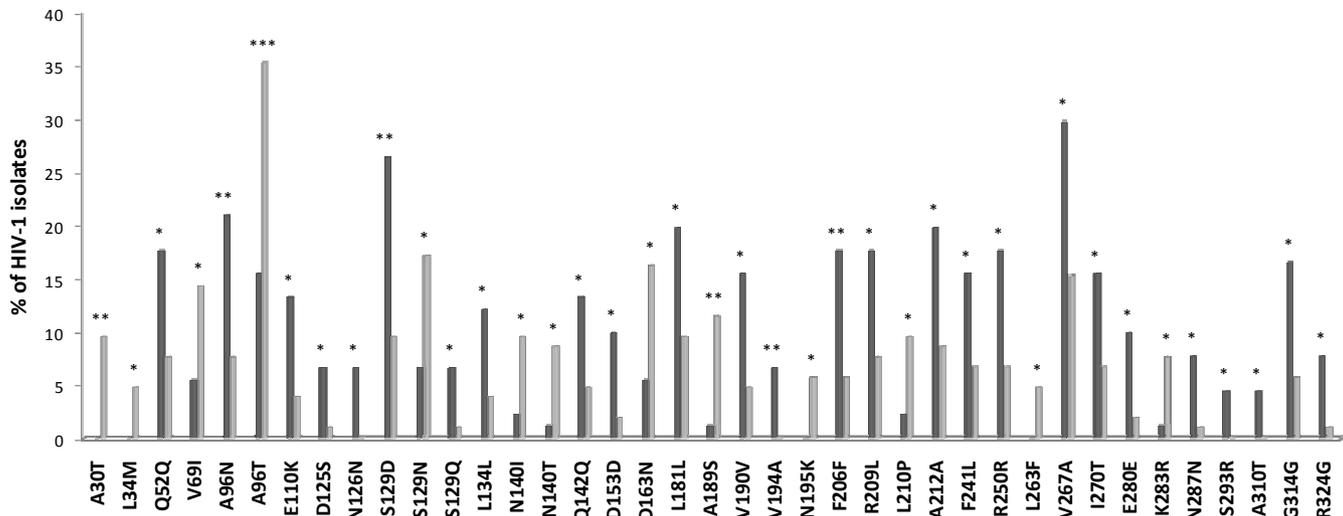


Figure 1

