

Original Research Article

Co receptor Usage of Human Immunodeficiency Virus Type 1 Strains in Ibadan, Nigeria

ABSTRACT

Aims To evaluate HIV-1 Coreceptor usage in Nigeria.

Study design: Cross-sectional study.

Place and Duration of Study: Department of Virology, College of Medicine, University College Hospital, Ibadan from January 2004 to December 2006.

Methodology: Genomic DNA was extracted from blood samples of 85 (42 males, 43 females; age range 18-58 years) consenting HIV-1 infected antiretroviral therapy-naïve individuals presenting at a voluntary counseling and testing centre. The *env* C2-V3 region of HIV-1 proviral DNA was amplified by nested PCR, successfully sequenced, manually edited and evolutionary relationships determined by maximum likelihood using MEGA 5.03^R from 64 of the blood samples of the HIV-1 infected patients. HIV-1 coreceptor usage was predicted based on genotypic analysis of HIV-1 *env* V3 loop sequences.

Results: Phylogenetic analysis showed HIV-1 subtypes A, G, CRF02_AG, CRF06_cpx, and CRF35_AD among the study participants. The V3 loop region of the viruses had amino acid sequence conservation in the base positions 1-8 and 26-35 and tip regions and sequence variability, including mutations and deletions at positions 9-25. Most (76.6%) of the sequences had the GPGQ crown motif while the GPGQ/L/R/K substitution was observed in 18.8%. The number of N-linked glycosylation sites ranged from 0 to 4 per *env* C2-V3 amino acid sequence with only 37.5% of the sequences having all 4 N-linked glycosylation sites. Predicted frequencies for CCR5 and CXCR4 genotypes were 31.2% and 68.8%, respectively, while 10% of the CCR5-tropic viruses showed Maraviroc-associated resistant mutations.

Conclusion: CXCR4-tropic viruses predominate among the studied population irrespective of HIV-1 subtype and it is associated with multiple amino acid deletions and mutations in V3 and the loss of one or more N-linked glycosylation sites. This data suggest the need for further studies involving a larger sample size prior to introduction of coreceptor inhibitors like Maraviroc for management of HIV infection in Nigeria..

Keywords: Coreceptor usage, HIV-1 strains, CCR5, CXCR4, C2-V3 region, V3 loop

1. INTRODUCTION

Early in infection, HIV-1 generally uses the CCR5 chemokine coreceptor, along with CD4, for cellular entry [1]. In many HIV-1 infected individuals, viral genotypic changes arise that allow the virus to use CXCR4, either in addition to CCR5 or alone, as an entry receptor [2] and this change is associated with accelerated CD4 decline and more rapid progression of HIV-1 disease [3]. Based on chemokine coreceptor usage, HIV-1 viruses can be classified as CCR5-tropic, CXCR4-tropic or dual-tropic (R5X4) viruses [4]. The V3 loop of gp120 largely determines coreceptor usage and the development of CXCR4 viruses is gradual and involves the accumulation of multiple amino acid changes in the V3 loop [5, 6].

With the introduction of HIV-1 coreceptor inhibitors as components of antiretroviral therapy [7], it is increasingly imperative to screen HIV positive individuals before they commence therapy with coreceptor inhibitors like Maraviroc. Maraviroc is an HIV-1 coreceptor antagonist that has shown good efficacy and tolerability in treatment-naïve and treatment-experienced patients harboring CCR5-tropic virus [8]. The use of Maraviroc in treatment simplification in patients requires analysis of HIV-1 DNA.

However, there is little information regarding bioinformatics prediction or analysis of HIV-1 coreceptor usage in Nigeriawith the only reported data by Ajoge and colleagues [9]. Yet this information is important for drug design and clinical management of patients before coreceptor inhibitors are used for treatment. There is also growing evidence of increasing viral resistance (primary and secondary – fuelled by both virus mutation and non-adherence issues) to commonly used anti-retroviral drugs [10-12] available in the country and there remains a possibility that CCR5-inhibitors like Maraviroc, can be used for HIV-1 antiretroviral therapy. This study was therefore carried out to evaluate virus coreceptor usage by bioinformatics method based on the V3 loop sequences of HIV-1 from HIV-1 infected antiretroviral therapy-naïve individuals.

2. METHODOLOGY

2.1 Study location

This study was carried out in the Department of Virology, College of Medicine, University College Hospital, Ibadan. The department offers HIV counseling and testing services as well as monitoring of Anti-Retroviral Therapy (ART) response in HIV infected patients. The department is also the Center for the WHO National Polio laboratory, HIV reference laboratory and WHO Collaborating Center for Influenza surveillance. Patients enrolled for the study were mostly from Ibadan and other locations in the southwestern region of Nigeria.

2.2 Study population

The study population consisted of 85 consenting HIV-1 infected anti-retroviral therapy-naïve individuals. They included 42 males and 43 females with median age of 37 years (range 18-58 years) who presented at the HIV counseling and testing centre in the Department of Virology, College of Medicine, University College Hospital, Ibadan from January 2004 to December 2006. The University of Ibadan/UCH ethical review board approved the study protocol and written informed consent was obtained from every individual whose blood sample was used for the study.

2.3 Sample collection, preparation and storage

Blood sample was collected by venepuncture from each patient into sterile vacutainer blood collection tubes containing EDTA (ethylene diamine tetra acetic acid) as coagulant. Each specimen was labelled with the date of collection and laboratory identification number. Plasma was separated from each sample and both plasma and packed cells were stored at -20°C until analysed. Only samples that were positive by Western blot for HIV-1 antibodies were analysed further for the study.

2.4 Laboratory analysis

2.4.1 Serology

Initial HIV screening was by Genscreen Ultra HIV Ag-Ab (BIORAD, France) and confirmation was by the Western blot assay using New Lav Blot 1 (BIORAD, France). The results were interpreted according to the manufacturer's instructions.

2.4.2 Molecular analysis

75 Genomic DNA was extracted from whole blood using the QiaAmp DNA Blood Mini kit
76 (Qiagen, Maryland, USA) according to manufacturer's instruction. The *env* C2-V3 region of
77 HIV-1 from the genomic DNA extract was amplified by nested PCR method using primers
78 WT1 and WT2, and KK40 and KK30 for the first and second rounds, respectively[13].
79 Amplified DNA was directly sequenced with the second round primers in the Genetic
80 Analyzer 3130xl (Applied Biosystems, California, USA) in both forward and reverse
81 directions using Sequencing kit v3.1 (Applied Biosystems, California, USA). The sequences
82 were manually edited using the sequencing software v3.1 (Applied Biosystem, California,
83 USA) and MEGA 5.03, and analysed for phylogenetic relationships with reference
84 sequences from the HIV database (Los Alamos, New Mexico, USA) by ClustalW and
85 Maximum Likelihood methods in MEGA 5.03 software based on their nucleotide
86 (approximately 350bp) and translated amino acid sequences[14].
87

88 **2.4.3 Virus phenotype determination**

89 The virus coreceptor usage was predicted by combined criteria using a genotypic rule based
90 on one of the following criteria for predicting CXCR4 (R4) coreceptor usage (i) Arginine (R)
91 or Lysine (K) at position +11 of V3 and/K at position +25, (ii) R at position +25 of V3 and a
92 net charge $\geq +5$, or (iii) a net charge of $\geq +6$ or (iv) bioinformatic analysis by comparison
93 between Cons B and query V3 loop sequence in Geno2Pheno (10% FPR) which uses the
94 presence of X4 specific mutations in the V3 loop as a predictor of X4 phenotype[15-19].
95

96 **3. RESULTS AND DISCUSSION**

98 **3.1 HIV-1 subtype**

99 The target region of approximately 350bp was successfully sequenced from 64 (75.3%) of
100 the 85 HIV-1 infected antiretroviral therapy-naïve individuals. The 64 HIV-1 infected
101 individuals comprised of 34 males and 30 females with a median age of 37 years (range 18-
102 58 years). Overall, 3.1%, 53.1%, 28.1%, 14.1% and 1.6% of the HIV-1 *env* C2-V3 nucleotide
103 sequences were identified as subtype A, G, CRF02_AG, CRF06_cpx, and CRF35_AD,
104 respectively [Figure 1].

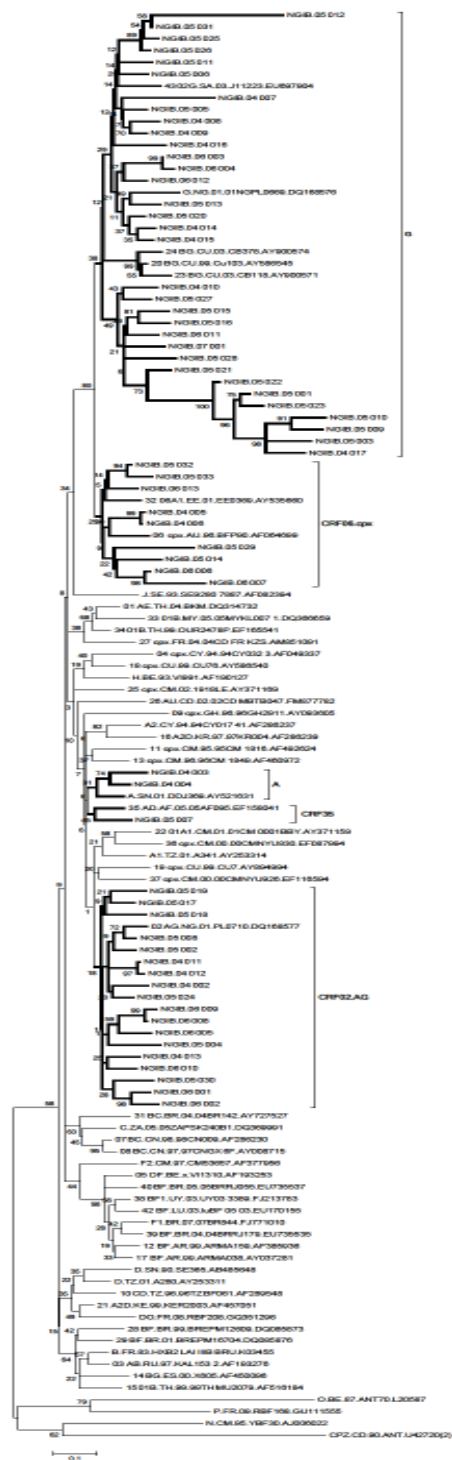


FIGURE 1: PHYLOGENETIC DISTRIBUTION OF HIV-1 STRAINS AMONG HIV-1 POSITIVE INDIVIDUALS

111 **3.2 V3 loop sequence analysis**

112 The sequence analysis of the V3 loop showed total conservation for 3 (P16, G17, and A33)
113 out of the 35 amino acid residues (8.6% of conservation). P16 and G17 are motifs within the
114 GPGQ motif of the V3 tip and A33 is in the base. The positions in which variation did not
115 exceed 10% were concentrated in the base (R3-4.7%, P4-7.8%, I30-4.7%, R31-7.8%), the
116 tip (G15-4.7%) and the stem (N6-3.1%, I26-6.3%). The most variable position was the 25th,
117 which was the determinant for the 11/25 rule used for predicting HIV-1 coreceptor usage.
118 The variation in the Cysteine residues limiting the V3 loop was 3.1% for C1 and 1.6% for
119 C35 [Figure 2].

120

121 The conserved GPGQ crown motif was the most common sequence observed at the V3 tip
122 occurring in 76.6% (49) of the translated HIV-1 V3 loop amino acid sequences. There was
123 substitution from GPGQ to other V3 crown motifs like GPGL (8), GPGR (3), GPGK (1),
124 APGQ (2) and RPGQ (1), as shown in figure 2.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	Subtype	Net Charge	Geno2Pheno (10%)	Coreceptor		
Cons B	C	T	R	P	N	N	N	T	-	R	K	S	I	H	I	G	P	G	R	A	F	Y	T	T	G	E	I	I	G	D	I	R	Q	A	H	C					
NGIB.04.001	C	T	R	P	G	N	N	T	-	R	R	-	Y	K	N	G	P	G	Q	S	F	Y	A	T	G	D	I	I	G	D	T	R	Q	A	H	C	G	3	R5	R5	
NGIB.04.002	C	T	R	P	N	N	N	T	-	R	K	C	-	T	Y	G	P	G	Q	T	I	Y	A	T	G	A	I	I	G	E	I	R	Q	A	H	C	CRF_02_AG	3	X4	X4	
NGIB.04.003	C	T	R	P	N	N	N	T	-	R	K	C	K	N	-	G	P	G	Q	A	F	Y	A	T	G	A	I	T	G	D	I	I	K	A	Y	C	A	4	R5	R5	
NGIB.04.004	C	T	R	P	N	N	N	T	-	R	K	C	N	N	-	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	T	R	K	A	H	C	A	3	X4	X4	
NGIB.04.005	C	Y	R	P	N	N	N	T	-	R	K	Y	I	-	I	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	CRF_06_cpx	2	R5	R5	
NGIB.04.006	C	Y	R	P	N	N	N	T	-	R	K	Y	I	-	I	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	CRF_06_cpx	2	R5	R5	
NGIB.04.007	Y	-	R	P	N	-	-	T	-	-	R	K	Y	K	N	G	P	G	Q	T	F	Y	A	T	-	-	-	I	G	N	I	R	Q	A	-	C	G	6	X4	X4	
NGIB.04.008	C	T	R	P	S	N	N	T	-	R	R	-	Y	K	N	G	P	G	Q	A	F	Y	A	T	G	A	V	T	G	D	I	R	N	A	H	C	G	4	R5	R5	
NGIB.04.009	C	T	R	P	N	N	N	T	-	-	R	K	Y	R	N	G	P	G	Q	A	F	Y	A	T	G	A	I	I	G	D	I	R	Q	A	H	C	G	4	X4	X4	
NGIB.04.010	C	T	R	P	N	N	N	T	-	R	K	H	T	P	-	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	E	A	H	C	G	1	R5	R5	
NGIB.04.011	C	S	R	P	G	N	N	T	-	-	R	Q	Y	T	Y	G	P	G	Q	A	F	Y	A	T	G	D	I	T	G	D	I	R	Q	A	H	C	CRF_02_AG	1	R5	R5	
NGIB.04.012	C	S	R	P	G	N	N	T	-	R	R	-	Y	T	Y	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	CRF_02_AG	2	X4	X4	
NGIB.04.013	C	T	R	P	G	N	N	T	-	-	R	K	Y	T	Y	G	P	G	Q	V	F	Y	T	D	-	-	-	I	I	G	D	I	R	Q	A	H	C	CRF_02_AG	3	X4	X4
NGIB.04.014	C	T	R	P	N	N	N	T	-	R	K	H	K	N	-	G	P	G	Q	M	F	Y	A	T	G	D	I	T	G	D	I	R	Q	A	H	C	G	3	R5	R5	
NGIB.04.015	C	T	R	P	N	N	N	T	-	-	R	K	Y	K	I	-	G	P	G	Q	T	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	Y	C	G	3	R5	X4
NGIB.04.016	C	T	R	A	N	N	N	T	-	-	R	K	A	N	K	N	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	K	Q	A	H	C	G	3	R5	X4
NGIB.04.017	C	T	T	P	N	N	Y	T	-	K	K	D	I	-	I	G	P	G	L	A	F	R	A	K	G	N	I	I	G	D	I	R	Q	A	H	C	G	3	X4	X4	
NGIB.05.001	C	T	R	P	N	N	Y	T	-	K	K	D	I	-	I	G	P	G	L	A	F	R	A	K	G	N	I	I	G	D	I	R	Q	A	H	C	G	4	X4	X4	
NGIB.05.002	C	T	R	P	G	N	N	T	-	-	R	K	D	I	Y	G	P	G	Q	T	F	Y	A	T	G	D	I	V	G	D	I	R	Q	A	H	C	CRF_02_AG	2	R5	X4	
NGIB.05.003	C	T	R	P	N	N	Y	T	-	K	K	D	I	-	I	G	P	G	L	A	F	R	A	K	G	T	I	I	G	D	I	K	Q	A	H	W	G	4	X4	X4	
NGIB.05.004	C	T	R	L	G	N	N	T	-	R	R	-	Y	T	Y	G	P	G	R	A	F	F	T	S	G	D	V	T	G	D	I	R	K	A	Y	C	CRF_02_AG	4	X4	X4	
NGIB.05.005	C	T	R	P	G	N	N	T	-	-	I	R	Y	K	N	G	P	G	Q	A	F	Y	A	T	G	K	I	T	G	D	I	R	K	A	H	C	G	5	X4	X4	
NGIB.05.006	C	T	R	P	N	N	N	T	-	R	R	-	Y	N	I	A	P	G	Q	A	F	Y	T	T	G	A	I	I	G	D	I	R	Q	A	Y	C	G	3	R5	R5	
NGIB.05.007	C	T	R	P	N	N	N	T	-	R	K	C	-	T	Y	G	P	G	Q	S	F	F	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	CRF_35_AD	2	R5	R5	
NGIB.05.008	C	T	R	P	N	N	N	T	-	R	K	C	-	T	Y	G	P	G	Q	T	F	Y	A	T	G	D	I	I	G	N	I	R	Q	A	H	C	CRF_02_AG	3	R5	R5	
NGIB.05.009	C	T	R	P	N	N	Y	T	-	K	K	D	I	-	I	G	P	G	L	A	F	R	A	K	G	N	I	I	G	D	I	R	Q	A	H	C	G	4	X4	X4	
NGIB.05.010	C	T	R	P	N	N	Y	T	-	K	K	D	I	-	I	G	P	G	L	A	F	P	A	K	G	N	I	I	G	D	I	R	Q	A	H	C	G	3	X4	X4	
NGIB.05.011	C	T	R	P	G	N	N	T	-	-	R	K	D	I	-	I	G	P	G	A	I	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	G	2	X4	X4	
NGIB.05.012	W	T	R	P	S	I	N	T	-	-	R	I	Y	R	D	G	P	G	Q	A	V	S	A	T	V	S	I	L	V	N	L	-	I	A	D	C	G	1	X4	X4	
NGIB.05.013	C	I	R	P	N	N	N	T	-	R	K	H	K	N	-	G	P	G	Q	T	F	Y	T	T	G	A	I	I	G	N	I	R	Q	A	H	C	G	5	R5	R5	
NGIB.05.014	C	I	R	P	G	N	N	T	-	K	R	Y	I	-	L	A	P	G	Q	V	F	Y	A	T	-	-	-	I	I	G	D	I	R	Q	A	H	C	CRF_06_cpx	3	X4	X4
NGIB.05.015	C	T	R	P	N	N	N	T	-	-	R	K	Y	T	T	G	P	G	Q	A	I	Y	A	T	G	A	I	I	G	D	I	R	Q	A	H	C	G	3	X4	X4	
NGIB.05.016	C	S	R	P	N	N	N	T	-	-	R	K	Y	T	T	G	P	G	Q	A	I	Y	A	T	G	A	I	I	G	D	I	R	Q	A	H	C	G	3	X4	X4	
NGIB.05.017	C	T	R	P	G	N	N	T	-	R	K	C	A	Y	-	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	K	A	H	C	CRF_02_AG	3	X4	X4	
NGIB.05.018	C	T	R	P	N	N	N	T	-	V	Q	G	-	A	C	G	P	G	Q	T	F	Y	A	K	D	D	I	I	G	D	I	R	Q	A	H	C	CRF_02_AG	0	R5	R5	
NGIB.05.019	C	I	R	P	S	N	N	T	-	-	R	K	F	S	Y	G	P	G	L	K	F	Y	T	T	G	N	I	I	G	D	I	R	K	A	H	C	CRF_02_AG	5	X4	X4	
NGIB.05.020	C	T	R	P	G	N	N	T	-	-	R	K	Y	K	N	G	P	G	Q	V	F	Y	A	T	G	E	I	I	G	D	I	R	Q	A	Y	C	G	3	X4	X4	
NGIB.05.021	C	T	R	P	N	N	N	T	-	R	K	Y	I	-	I	G	P	G	Q	A	L	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	G	2	R5	R5	
NGIB.05.022	C	T	R	P	N	N	Y	T	K	K	K	D	I	-	I	G	P	G	L	A	F	R	A	K	G	N	I	I	G	D	I	R	Q	A	H	C	G	5	X4	X4	
NGIB.05.023	C	T	G	P	G	N	N	T	-	K	K	D	I	-	L	A	P	G	Q	T	F	F	A	T	G	N	I	I	G	D	I	R	Q	A	H	C	G	3	X4	X4	
NGIB.05.024	C	T	R	P	N	N	N	T	-	R	K	H	A	Y	-	G	P	G	Q	T	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	CRF_02_AG	2	X4	X4	
NGIB.05.025	C	T	G	P	H	N	N	T	-	R	R	-	Y	K	D	G	P	G	Q	A	F	Y	A	T	G	E	I	I	G	N	I	R	Q	A	H	C	G	2	R5	R5	
NGIB.05.026	C	T	R	P	N	N	N	T	-	R	R	-	Y	K	D	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	N	I	R	Q	A	H	C	G	3	R5	R5	
NGIB.05.027	C	T	R	P	N	N	N	T	-	-	R	K	Y	N	I	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	Y	C	G	2	R5	X4	
NGIB.05.028	C	T	R	P	G	N	N	T	-	R	K	D	-	D	I	G	P	G	R	V	F	Y	A	T	S	D	V	V	G	D	I	R	Q	A	H	C	G	1	X4	X4	
NGIB.05.029	C	T	R	P	N	N	N	T	-	R	K	Y	I	P	-	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	R	A	Y	C	CRF_06_cpx	3	R5	R5	
NGIB.05.030	C	T	R	P	D	N	N	T	-	-	R	Q	C	A	Y	G	P	G	Q	V	F	Y	A	N	K	D	I	I	G	N	I	R	Q	A	H	C	CRF_02_AG	2	X4	X4	
NGIB.05.031	C	T	R	P	N	N	N	T	-	R	R	-	Y	R	D	G	P	G	Q	A	F	Y	A	T	G	A	I	V	G	N	I	R	Q	A	H	C	G	4	X4	X4	
NGIB.05.032	C	T	R	P	N	N	N	T	-	R	K	Y	T	H	-	G	P	G	Q	A	F	Y	A	T	G	A	I	I	G	N	I	R	Q	A	H	C	CRF_06_cpx	4	R5	R5	
NGIB.05.033	C	T	R	P	N	N	N	T	-	R	R	Y	T	H	-	G	P	G	R	V	F	Y	A	T	G	D	I	I	G	D	I	R	R	A	H	C	CRF_06_cpx	4	X4	X4	
NGIB.06.001	C	S	R	P	S	N	N	T	-	R	R	D	-	R	Y	G	P	G	Q	M	F	Y	A	A	G	E	I	I	G	D	I	R	Q	A	H	C	CRF_02_AG	2	X4	X4	
NGIB.06.002	C	T	R	P	S	N	N	T	-	R	K	D	-	R	Y	G	P	G	Q	M	F	Y	A	A	G	D	I	T	G	D											

Most potential N-linked glycosylation sites were conserved in the immediate vicinity of and within the V3 loop of the HIV-1 sequences analyzed. The NxS/Ty sequon for the potential N-glycosylation site centered at the 5'-Cys V3 loop residue was conserved in 46 (71.9%) of the HIV-1 C2-V3 amino acid sequences with 41 (89.1%) presenting with the NCT (Asparagine-Cysteine-Threonine) sequon [Figure 2]. Potential N-linked glycosylation sites located six amino acids downstream from the 5' cysteine residue in the V3 loop was conserved in 61 (95.3%) of HIV-1 infected individuals with 54 (88.5%) presenting with the NNT (Asparagine-Asparagine-Threonine) sequon.

Interestingly, one individual had no potential N-linked glycosylation site in the HIV-1 C2-V3 amino acid sequence analyzed. This sequence (NGIB.04__007) is a subtype G HIV-1 virus having the GPGQ V3 tip motif with a predicted X4 phenotype.

3.3 Prediction of Coreceptor usage

Using the 4 combined criteria, 68.8% (44) of V3 loop amino acid sequences had predicted X4 virus phenotype while the remainder (31.2%) were CCR5-tropic. Of these 44 V3 loop amino acid sequences with CXCR4 tropism, 7 V3 loops were classified as X4 phenotype based wholly on the 11/25 rule while the remaining 37 were predicted based on a combination of the other 3 criteria.

On further analysis of the V3 loop amino acid sequences that met each criterion, 16 of the V3 loop amino acid sequences had Lysine residue at position +11 while only one had Arginine residue at position +11. The only V3 loop sequence with the Arginine residue at position +11 also had a lysine residue at position +25. One V3 loop amino acid sequence had a net charge of +6. Bioinformatics analysis with Geno2Pheno at 10% false positive rate (FPR) predicted 38 (including some V3 loop sequences that met the other criteria) as X4 viruses. All V3 loop sequences with 33 and 31 amino acid residues, and those not limited by the Cysteine residues, were predicted as X4 virus phenotype. In addition, HIV-1 subtypes G and CRF02_AG had a higher prevalence of CXCR4 virus phenotypes compared to HIV-1 CRF06_cpx in the study population [Table 1]

TABLE 1. FREQUENCY OF CORECEPTOR USAGE AMONG HIV-1 SUBTYPES

	Subtype A	Subtype G	CRF02_AG	CRF06_cpx	TOTAL
Subtype	(%)	(%)	(%)	(%)	(%)
Total	2	34	18	9	63
X4 Phenotype	1(50.0)	25(73.5)	14(77.8)	4(44.4)	44(68.8)
R5 Phenotypy	1(50.0)	9(26.5)	4(22.2)	5(55.6)	20(31.2)

161 DISCUSSION.

162

163 The increasing interest in HIV-1 tropism is related to the introduction of CCR5 inhibitors as
164 components of antiretroviral therapy[7]. Although, the “gold standard” for characterization of
165 HIV-1 tropism is a recombinant virus phenotypic entry assay, genotypic methods based on
166 the V3 amino acid sequence analysis and characteristics have been successful with different
167 HIV-1 subtypes using sequence-based criteria derived from HIV-1 subtype B[20-
168 22]. Previous authors have correlated the use of sequence based criteria derived from HIV-1
169 subtype B for analysis of non-B subtypes and found it appropriate[6, 21]

170 Prediction of HIV-1 coreceptor usage by genotypic methods are easier to do, more cost
171 effective and have been found to be useful as screening strategy in medical practice[23, 24].
172 Moreover, the anti-CCR5 inhibitor, Maraviroc, can be taken with or without food, and does
173 not require refrigeration and thus, can be very suitable for resource-limited settings like sub-
174 Saharan Africa.

175

176 The presence of CXCR4-tropic HIV-1 isolates in HIV-1 infected individuals is associated with
177 a rapid decline of CD4+ T cells, rapid disease progression, and reduced survival time after
178 AIDS diagnosis[25]. Various studies on motif analysis of the V3 region of *env* have shown
179 certain characteristics that are strongly associated with CXCR4 usage and they include: the
180 presence and accumulation of CXCR4-associated mutations; insertions and deletions in V3,
181 the presence of basic amino acids at V3 positions 11 and/or 25, the loss of a glycosylation
182 site and an increased positive amino acid charge ($\geq +6$) [10-14].

183

184 Although the typical V3 loop amino acid sequence is 35 amino acids, V3 amino acid
185 sequences of HIV-1 subtypes and CRFs found in this study showed that all, except three
186 sequences, were 34-amino acids in length. Short length V3 amino acid sequences have
187 been reported by other investigators and are frequently associated with CXCR4 usage [5-6].
188 Sequence conservation in V3 loop was almost entirely consistent when considering the first
189 8 amino acid residues proximal to the terminal Cysteine residue in the base region across
190 the subtypes evaluated in this study. It has been reported that sequence conservation in the
191 V3 region across different HIV-1 subtypes is consistent with a V3 interaction with an
192 invariable cellular protein[26]. The sequence variability in the stem and turn regions is as a
193 result of site specific mutations and consequently, most (7/10) of the X4-associated
194 mutations detected in this study occur in the stem and turn regions (positions 9-25) of the
195 V3 loop. This observation is consistent with findings by previous researchers and these site-
196 specific mutations at either side of the V3 loop crown motif are important determinants of
197 CXCR4 coreceptor usage[26, 13].

198

199 Some authors have shown that the V3 loop of HIV-1 has a highly conserved crown motif and
200 the GPGQ motif (which was in the majority in this study) in particular, is generally associated
201 with African HIV isolates and it occurs irrespective of coreceptor usage [13,16]. However,
202 HIV-1 subtype B has a GPGR crown motif irrespective of coreceptor usage while in HIV-1
203 subtype C, viruses using the CCR5 coreceptor have conserved GPGQ sequence tip and
204 CXCR4 viruses have a substitution which commonly presents as GPGR [20]. In addition, the
205 substitution of GPGQ to GPGL/R/K as observed in 18.8% of the V3 loop sequences of HIV-1
206 subtypes and CRFs in this study, predict CXCR4 tropism, which is in agreement with some
207 studies[27, 28].

208

209 Another determinant of CXCR4 coreceptor usage of HIV-1 isolates which had been reported
210 from previous studies is the presence of basic amino acids at either positions 11 and/or 25 of
211 the V3 loop. Sixteen (94.1%) of the 17 V3 loop sequences with basic amino acids at
212 positions 11 had Arginine residues, while one sequence had Lysine and Arginine residues

at positions 11 and 25, respectively. Fouchier and colleagues[25] reported a strong association between the syncytium-inducing (SI) capacity of HIV-1 and the presence of positively charged amino acid residues at positions 11 and 25 of HIV-1 V3 loop, indicating that the V3 genotype may be useful for prediction of the viral phenotype. Also, studies of 3D molecular modeling of the V3 loops from primary isolates whose coreceptor usage was experimentally defined have revealed a charged "patch" on the surface of V3 that correlates with coreceptor usage. This V3 surface patch is positively charged in X4-tropic viruses and negatively charged or neutral in R5-tropic viruses, and is formed by two amino acids, at position 11 and at position 24 or 25; amino acids 11 and 24 or 11 and 25 contact each other in 3D space[29]. However, a study showed that sequence variability in V3 loop among subtype C viruses was not associated with the presence of basic amino acid substitutions which is in contrast to that observed in subtype B where sequence variability is associated with basic substitutions and is a determinant of altered coreceptor usage[30].

The V3 loop sequence of one sample had a net charge of +6 and multiple amino acid deletions, including the deletion of the N-linked glycosylation site. Some studies have shown that a high net charge of +5 and above in V3 loop with a loss of one or more N-linked glycosylation sites in the C2-V3 region of HIV-1 are characteristics of X-tropic viruses[28, 31-33].

Majority (68.8%) of the V3 loop sequences of the HIV-1 strains from this study had one or more X4-associated mutations including the 16 sequences that had only basic amino acid at V3 loop position 11 and/or 25. These X-4 associated mutations are consistent with report from previous studies. A review that compared different genotypic coreceptor usage prediction methods and analyses showed that the presence of mutations in the V3 loop is predictive and indicative of CXCR4 coreceptor usage[34]. Sing *et al.*[14] using univariate analyses for the prediction of HIV coreceptor use from clonal HIV V3 loop sequences identified 41 V3 mutations significantly associated with coreceptor usage. Marcelin and colleagues[35] reported that one reason for the failure of CCR5 inhibitors was the selection of resistance to CCR5 antagonists through amino acid changes in V3 loop.

Furthermore, other studies have shown that these X4-associated mutations in the V3 loop of *env* and the presence of basic amino acids at V3 positions 11 and 25 are strongly associated with syncytium induction and CXCR4 usage. Chesebro and colleagues in their studies on HIV infectivity and replication in different permissive cells, observed that infectivity and replication in macrophages was affected by the presence of basic amino acids in positions 11 and 25 of V3 with viruses harboring such mutations being T-cell tropic and syncytium inducing [10-12]. These observations were also reported by Fouchier and colleagues in their studies[25, 36, 37]. Hoffman *et al.*[38] showed that these basic amino acids in V3 is phenotype associated and results in the use of the CXCR4 coreceptor. Studies on HIV-1 V3 motif analysis by Jensen *et al.*[5] also observed the association between basic amino acids in V3 and CXCR4 coreceptor usage. Raymond *et al.*[24] Soulie *et al.*[39] and Ajoge *et al.* [9] in their various studies also reported that the presence of basic amino acids at V3 positions 11 and 25 was associated with CXCR4 coreceptor usage.

The 4 N-linked glycosylation sites (NGS) within the C2V3 region are present in different proportions in all, except one, of the sequences. Only 37.5% of the C2V3 sequences had the 4 NGS. The loss of one or more N-linked glycosylation sites has been shown to be associated with the development of X4 viruses [23], assist in more efficient use of CXCR4 and might be an important factor in the switch of CCR5 to CXCR4 viruses, and therefore, virus tropism[40, 41]. Pollaskis *et al.*[31] also reported that the loss of an N-linked glycosylation site within the V3 region had a major influence on virus switching from R5 to X4 phenotype.

Overall, the predicted coreceptor usage using 4 combined criteria [14, 29] for the V3 loop amino acid sequences of the HIV-1 subtypes and CRFs in this study was 31.2% and 68.8% for CCR5 and CXCR4, respectively, independent of CD4 and viral load status of the individuals, and, irrespective of subtype. The use of the combined criteria has been shown to increase both the sensitivity and the specificity of the prediction as reported by Raymond *et al.* [19] and Ajoge *et al.* [35]. The high proportion of CXCR4 in this study might be an indication that the HIV-1 infected individuals are in the advanced stage of infection associated with accelerated disease progression, or they may be exposed to anti-retroviral therapy, or both[32, 42-44].

The X4 phenotype predominates in both subtype G and CRF02_AG viruses but is not the predominant predicted phenotype in CRF06_cpx. The higher usage of CXCR4 by HIV-1G from this study is in agreement with the results of a previous study and further corroborates the evidence that HIV-1G may have a higher proportion of X4 variants [35], although more phenotypic assays for coreceptor usage for subtype G with corresponding genotypic testing needs to be done. However, the high usage of CXCR4 by HIV-1 CRF02_AG is contrary to reports from studies by Raymond *et al.* [17] and Ajoge *et al.*[35] that reported CRF02_AG HIV-1 strains to be predominantly CCR5-tropic. The prediction on CRF06_cpx shows that individuals infected with this subtype of HIV-1 may benefit from coreceptor antagonists like Maraviroc, although there is need for better characterization of HIV-1 CRF06_cpx tropism, and HIV-1G, in Nigeria.

In addition, 2 (10%) of the predicted CCR5-using HIV-1 from this study had the A22T mutation that confers resistance to CCR5 inhibitors like Maraviroc. The high prevalence of X4 tropic viruses and the presence of Maraviroc resistant mutations in some R5 viruses indicate that coreceptor usage ability of patients' viral population will clearly be required before clinical administration of Maraviroc. Moreover, if future phenotypic and genotypic studies confirm the predominance of Maraviroc-resistant CCR5, and CXCR4 phenotypes in Nigeria, a wide scale introduction and clinical application of Maraviroc in HIV-1 treatment and control will be limited.

4. CONCLUSION

This data show a high prevalence of HIV-1 subtype with the GPGQ crown motif in the V3 region suggesting a functional use of this motif. Moreover, X4 viruses predominate in the study population and it is associated with multiple amino acid deletions and mutations in V3 and the loss of one or more N-linked glycosylation sites with some CCR5-tropic viruses showing Maraviroc resistant mutations.

It will be important to carry out genotypic prediction tests for HIV-1 coreceptor usage in HIV patients' population encompassing the various regions of the country and in the predominant circulating HIV-1 subtypes before the introduction of CCR5 antagonists for clinical management of infected persons in Nigeria.

The sequences have been submitted in GenBank with Accession numbers **KF437580-KF437619**.

313 **CONSENT**

314

315 Written informed consent was obtained from every individual whose blood sample was used
316 for the study.

317

318 **ETHICAL APPROVAL**

319 The University of Ibadan/UCH ethical review board approved the study protocol and written
320 informed consent was obtained from every individual whose blood sample was used for the
321 study.

322

323 **References**

- 324 1. Bjorndal A, Sonnerborg A, Tscherning C, Albert J, Fenyo EM. Phenotypic characteristics of human immunodeficiency
325 virus type 1 subtype C isolates of Ethiopian AIDS patients. *AIDS Res Hum Retroviruses*. 1999;15(7):647-53.
326 doi:10.1089/088922299310944.
- 327 2. Kupfer B, Kaiser R, Rockstroh JK, Matz B, Schneeweis KE. Role of HIV-1 phenotype in viral pathogenesis and its
328 relation to viral load and CD4+ T-cell count. *J Med Virol*. 1998;56(3):259-63. doi:10.1002/(SICI)1096-
329 9071(199811)56:3<259::AID-JMV14>3.0.CO;2-9 [pii].
- 330 3. Koot M, Vos AH, Keet RP, de Goede RE, Dercksen MW, Terpstra FG et al. HIV-1 biological phenotype in long-term
331 infected individuals evaluated with an MT-2 cocultivation assay. *AIDS*. 1992;6(1):49-54.
- 332 4. Berger A, Doerr HW, Weber B. Human immunodeficiency virus and hepatitis B virus infection in pregnancy: diagnostic
333 potential of viral genome detection. *Intervirology*. 1998;41(4-5):201-7. doi:24937 [pii]
334 24937.
- 335 5. Jensen MA, Li FS, van 't Wout AB, Nickle DC, Shriner D, He HX et al. Improved coreceptor usage prediction and
336 genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop
337 sequences. *J Virol*. 2003;77(24):13376-88.
- 338 6. Jensen MA, van 't Wout AB. Predicting HIV-1 coreceptor usage with sequence analysis. *AIDS Rev*. 2003;5(2):104-12.
- 339 7. Dorr P, Westby M, Dobbs S, Griffin P, Irvine B, Macartney M et al. Maraviroc (UK-427,857), a potent, orally
340 bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human
341 immunodeficiency virus type 1 activity. *Antimicrob Agents Chemother*. 2005;49(11):4721-32. doi:49/11/4721 [pii]
342 10.1128/AAC.49.11.4721-4732.2005.
- 343 8. Meini G, Materazzi A, Saladini F, Rosi A, Vicenti I, Mancini M et al. Stability of unfrozen whole blood DNA for remote
344 genotypic analysis of HIV-1 coreceptor tropism. *BMC Infect Dis*. 2013;13:508. doi:10.1186/1471-2334-13-508
345 1471-2334-13-508 [pii].
- 346 9. Ajoge HO, Gordon ML, de Oliveira T, Green TN, Ibrahim S, Shittu OS et al. Genetic characteristics, coreceptor usage
347 potential and evolution of Nigerian HIV-1 subtype G and CRF02_AG isolates. *PLoS One*. 2011;6(3):e17865.
348 doi:10.1371/journal.pone.0017865.
- 349 10. Hawkins CA, Chaplin B, Idoko J, Ekong E, Adewole I, Gashau W et al. Clinical and genotypic findings in HIV-infected
350 patients with the K65R mutation failing first-line antiretroviral therapy in Nigeria. *J Acquir Immune Defic Syndr*.
351 2009;52(2):228-34. doi:10.1097/QAI.0b013e3181b06125.
- 352 11. Chaplin B, Eisen G, Idoko J, Onwujekwe D, Idigbe E, Adewole I et al. Impact of HIV type 1 subtype on drug resistance
353 mutations in Nigerian patients failing first-line therapy. *AIDS Res Hum Retroviruses*. 2010;27(1):71-80.
354 doi:10.1089/aid.2010.0050.
- 355 12. Imade GE, Sagay AS, Chaplin B, Chebu P, Musa J, Okpokwu J et al. Short communication: Transmitted HIV drug
356 resistance in antiretroviral-naive pregnant women in north central Nigeria. *AIDS Res Hum Retroviruses*. 2013;29(2):127-33.
357 doi:10.1089/AID.2013.0074.
- 358 13. Kanki PJ, Hamel DJ, Sankale JL, Hsieh C, Thior I, Barin F et al. Human immunodeficiency virus type 1 subtypes differ
359 in disease progression. *J Infect Dis*. 1999;179(1):68-73. doi:JID980530 [pii]
360 10.1086/314557.
- 361 14. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version
362 4.0. *Mol Biol Evol*. 2007;24(8):1596-9. doi:msm092 [pii]
363 10.1093/molbev/msm092.

15. Chesebro B, Nishio J, Perryman S, Cann A, O'Brien W, Chen IS et al. Identification of human immunodeficiency virus envelope gene sequences influencing viral entry into CD4-positive HeLa cells, T-leukemia cells, and macrophages. *J Virol.* 1991;65(11):5782-9.

16. Chesebro B, Wehrly K, Nishio J, Perryman S. Macrophage-tropic human immunodeficiency virus isolates from different patients exhibit unusual V3 envelope sequence homogeneity in comparison with T-cell-tropic isolates: definition of critical amino acids involved in cell tropism. *J Virol.* 1992;66(11):6547-54.

17. Chesebro B, Wehrly K, Nishio J, Perryman S. Mapping of independent V3 envelope determinants of human immunodeficiency virus type 1 macrophage tropism and syncytium formation in lymphocytes. *J Virol.* 1996;70(12):9055-9.

18. Holm-Hansen C, Grothues D, Rustad S, Rosok B, Pascu FR, Asjo B. Characterization of HIV type 1 from Romanian children: lack of correlation between V3 loop amino acid sequence and syncytium formation in MT-2 cells. *AIDS Res Hum Retroviruses.* 1995;11(5):597-603.

19. Sing T, Low AJ, Beerenwinkel N, Sander O, Cheung PK, Domingues FS et al. Predicting HIV coreceptor usage on the basis of genetic and clinical covariates. *Antivir Ther.* 2007;12(7):1097-106.

20. Jensen MA, Coetzer M, van 't Wout AB, Morris L, Mullins JI. A reliable phenotype predictor for human immunodeficiency virus type 1 subtype C based on envelope V3 sequences. *J Virol.* 2006;80(10):4698-704. doi:80/10/4698 [pii] 10.1128/JVI.80.10.4698-4704.2006.

21. Monno L, Brindicci G, Saracino A, Punzi G, Cibelli D, Lagioia A et al. V3 sequences and paired HIV isolates from 52 non-subtype B HIV type 1-infected patients. *AIDS Res Hum Retroviruses.* 2010;26(3):365-72. doi:10.1089/aid.2009.0224.

22. Raymond S, Delobel P, Mavigner M, Cazabat M, Souyris C, Encinas S et al. Genotypic prediction of human immunodeficiency virus type 1 CRF02_AG tropism. *J Clin Microbiol.* 2009;47(7):2292-4. doi:JCM.02439-08 [pii] 10.1128/JCM.02439-08.

23. Chueca N, Garrido C, Alvarez M, Poveda E, de Dios Luna J, Zahonero N et al. Improvement in the determination of HIV-1 tropism using the V3 gene sequence and a combination of bioinformatic tools. *J Med Virol.* 2009;81(5):763-7. doi:10.1002/jmv.21425.

24. Raymond S, Delobel P, Mavigner M, Cazabat M, Souyris C, Sandres-Saune K et al. Correlation between genotypic predictions based on V3 sequences and phenotypic determination of HIV-1 tropism. *AIDS.* 2008;22(14):F11-6. doi:10.1097/QAD.0b013e32830ebcd4.

25. Fouchier RA, Brouwer M, Broersen SM, Schuitemaker H. Simple determination of human immunodeficiency virus type 1 syncytium-inducing V3 genotype by PCR. *J Clin Microbiol.* 1995;33(4):906-11.

26. Patel MB, Hoffman NG, Swanstrom R. Subtype-specific conformational differences within the V3 region of subtype B and subtype C human immunodeficiency virus type 1 Env proteins. *J Virol.* 2008;82(2):903-16. doi:JVI.01444-07 [pii] 10.1128/JVI.01444-07.

27. Coetzer M, Cilliers T, Ping LH, Swanstrom R, Morris L. Genetic characteristics of the V3 region associated with CXCR4 usage in HIV-1 subtype C isolates. *Virology.* 2006;356(1-2):95-105. doi:S0042-6822(06)00507-1 [pii] 10.1016/j.virol.2006.07.030.

28. Duri K, Soko W, Gumbo F, Kristiansen K, Mapingure M, Stray-Pedersen B et al. Genotypic analysis of human immunodeficiency virus type 1 env V3 loop sequences: bioinformatics prediction of coreceptor usage among 28 infected mother-infant pairs in a drug-naïve population. *AIDS Res Hum Retroviruses.* 2011;27(4):411-9. doi:10.1089/aid.2010.0142.

29. Cardozo T, Kimura T, Philpott S, Weiser B, Burger H, Zolla-Pazner S. Structural basis for coreceptor selectivity by the HIV type 1 V3 loop. *AIDS Res Hum Retroviruses.* 2007;23(3):415-26. doi:10.1089/aid.2006.0130.

30. Ping LH, Nelson JA, Hoffman IF, Schock J, Lamers SL, Goodman M et al. Characterization of V3 sequence heterogeneity in subtype C human immunodeficiency virus type 1 isolates from Malawi: underrepresentation of X4 variants. *J Virol.* 1999;73(8):6271-81.

31. Pollakis G, Kang S, Kliphuis A, Chalaby MI, Goudsmit J, Paxton WA. N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization. *J Biol Chem.* 2001;276(16):13433-41. doi:10.1074/jbc.M009779200 M009779200 [pii].

32. Johnston ER, Zijenah LS, Mutetwa S, Kantor R, Kittinunvorakoon C, Katzenstein DA. High frequency of syncytium-inducing and CXCR4-tropic viruses among human immunodeficiency virus type 1 subtype C-infected patients receiving antiretroviral treatment. *J Virol.* 2003;77(13):7682-8.

33. Kaleebu P, Nankya IL, Yirell DL, Shafer LA, Kyosiimire-Lugemwa J, Lule DB et al. Relation between chemokine receptor use, disease stage, and HIV-1 subtypes A and D: results from a rural Ugandan cohort. *J Acquir Immune Defic Syndr.* 2007;45(1):28-33. doi:10.1097/QAI.0b013e3180385aa0.

34. Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. Bioinformatics prediction of HIV coreceptor usage. *Nat Biotechnol.* 2007;25(12):1407-10. doi:nbt1371 [pii] 10.1038/nbt1371.

35. Marcelin AG, Ceccherini-Silberstein F, Perno CF, Calvez V. Resistance to novel drug classes. *Curr Opin HIV AIDS.* 2009;4(6):531-7. doi:10.1097/COH.0b013e328331d4b1 01222929-200911000-00013 [pii].

36. Fouchier RA, Brouwer M, Kootstra NA, Huisman HG, Schuitemaker H. HIV-1 macrophage tropism is determined at multiple levels of the viral replication cycle. *J Clin Invest*. 1994;94(5):1806-14. doi:10.1172/JCI117529.
37. Fouchier RA, Groenink M, Kootstra NA, Tersmette M, Huisman HG, Miedema F et al. Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. *J Virol*. 1992;66(5):3183-7.
38. Hoffman NG, Seillier-Moisewitsch F, Ahn J, Walker JM, Swanstrom R. Variability in the human immunodeficiency virus type 1 gp120 Env protein linked to phenotype-associated changes in the V3 loop. *J Virol*. 2002;76(8):3852-64.
39. Soulie C, Malet I, Lambert-Niclot S, Tubiana R, Thevenin M, Simon A et al. Primary genotypic resistance of HIV-1 to CCR5 antagonists in CCR5 antagonist treatment-naive patients. *AIDS*. 2008;22(16):2212-4. doi:10.1097/QAD.0b013e328313bf9c00002030-200810180-00020 [pii].
40. Nabatov AA, Pollakis G, Linnemann T, Kliphuis A, Chalaby MI, Paxton WA. Inpatient alterations in the human immunodeficiency virus type 1 gp120 V1V2 and V3 regions differentially modulate coreceptor usage, virus inhibition by CC/CXC chemokines, soluble CD4, and the b12 and 2G12 monoclonal antibodies. *J Virol*. 2004;78(1):524-30.
41. Polzer S, Dittmar MT, Schmitz H, Schreiber M. The N-linked glycan g15 within the V3 loop of the HIV-1 external glycoprotein gp120 affects coreceptor usage, cellular tropism, and neutralization. *Virology*. 2002;304(1):70-80. doi:S0042682202917605 [pii].
42. Westby M, van der Ryst E. CCR5 antagonists: host-targeted antivirals for the treatment of HIV infection. *Antivir Chem Chemother*. 2005;16(6):339-54.
43. Westby M. Resistance to CCR5 antagonists. *Curr Opin HIV AIDS*. 2007;2(2):137-44. doi:10.1097/COH.0b013e328014200701222929-200703000-00010 [pii].
44. Connor RI, Sheridan KE, Ceradini D, Choe S, Landau NR. Change in coreceptor use correlates with disease progression in HIV-1-infected individuals. *J Exp Med*. 1997;185(4):621-8.