ABSTRACT

HIV-1 gains entry into host cells by binding to CD4 and a coreceptor, predominantly CCR5 or CXCR4. Viruses that use CCR5 are termed R5, those able to use CXCR4 are termed X4 while viruses able to use both coreceptors are referred to as R5X4. Accelerated CD4 decline and disease progression within an infected HIV-1 subtype B infected individual is often associated with the emergence of viruses able to use CXCR4. However, CXCR4 coreceptor usage appears to occur less frequently among HIV-1 subtype C viruses, the most predominant strain circulating globally, including South Africa. The aim of this study was to investigate the genetic determinants of CXCR4 usage in HIV-1 subtype C isolates.

The V3 region of the envelope glycoprotein is the major determinant of coreceptor usage. In Chapter 2, 32 subtype C isolates with known phenotypes (16 R5, 8 R5X4 and 8 X4 isolates) were assessed using a subtype C specific V3-heteroduplex tracking assay. Results indicated that there were sufficient genetic differences to discriminate between R5 viruses and those able to use CXCR4 (both R5X4 and X4). In general, R5 isolates had a mobility ratio >0.9 whereas CXCR4-using isolates were usually <0.9. Sequence analysis of the V3 region showed that CXCR4-using viruses were often associated with an increased positive amino acid charge, insertions and loss of a glycosylation site, similar to HIV-1 subtype B. In contrast, where subtype B consensus V3 has a GPGR crown motif irrespective of coreceptor usage, all 16 subtype C R5 viruses had a conserved GPGQ sequence at the tip of the loop, while 12 of the 16 (75%) CXCR4-using viruses had substitutions in this motif, most commonly arginine (R). Thus, the rare occurrence of CXCR4-using viruses in subtype C may be due to the highly conserved nature of the GPGQ crown that may limit the potential for the development of X4 viruses.

The usefulness of available genotype-based methods for predicting viral phenotypes in subtype C was explored in Chapter 3. Results indicated that commonly used prediction methods could detect R5 viruses, but were not very sensitive at identifying X4 viruses.

We therefore developed a subtype C specific predictor based on position specific scoring matrices (PSSM). Similar methodology, as used in developing the subtype B PSSM, was applied on a training set of 280 subtype C sequences of known phenotype (229 NSI/CCR5 and 51 SI/CXCR4). The C-PSSM had a specificity of 94% (C.I. [92%-96%]) and sensitivity of 75% (C.I. [68%-82%]), indicating that the C-PSSM had improved sensitivity in predicting CXCR4 usage. This method also highlighted amino acid positions within V3 that could contribute differentially to phenotype prediction in subtypes B and C. A reliable phenotype prediction method, such as the C-PSSM, could provide a rapid and less expensive approach to identifying CXCR4 variants, and thus increase our knowledge of subtype C coreceptor usage.

In Chapter 4 we examined the genetic changes in full-length gp160 envelope genes of 23 sequential isolates from 5 patients followed for two to three years. Three of the patients' isolates used CCR5 at all time points while 2 patients underwent a coreceptor switch with disease progression. The genetic changes observed over time indicated changes in length of variable loops particularly the V1, V4 and V5 and shifting N-glycosylation sites, particularly in the 2 patients that used CXCR4. Changes in the V3 were only noted in the 2 patients' that used CXCR4 which included substitutions of specific amino acids including those in the crown and increased amino acid charge in the V3 region. Both of these patients were dually infected suggesting that recombination may contribute to the rapid emergence of X4 viruses.

The *in vitro* and *in vivo* development of CXCR4 usage was analysed in a pediatric patient that experienced a coreceptor switch during disease progression (Chapter 5). Biological and molecular clones were generated and the V1-V5 regions sequenced. Analyses of the V3 region indicated that the evolution to CXCR4 usage happens in a step-wise manner that included increased charge and changes in the crown motif. The intermediate variants with predicted dualtropism were also associated with increased V1-V2 lengths, suggesting that other regions may contribute to coreceptor switching. Furthermore, the development of CXCR4 usage within this patient was due to two mutational pathways, in which one resulted in R5X4 viruses and the other X4 variants.

In Chapter 6, the impact and treatment of acute TB on HIV-1 diversity in co-infected patients was investigated, specifically to determine the genetic characteristics of the viral populations present before, during and after TB treatment. Plasma samples from 18 HIV-1 infected patients were analysed using the C2V3 region, six of whom showed a high degree of variation using a V3-HTA and were selected for further analyses. All patients were predicted as R5 with no evidence of coreceptor switching over time. There was no correlation between the degree of genetic diversity and viral load, although both showed fluctuations over time. Phylogenetic and pairwise genetic distance analysis indicated that there was amplification of existing variants in 3 patients while in the other 3 patients there were dramatic shifts in viral populations suggesting selection of viral sub-populations over time. Thus in some co-infected patients, TB can affect HIV-1 genetic heterogeneity although there was no evidence of a shift towards CXCR4 usage despite the presence of an AIDS defining illness.

Observations in this study have shown that the V3 region is the major determinant of coreceptor usage within HIV-1 subtype C, similar to HIV-1 subtype B. Characteristics such as increased charge length variability of the V3 region and loss of the glycosylation site within this region are associated with CXCR4 usage. The limited number of X4 viruses in subtype C does suggest some restricting mechanisms for CXCR4 usage. In this study we looked at genetic determinants and found that the rare occurrence of CXCR4-using viruses in subtype C, may be due to the highly conserved nature of the GPGQ crown that may limit the potential for the development of subtype C X4 viruses. Furthermore, the development of CXCR4 usage happened in a step-wise manner, with R5X4 viruses intermediates, in which an increased V1-V2 was observed suggesting that other regions within the envelope protein do contribute to coreceptor usage. Thus, regions such as V1-V2 and V4-V5 did contribute to coreceptor usage, but the V3 region remained the most important determinant of coreceptor usage in HIV-1 subtype C isolates. Collectively these findings have provided important data on the genetic determinants of CXCR4 usage in HIV-1 subtype C and an understanding of how they might evolve within a patient.