Article

Not peer-reviewed version

Identification of a Clade-Specific HLA-C*03:02-Restricted Epitope GY9 Derived from the HIV-1 p17 Matrix Protein

Samuel Kyobe * , Savannah Mwesigwa , Gyaviira Nkurunungi , Gaone Retshabile , Moses Egesa , Joel Kabali , Grace P Kisitu , Eric Katagirya , Marion Amujal , Misaki Wayengera , Fred Katabazi , Edgar Kigozi , Fahim Yiga , Busisiwe C Mlotshwa , Lesedi Williams , Hakim Sendagire , Sununguko Wata Mpoloka , Moses L Lutaakome Joloba , Dithan Kiragga , Graeme Mardon , Adeodata Kekitiinwa , Mogomotsi Matshaba , Neil A Hanchard , Jacqueline L Kyosiimire , David Robinson

Posted Date: 8 October 2023

doi: 10.20944/preprints202310.0441.v1

Keywords: HIV control; HLA-C*03:02; Immunoinformatics

Article

Identification of a Clade-Specific HLA-C*03:02-Restricted Epitope GY9 Derived from the HIV-1 p17 Matrix Protein

Samuel Kyobe, [1,*,†] Savannah Mwesigwa [1,2], Gyaviira Nkurunungi [3,4], Gaone Retshabile [5], Moses Egesa [3,4], Joel Kabali [3], Grace P. Kisitu [6], Eric Katagirya [2], Marion Amujal [2], Misaki Wayengera [2], Fred Katabazi [2], Edgar Kigozi [2], Fahim Yiga [2], Busisiwe C. Mlotshwa [5], Lesedi Williams [5], Hakim Sendagire [1], Sununguko W. Mpoloka [5], Dithan Kiragga [6], Graeme Mardon [7,8], Adeodata Kekitiinwa [6,9], Mogomotsi Matshaba [9,10], Neil A. Hanchard [11], Jacqueline Kyosiimire-Lugemwa [3], and David Robinson [12], on behalf of the Collaborative African Genomics Network (CAfGEN) of the H3Africa Consortium

[1] Department of Medical Microbiology, College of Health Sciences, Makerere University, Kampala, Uganda
[2] Department of Immunology and Molecular Biology, College of Health Sciences, Makerere University, Kampala, Uganda
[3] The Medical Research Council (MRC)/UVRI & LSHTM Uganda Research Unit
[4] Department of Infection Biology, London School of Hygiene & Tropical Medicine, London, United Kingdom
[5] School of Allied Health Professions, Faculty of Health Sciences, University of Botswana, Gaborone, Botswana
[6] Baylor College of Medicine Children's Foundation, Kampala, Uganda
[7] Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA
[8] Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX 77030, USA
[9] Pediatric Retrovirology, Department of Pediatrics, Baylor College of Medicine, Houston, TX 77030, USA
[10] Botswana-Baylor Children's Clinical Centre of Excellence, Gaborone, Botswana
[11] National Human Genome Research Institute, National Institutes of Health, USA
[12] Department of Chemistry and Forensics, School of Science and Technology, Nottingham Trent University Clifton Lane, Nottingham, NG11 8NS, United Kingdom
* Correspondence: samuelkyobe@gmail.com
† Current address: Affiliation 1.

Abstract: Efforts towards an effective HIV-1 vaccine have remained mainly unsuccessful. There is increasing evidence for a potential role of HLA-C-restricted CD8$^+$ T cell responses in HIV-1 control, including our recent report of HLA-C*03:02 among African children. However, there are no documented optimal HIV-1 CD8$^+$ T cell epitopes restricted by HLA-C*03:02; additionally, the structural influence of HLA-C*03:02 on epitope binding is undetermined. Immunoinformatics approaches provide a fast and inexpensive method to discover HLA-restricted epitopes. Here, we employed immunopeptidomics to identify HLA-C*03:02-restricted CD8$^+$ T cell epitopes. We identified a clade-specific Gag-derived GY9 (GTEELRSLY) HIV-1 p17 matrix epitope exclusively restricted to HLA-C*03:02. Residues E62, T142, and E151 in the HLA-C*03:02 binding groove and positions p3, p6, and p9 on the GY9 epitope are crucial in shaping and stabilizing the epitope binding. Our findings support the growing evidence of the contribution of HLA-C molecules to HIV-1 control and provide a prospect for vaccine strategies.

Keywords: HIV control; HLA-C*03:02; immunoinformatics

---

1. Introduction

Host genetic factors play an important role in HIV-1 control [1]. Outside the Δ32 mutation in the chemokine receptor (CCR5) gene, genome-wide association studies (GWAS) have consistently identified variants in the major histocompatibility complex (MHC, also known as human leucocyte antigens [HLA]) class I alleles to play a significant role in the control of HIV-1 infection [2,3]. The

HLA class I alleles predominantly display intracellularly processed viral antigens on cell surfaces to elicit CD8$^+$ T lymphocytes (CTL) in the adaptive immune responses [4,5]. This cell-mediated immune response is responsible for the clearance of virally infected cells [4,5]. Therefore, the mechanisms of intracellular antigen processing, including proteasome cleavage, peptide loading, and transportation into the endoplasmic reticulum via the transporter associated with antigen processing and peptide stabilization on the MHC molecule for stable cell surface presentation are a subject of interest to understand the contribution of the HLA class I in viral control [6–10].

We recently documented the putative role of HLA-C*03:02, HLA-B*57:03, and HLA-B*58:01 in long-term non-progression (LTNP) of HIV-1 among children in Uganda and Botswana [11]. In contrast to HLA-B*57:03 and HLA-B*58:01, the specific mechanisms underlying HLA-C*03:02-mediated HIV-1 control have not been fully elucidated. As such, the molecular and immunological basis of how HLA-C*03:02 confers its protective effects against HIV-1 infection remains unknown. Prior to our work, the HLA-C*03:02 allele was demonstrated a have significant correlation with both reduced viral load and elevated CD4$^+$ T cell count within the South African population, although these associations did not attain statistical significance [12]. Considering these findings, it is evident that HLA-C*03:02 holds potential significance in HIV control, stimulating further investigation into the mechanisms underlying its immune-mediated effects. HLA-B*57:03 and HLA-B*58:01 molecules display highly restricted HIV-1 epitopes in the structural and non-structural proteins that mediate HIV control [10,13–15]. Most of these peptides are derived from the Gag protein; however, high immunogenicity has been demonstrated in the non-structural proteins Nef, Vif, and Vpr [15–18]. Additional research is therefore needed to unravel the intricate interplay between HLA-C*03:02 and HIV-1 proteins, as well as the immune responses triggered by this particular HLA allele. The HLA-restricted epitopes are characterized by their ability to induce effective qualitative and quantitative cellular immune responses [19,20]. These epitopes are crucial in driving robust and polyfunctional cellular immune responses, contributing to the recognition and targeting of HIV-infected cells [19–22]. Furthermore, HIV-1 epitopes have been demonstrated to elicit humoral immune responses, generating broadly neutralizing antibodies [23]. The selective pressure of protective HLA alleles is known to drive the emergence of escape mutants, though at the expense of viral replication fitness, which other compensatory mutations may counter [13,24]. However, accumulation of escape mutations in HLA-B*57:03/B*58:01-restricted epitopes abrogates the protective effect through various mechanisms, including qualitative binding to killer immunoglobulin-like receptors [13,25]. Nonetheless, developing epitope-based vaccines that efficiently elicit both humoral and cellular immune responses has re-emerged as a strategy to control the global HIV-1 epidemic [5,26,27].

The success of multi-epitope HIV-1 vaccines remains generally challenging due to the rapid genetic evolution of the virus, diverse HLA genetic polymorphism, and viral-clade geographical diversity [11,28,29]. Previous research has predominantly focused on characterizing HLA-restricted epitopes specific to protective HLA-A and HLA-B alleles in the context of HIV. At the same time, comparatively limited consideration has been given to exploring the protective HLA-C alleles [10]. Therefore, identifying and prioritizing protective HLA-C-restricted epitopes from locally prevalent HIV-1 clade remains viable for designing an optimal vaccine candidate. The scientific literature presents many methodological approaches for identifying optimal HIV-1 CTL epitopes, each yielding diverse outcomes [14,30]. This diversity underscores the complexity of epitope prediction and necessitates careful consideration of the most appropriate methodologies for accurate and comprehensive epitope discovery. However, it's crucial to emphasize that these approaches consistently exhibit a strong agreement between predictive and experimental methods [31]. In this study, we employed an immunoinformatics approach and identified four potentially HLA-C*03:02-restricted CD8$^+$ T cell epitopes. Furthermore, using an ELISpot assay, we experimentally validated that a clade-specific GY9 epitope derived from the p17 HIV-1 matrix protein is exclusively restricted to HLA-C*03:02 alleles in our population. Our observations further support the growing evidence of the contribution of HLA-C molecules to HIV-1 control and provide an opportunity for innovative vaccine strategies.

2. Results

2.1. *HIV-1 clades C and A have private and shared HLA-C*03:02-epitopes and preferentially accommodate hydrophobic residues in the distal pocket*

We used the NetMHCpan 4.1 and MotifScan servers to predict the epitope repertoire of HLA-C*03:02, and a total of 42,679 and 92 peptides were predicted, respectively. Expectedly the *env* and *pol* genes contribute the largest number of epitopes (Figure 1A). Among the NetMHCpan-predicted epitopes, 75 and 321 were predicted to meet the strong and weak binders' threshold, respectively (Table S1). The thresholds are expressed in terms of %Rank, the percentile of the predicted binding affinity compared to the distribution of binding affinities calculated on a set of random natural peptides. A similar number (and proportion) of strong and weak binders were predicted from the HIV-1 A and C proteome, and 76 (23.5%) peptides were found to be shared among the clades (Figures 1B and C). We then used NetChop 3.0 to determine a final set of 238 epitopes predicted to undergo proteasomal cleavage (Figure 1D). These epitope sequences range from 8–13mers with a predominance of 9-mers (65%, Figure 1E). Further analysis of their amino acid sequence pattern at the HLA-C*03:02 motif using sequence logos (Figure 1F) found that certain amino acids are predominant or conserved at positions 1 (p1), 2 (p2), and 9 (p9, C-terminus). The P9 position of the HLA-C*03:02 motif is occupied by leucine, a large hydrophobic amino acid, but the position also accepts large hydrophobic and neutral residues phenylalanine and tyrosine, respectively (Figure 1F). At position p2, the small hydrophobic residue alanine is preferred, but the small hydrophilic and neutral residues valine and threonine are also accommodated. Similar to p9, position p1 equally favours large hydrophobic residues phenylalanine, isoleucine, and lysine but also accepts a large neutral residue, tyrosine (Figure 1F).

2.2. *C*03:02-restricted stable epitopes are mainly derived from structural proteins of HIV*

Next, we performed in silico docking to determine and characterize the top-ranked HLA-C*03:02 epitopes that preferentially elicit CD8$^+$ T cell responses accounting for the putative protective effect [11]. First, we designed a 3D structural model of the HLA-C*03:02 molecule. The best template for model building was protein data bank (PDB) ID 5w6a.2 (HLA-C*06:02), with high sequence identity (94.7) and coverage, resulting in a model with high confidence scores, favorable stereochemistry, and stability, suitable for ligand binding studies (molecular docking) (Figures S1 and S2 and Table S2). According to our docking protocol, we found eight top-ranked conformations (peptides); with the best energetically favoured docking scores and extensive strong peptide-HLA (pHLA) hydrogen bonds (Figure 2A-D, Figure S3, Table 1). Four epitopes were found in structural HIV-1 proteins, including ^{71}GTEELRSLY^{79} (GY9) located on the gag gene derived from the p17 matrix protein, ^{43}GAERQGTLNF^{52} (GF10), and ^{324}AQNPEIVIY^{332} (AY9) encoded on the pol gene and ^{58}KAYETEMHN^{66} (KN9) located in the env gene derived from the gp120 protein. Other epitopes were derived from non-structural HIV-1 proteins, such as ^{84}GAFDLSFFL^{92} (GL9) and ^{114}WVYNTQGYF^{122} (WF9) from the Nef protein, while ^{128}VVSPRCEY^{135} (VY8) and ^{109}VSVESPVIL^{117} (VL9) are derived from Vif and Rev proteins respectively. To establish the structural basis of the stability of these predicted pHLA complexes, we performed an extensive conventional molecular dynamic (MD) simulation. We performed all-atom MD simulations of HLA-C*03:02 in the unbound form and on each of the eight pHLA complexes. The root means square deviation (RMSD) of protein atom from their initial structural position over time provides an assessment of the stability of the protein-ligand complexes. We calculated and compared the average RMSD of the Cα atoms of the pHLA complexes and the unbound HLA-C*03:02 (1.05Å). Four pHLA complexes exhibited convergence, particularly evident within the final 100ns of MD, as depicted in Figure 3A-D. The achieved convergence is reflected in notably reduced average RMSD values: 0.66Å (GY9), 0.67Å (GF10), 0.70Å (AY9), and 0.59Å (VL9) compared to the free HLA-C*03:02 molecule (Figure 3 A-D). The remaining pHLA complexes with KN9, GL9, WF9, and VY8 epitopes exhibited a lack of stability (Figure S3G and H). In our subsequent molecular dynamics (MD) trajectory analyses, we focused on analyzing the last 100 nanoseconds.

This data indicates that the molecules GY9, GF10, AY9, and VL9 exhibit enhanced conformational stability of the HLA-C*03:02 molecule upon binding.

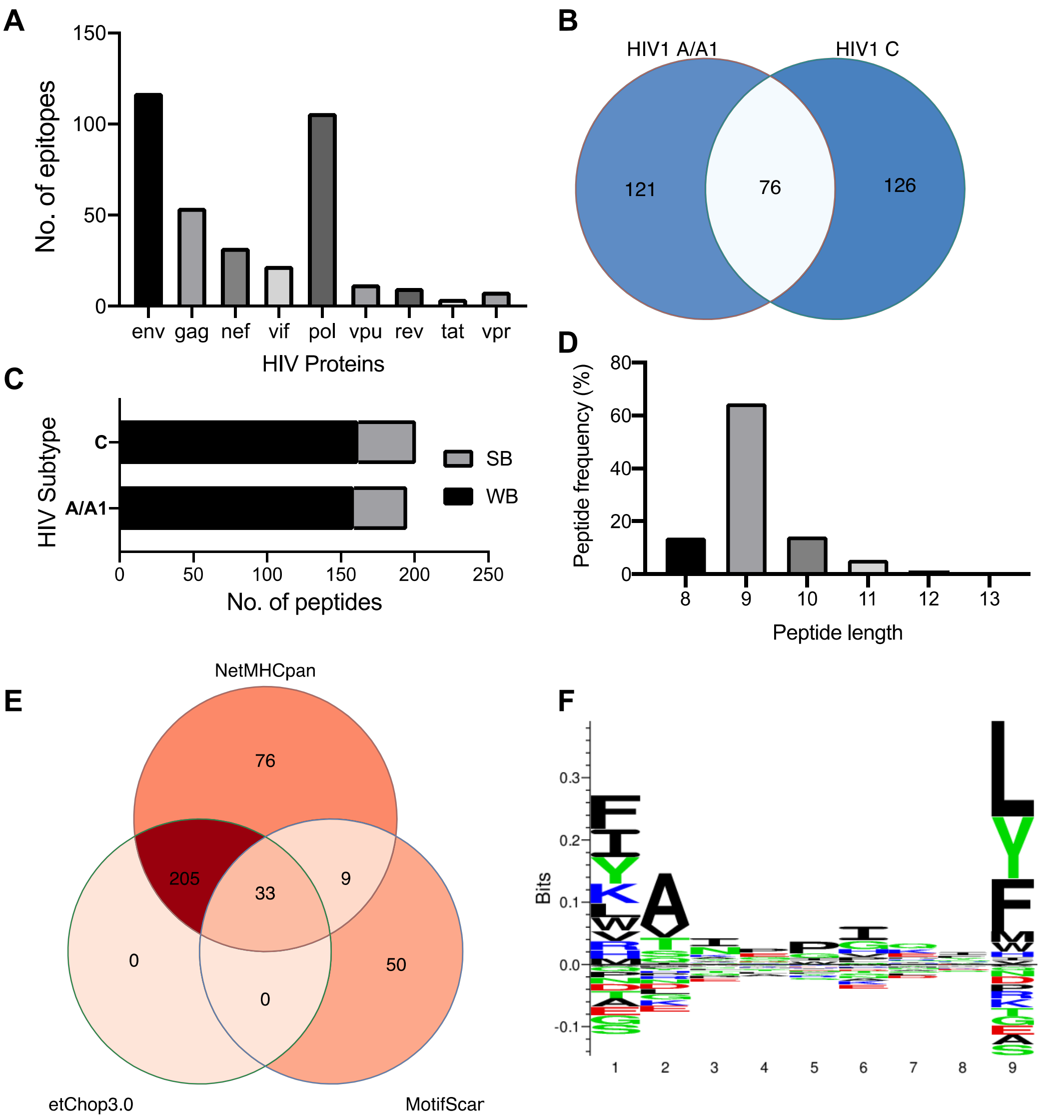

Figure 1. HIV 1 subtypes A/A1 and C predicted epitopes. (**A**) Distribution of predicted peptide length of HLA-C*03:02 allele in the HIV-1 A/C consensus peptidome. (**B**) Venn diagram showing the number of shared predicted epitopes between HIV-1 subtype A/A1 and C. (**C**) Proportion of HIV 1 predicted to be strong or weak binders. (**D**) Venn diagram showing the number of peptides predicted by NetMHCpan and Motif Scan methods. Also, the number of epitopes predicted to undergo proteasomal cleavage by TAP is shown. (**E**) Number of epitopes predicted by HIV protein. (**E**) Frequency of the length of predicted peptides. (**F**) Amino acid sequence logo representation of the most abundant residue at each position in the epitope. Prominent amino acid symbols indicate the most preferred amino acid in that epitope position. The sequence logo was calculated using clustering and pseudo counts with a weight on prior at 200, and data was handled with probability-weighted Kullback–Leibler probability distribution. Abbreviations: SB, strong binders; WB, weak binders.

Table 1. Molecular docking results of top-ranked HIV epitopes docked with C*0302.

HIV-1 clade	**Protein and Position**	**Peptide Sequence**		**No. of H bonds**	**RMSD**[a]	**Binding energy (kcal/mol)**	**Convex-PL score**	**Amino acids involved in H bond interaction**
	Structural							
C/A1	Gag^{71}*	GTEELRSLY	GY9	7	0.811	–8.6	7.23	Lys65, Tyr83, Tyr98, Thr142, Trp146 & Tyr158
A1	Env^{58}	KAYETEMHN	KN9	11	0.202	–7.4	7.36	Tyr6, Arg61, Glu62, Lys65, Tyr66, Ser76, Glu151 & Tyr158
C	Pol^{43}*	GAERQGTLNF	GF10	7	1.165	–8.7	7.22	Tyr6, Glu62, Lys65, Tyr98, Thr142, Lys145 & Gln154
C	Pol^{324}*	AQNPEIVIY	AY9	6	1.547	–7.7	7.16	Glu62, Lys65, Lys145, Tyr158 & Tyr170
	Regulatory							
C /A1	Nef^{84}	GAFDLSFFL	GL9	7	1.030	–9.1	7.24	Tyr8, Glu62, Lys65, Thr72, Tyr98, Lys145 & Tyr158
C/A1	Nef^{114}	WVYNTQGYF	WF9	7	0.873	–9.7	7.53	Tyr6, Arg68, Tyr98, Thr142, Lys145, Trp146 & Glu151
A1	Vif^{128}	VVSPRCEY	VY8	8	1.473	–8.3	7.10	Gln69, Thr72, Ser76, Tyr83, Tyr98, Thr142, Ly145 & Glu151
A1	Rev^{109}*	VSVESPVIL	VL9	7	1.638	–8.0	7.27	Tyr8, Arg61, Thr72, Asn79, Tyr83 & Tyr98

[a] RMSD Root mean square deviation (in Å) in comparison to the native ligand (PDB ID: 5w6a.2) of the C*0602 allele

*stable on molecular dynamics

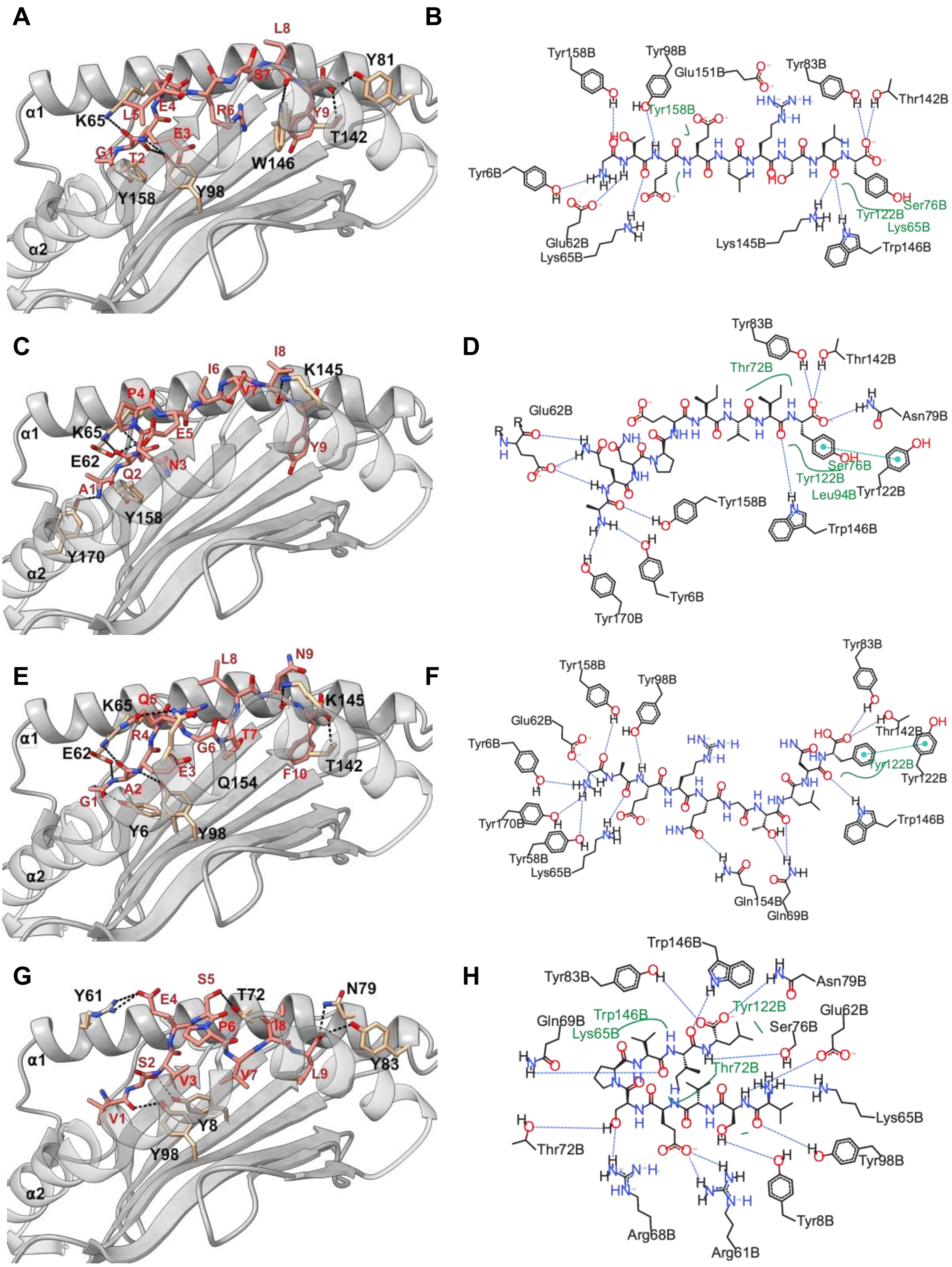

Figure 2. Molecular docking analysis and representation of HLA-C*03:02 interactions with docked epitopes GY9 (**A**&**B**), AY9 (**C**&**D**), GF10 (**E**&F), and VL9 (**G**&**H**). In the left panel is a 3D depiction of the HLA-C*03:02 molecule (ribbon representation) alongside the epitope (stick representation). Dashed lines highlight the presence of hydrogen bonds. Residues within HLA-C*03:02 contributing to hydrogen bond interactions are labeled in black, and residues in the epitope are labeled in red. The α2 chain has been rendered transparent, enabling clear visualization of the epitope. In the left panel is a 2D depiction (generated using PoseView, https://proteins.plus/) of the docked epitopes, with

hydrogen bonds depicted as black dashed lines and van der Waals forces shown in green. These illustrations provide an insightful view of the molecular interactions.

We then examined the formation of intermolecular interactions of the stable pHLA complexes. Hydrogen bonds play a significant role in forming and stabilizing pHLA complexes in the binding groove of HLA-C*03:02. We examined the hydrogen bond occupancy between the four epitopes and HLA-C*03:02 using the hydrogen bond module in VMD software [32]. We analyzed strong hydrogen bonds with an acceptor-donor atom distance ≤3.5Å and a hydrogen-to-donor-acceptor angle greater than 120°. We observed that residues Glu62, Thr142, and Lys170 in HLA-C*03:02 were involved in hydrogen bond formation with all four epitopes with more than 50% occupancy (Figure 2 and Fig 3E, Table S3). More than 50% hydrogen bond occupancy existed between Lys90, Trp171, and Glu151 in HLA-C*03:02 with at least three of the four epitopes (Figure 3E, Table S3). These results suggest that the HLA-C*03:02 binding groove favorably and stably binds three epitopes derived from structural HIV-1 proteins.

*2.3. Positions 62, 142, and 151 in HLA-C*03:02 and P6 in the GY9 epitope provide the structural basis for the preferential binding of GY9*

The binding free energy (ΔG) of pHLA complexes determines the stability of complex formation. Therefore, we applied the molecular mechanic/Poisson-Boltzmann surface area (MM/PBSA) method to estimate the binding free energies of GY9, AY9, GF1,0, and VL9 complexation with HLA-C*03:02. Generally, a more negative magnitude of the binding free energy corresponds to strong (high) binding affinities of pHLA complexes. Among the epitopes, GY9 showed a much stronger binding free energy of -88.41 kcal/mol, indicating a strong and favourable binding affinity to HLA-C*03:02 (Table 2). Notably, the van der Waals and electrostatic energies of -53.01 kcal/mol and -547.44 kcal/mol, respectively, between GY9 and HLA-C*03:02 contribute significantly to the binding (Table 2). We found that the electrostatic contribution of the GY9 epitope is much higher compared to other epitopes. Given the significant contribution of hydrogen bonding formation to the electrostatic energy, this means that hydrogen bonds are likely to play a critical role in GY9 binding to HLA-C*03:02. Also, the van der Waals energy is an indicator of the compactness of a ligand in the receptor binding groove; we found that GY9 also had the strongest value (-53.01 kcal/mol), suggesting a more favourable packing arrangement of GY9 in the HLA-C*03:02 binding groove (Table 2). This compactness is essential in pHLA complexes and affects efficient T cell receptor (TCR) engagement [33,34]. Overall, these findings provide valuable insights into the importance of van der Waals interactions, electrostatic interactions, and hydrogen bonding in the binding dynamics of HIV-1 epitopes to HLA-C*03:02 as well as the preference for GY9.

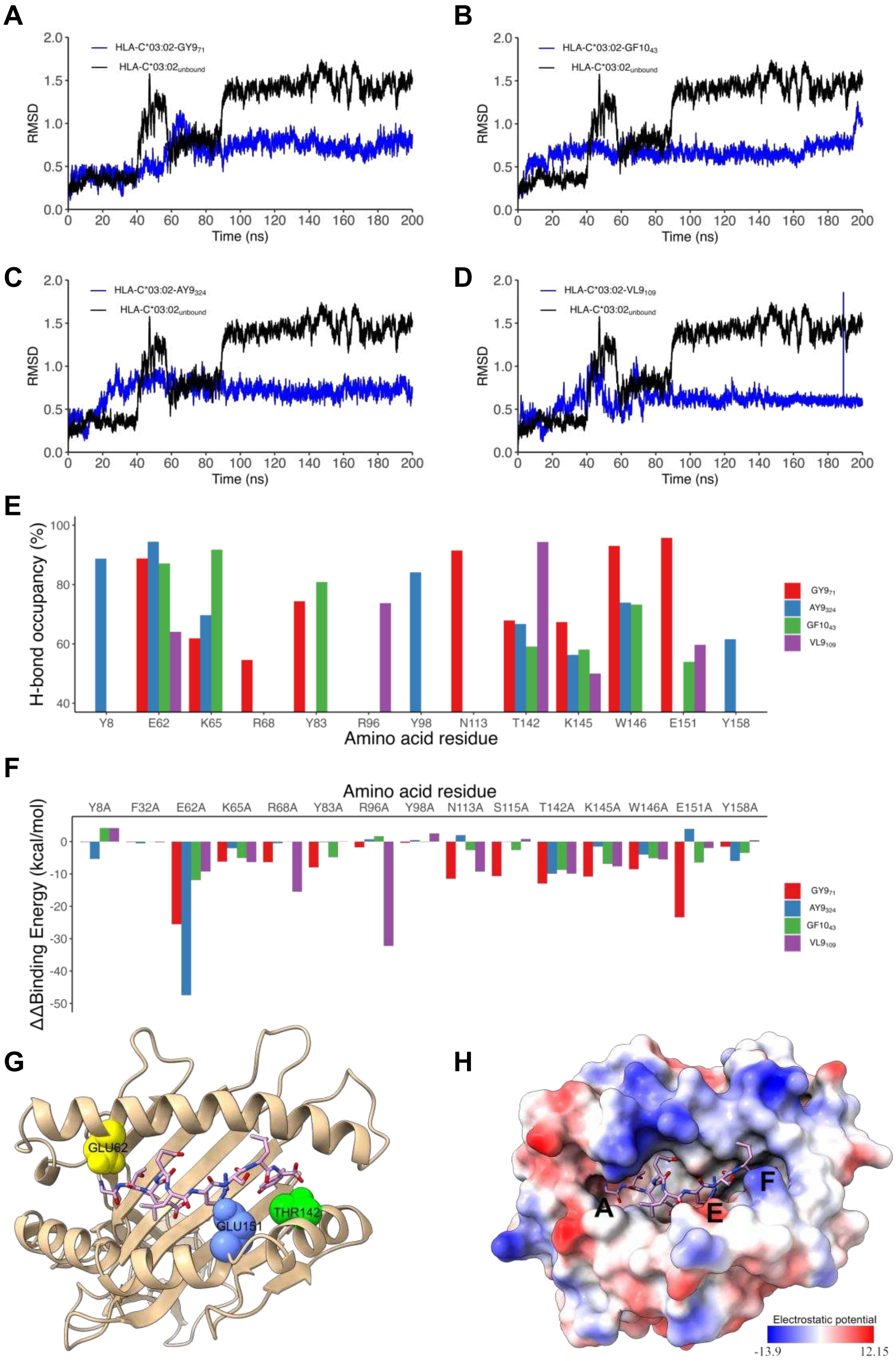

Figure 3. Molecular dynamics simulations analysis of HLA-C*03:02/peptide complexes. (**A**) through (**D**) are RMSD plots of the Cα backbone of HLA-C*03:02 in complex with VL9 (A), GY9 (**B**), GF10 (**C**) and AY9 (**D**). (**E**) The percent of hydrogen-bond occupancy for interactions between HLA-C*03:02 residues (donors or acceptors) and stable epitopes (VL9, GF10, GY9, and AY9) across the final 100ns. (**F**) Binding free energy change of key residues involved in HLA-C*03:02 binding with epitopes. (**G**)

3D structure of HLA-C*03:02 displaying the key residues Glu62, Glu151, and Thr142 in pocket A (yellow), E (blue), and F (green), respectively. (**H**) The electrostatic surface potential of HLA-C*03:02 Electrostatic potential was calculated and visualized using ChimeraX default settings. The color scale ranges from −13 (red) to +12 (blue) kT/e.

Table 2. Binding free energies obtained by the MM/PBSA method of pHLA complexes with GY9, GF10, AY9, and VL9 peptide.

HIV-1 subtype	Peptide Sequence		Energy Components (kcal/mol)			
			van der Waals	Electrostatics	Polar solvation	ΔG Binding energy
	Structural					
C/A1	GTEELRSLY	GY9	-53.01	-547.44	521.66	-88.41
C	GAERQGTLNF	GF10	-53.45	-335.96	347.44	-50.59
C	AQNPEIVIY	AY9	-68.75	-258.85	285.79	-51.24
	Regulatory					
A1	VSVESPVIL	VL9	-67.29	-367.42	394.36	-49.73

To gain insight into the individual contributions of the amino acids within the HLA-C*03:02 binding groove to the binding free energy, we performed a computational alanine scanning (CAS, or mutagenesis) based on the MM/PBSA method [35]. A negative value of ΔΔG indicates a favourable contribution for the wild-type residue in that position and vice versa. We mutated 35 amino acid residues within 5Å of the epitopes to alanine and computed the binding free energy difference between wild-type and mutant pHLA complexes. Notably, mutants E62A, T142A, and E151A in HLA-C*03:02 resulted in a significant loss of binding free energy with GY9 (Figure 3F, Table S4). For position 62, the mutation to alanine (E62A) resulted in a loss of binding free energy ranging from -9.25 kcal/mol to -47.47 kcal/mol across different peptide ligands (GY9, GF10, AY9, VL9). Similarly, for position 142, alanine mutation (T142A) led to a decrease in binding free energy ranging from -9.91 kcal/mol to -9.94 kcal/mol. Likewise, for position 151, alanine mutation (E151A) resulted in reduced binding free energy ranging from -1.98 kcal/mol to -23.40 kcal/mol, except for AY9, where a positive change in binding free energy of 3.93 kcal/mol was observed. Consequently, these three positions, E62, T142, and E151, found in the A, E, and F pockets of the HLA-C*03:02 peptide-binding groove (Figure 3G), are predicted to confer epitope specificity.

Previous studies have shown that point mutations within epitopes significantly diminish or abrogate immune responses [34]. We performed CAS on the GY9 epitope to establish the most influential positions to the binding affinity. Surprisingly, we noted a consistent trend in a decrease in the binding free energy (ΔΔG) across all the amino acid positions in GY9 except p1 with a small glycine residue (similar to alanine). However, p6 demonstrated a significant negative loss in binding free energy (ΔΔG -27.00 kcal/mol; Table 3), further reinforcing the importance of Arg6 at this position for binding affinity. We observed that

Table 3. Change in binding free energy and conservancy scores of the GY9 peptide.

HIV1 Clade[a]	Amino Acid Residue and Position								
	G1	**T2**	**E3**	**E4**	**L5**	**R6**	**S7**	**L8**	**Y9**
A1	.	.	.	.	.	R/K	.	.	Y/F
C	.	.	.	.	.	K	.	.	Y/F/H
D	.	.	.	.	I	K	.	.	Y/F
K	.	.	.	.	I	K	.	.	Y/F

ΔΔG GY9[b]	NA	-8.18	-10.3	-4.82	-2.83	-27.00	-5.85	-2.82	-7.33
Conservancy score[c]	5	4	2	3	3	2	5	5	2

[a] common clades in Botswana and Uganda populations; NA, glycine is of similar size to alanine

[b] Binding free energy change due to mutation of amino acid residue to alanine

[c] The scores are from one to nine to show the conservation level (low to high, respectively)

Arg at position p6 in GY9 leads to forming of three hydrogen bonds (donor) with residues E176, W171, and N138 (Table S3). Remarkably, these hydrogen bonds exhibit a high occupancy of over 90% throughout the MD simulation, indicating their persistent and stable nature. This suggests that stabilizing p6 is vital to prevent protrusion of the epitope out of the peptide binding groove that would considerably alter the structure of the pHLA-TCR binding platform. We computed the conservancy score by aligning viral sequences from all publicly available HIV-1 subtypes A, C, D, and K and their recombinants. We also found that p3, p6, and p9 had the lowest conservancy score (Table 3). These results suggest that p6 contributes favorably to GY9 binding and may serve as the primary anchor residue, while positions p3 and p9 are secondary anchor residues refining epitope binding

*2.4. The GY9 epitope elicits a clade-specific HLA-C*03:02-restricted IFN-γ response*

To discern the immunogenic potential of GY9 ex vivo, we assessed GY9-specific $CD8^+$ T cells, employing a dual color enzyme-linked immunospot (ELISpot) assay to measure the production of IFN-γ and IL-2. IFN-γ production indicates an active immune response, reflecting ongoing T cell effector functions. On the other hand, IL-2 secreted by activated T cells or NK cells plays a crucial role in driving the proliferation and differentiation of naive T cells, B cells, and NK cells, facilitating their transition into effector (such as Th1) and memory cells, and promoting the release of secondary cytokines. We used peripheral blood mononuclear cells (PBMC) from a study population that included 25 subjects on antiretroviral therapy (ART) recruited from Uganda, of whom 13 were expressing the HLA-C*03:02 allele (HLA typing is described elsewhere, Table S5) [11]. GY9-specific IFN-γ production ranging from 65 to 940 SFU/million PBMC, was found in 3/10 (30%) HLA-C*03:02^{+ve} subjects. Still, no response was found among any HLA-C*03:02^{-ve} subjects (Figure 4, Table S5). All GY9 responders were coincidentally infected with HIV-1 clade A1 (2/3) or C (Figure 4, Table S5). It should be noted that GY9 originated from both the A1 and C clade consensus sequences (https://www.hiv.lanl.gov/content/index). Except for three individuals for whom HIV-1 could not be typed, all non-responders to GY9 were found to be infected with HIV-1 clades A1, C, D, or A1D recombinant strains (Table S5). Unsurprisingly, in our cohort on chronic ART (1–121 months), we detected no IL-2 production in HLA-C*03:02^{+ve} or HLA-C*03:02^{-ve} individuals (Table S5) [36]. We have already demonstrated above that some positions with the GY9 epitope are under selective pressure (CAS and conservancy scores). We think the lack of response in HLA-C*03:02^{+ve} subjects infected with A1 may suggest the presence of escape mutants, especially in positions p6 > p3 > p9. These findings indicate that GY9 elicits a clade-specific immune response and exhibits non-promiscuity for HLA types.

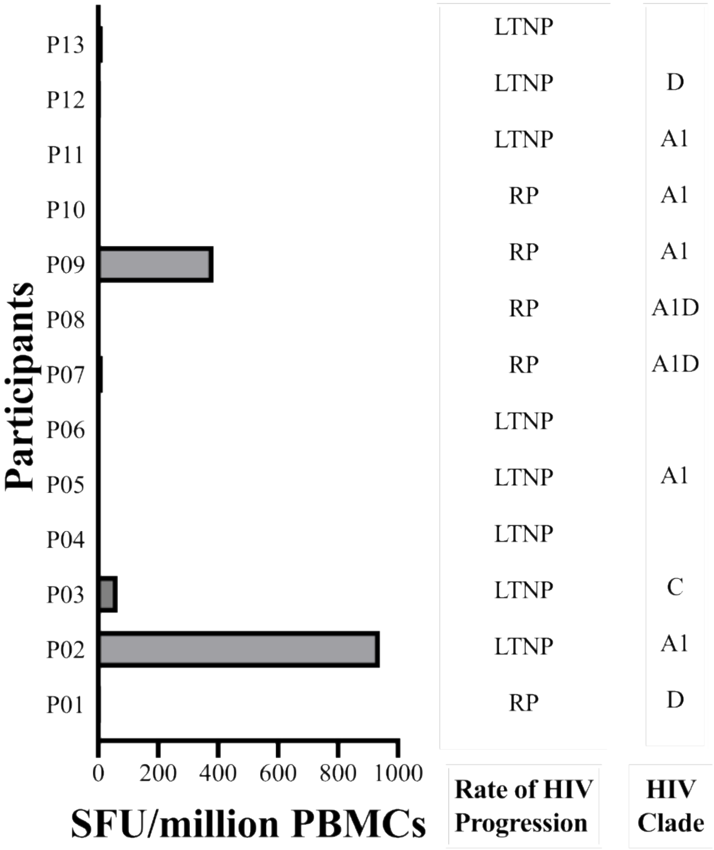

Figure 4. GY9-specific IFN-γ CD8⁺ T cell responses among HLA-C*03:02+ve participants. The magnitude of IFN-γ CD8⁺ T cell responses against GY9 epitope were measured among HLA-C*03:02^{+ve} and HLA-C*03:02^{-ve} (data not shown) participants by ELISpot assay. LTNP, long-term nonprogression; RP, rapid progression.

3. Discussion

Several approaches are being investigated to develop novel HIV-1 vaccines. Among these approaches is the search for multi-epitope vaccine candidates that elicit effective quantitative and qualitative humoral and cellular immune responses [26,37]. Theoretically, T-cell-based vaccines, utilizing peptides identified through in silico predictions, hold promise as effective vaccination strategies, particularly when focused on pinpointing the most immunogenic antigens [26]. In this study, we have utilized a synergy of computational techniques and empirical functional validation to uncover a previously unrecognized HIV-1 epitope, GY9. This epitope displays a distinctive potential for presentation by the HLA-C*03:02 allele, associated with effective HIV-1 control among African pediatric populations. Consistent with previous reports, the role of the HIV-1 Gag protein is prominent in providing the most immunogenic peptides presented by class I HLA molecules. Additionally, we report three epitopes in the Env and Pol proteins that map to HIV-1 subtypes A and C, suggesting that some control of HIV-1 may be attributable to HLA-C*03:02.

The HIV-1 Gag protein is preferred for T cell vaccine candidates because it is highly immunogenic and conserved across HIV-1 clades [16,38]. Several T cell candidate vaccines have so far shown variable immunogenicity [38]; however, thus far, this GY9 epitope has not been reported to show immunogenicity or restriction to HLA-C*03:02 and therefore has not been considered a potential vaccine candidate [38,39]. This could be attributed to the lack of prioritization of HLA-C preferential antigens. The role of HLA-C class I molecules in delaying HIV-1 progression has been historically considered less significant, primarily attributed to their lower surface expression levels. This phenomenon might arise from the underappreciation of HLA-C-preferred antigens. The historical perception regarding the impact of HLA-C class I molecules on the control of HIV progression has been relatively suppressed, often linked to their comparatively lower levels of surface expression and high LD with HLA-A and B alleles. Consequently, their contribution to HIV-1 control has not been prominently emphasized. Notably, to date, less than 10% (22/280) of optimal HIV-1 CTL epitopes ("*A list*") defined in the LANL HIV-1 epitope database are HLA-C-restricted epitopes. Evidence is progressively accumulating, highlighting the potential contribution of HLA-C molecules to control HIV-1, particularly when focusing on the conserved Gag epitopes [30,40,41]. In our study, we observed a variable magnitude of IFN-γ responses and no detectable IL-2 response upon stimulation of T cells from people living with HIV with the GY9 epitope, consistent with findings reported in previous studies [41,42]. This variable magnitude of cytokine (IFN-γ and IL-2) response could be attributed to several factors, such as immune exhaustion due to chronic infection and ART (IL-2 ablation and low IFN-γ production) among this cohort and viral escape within the GY9 epitope (no IFN-γ responses) [41,43–45]. Indeed, we measured the magnitude of response using study participants on highly active ART (HAART; average duration 32 months [1–121 months]) [45]. A striking absence of both IFN-γ and IL-2 responses were observed in a larger number of participants with the HLA-C*03:02 allele; this aligns with the likelihood of amino acid modifications/mutations within the GY9 epitope sequence. Indeed, when we calculated conservancy scores, p6 and p3/p9 had a very low score (2), which means that these positions are associated with a high rate of mutations/variation across the various HIV-1 clades A, C, D, and K. Similarly, our CAS studies detected significant differences in the pHLA relative binding free energy when residues in p6 and p3/p9 are mutated to alanine, suggesting a very high contribution to epitope binding. A recent report by Li et al. suggests that mutations within the epitope significantly impact pHLA binding due to conformational changes and eventually affect TCR recognition and antigen presentation [34]. Collectively, our data strongly indicate that p6 and p3/p9 within the GY9 epitope serve as primary and secondary anchor residues, respectively, crucial for robust binding within the HLA-C*03:02 antigen-binding cleft, thereby facilitating optimal T cell receptor (TCR) engagement. Indeed, Joglekar et al. demonstrated that peptide-MHC binding is essential for TCR binding and that peptide mutations play an important role in viral escape [46]. Therefore, we argue that these findings show a potential viral escape and immune evasion pathway within the GY9 epitope [34]. The absence of detectable IL-2 responses to the GY9 epitope underscores the impaired capacity to reactivate HIV-specific memory T cells elicited during chronic infection, indicating a compromised immune response [36]. This observation aligns with the known phenomenon that HIV-1 infection leads to an expansion of CD8^{+}CD28^{-} T cells, characterized by their compromised ability to produce IL-2 [47]. Our data shows that residues E62, T142 and E151 in the HLA-C*03:02 binding groove, along with positions p3, p6 and p9 on the GY9 epitope, are critical hot spots for binding. These residues play a crucial role in shaping and stabilizing the protein-protein interface, significantly contributing to its stability. These compelling results strongly indicate the prominent influence of the HLA-C*03:02-restricted GY9 epitope sequence in shaping HIV control.

Immunogenicity in HIV-1 is not restricted to the Gag protein since numerous studies have established the role of epitopes derived from other HIV-1 proteins [30]. Indeed, HIV-1 vaccine candidate studies have demonstrated an advantage of multi-epitope prototypes [38]. In this study, our immunoinformatic approach identified three potentially immunogenic epitopes, GF10/AY9 and VL9, derived from the Pol and Rev proteins, respectively; however, we did not find any detectable HIV-1-specific CD8^{+} T cell responses against these epitopes in our population. While the lack of responses could be explained by similar factors noted above, the epitopes GF10 and AY9 are derived from the HIV1-C subtype; all our HLA-C*03:02^{+ve} participants used for the dual IFN-γ/IL-2 ELISpot assay were infected with HIV1-A1, C, D and the A1D recombinant. When we performed a conservancy score, we found that many positions along the epitopes had a very low conservancy score (VL9>GF10>AY9, data not shown) that could explain these positions as escape mutations that abrogate responses to epitopes derived from other HIV-1 clades and the potential unsuitability of these epitopes [34]. Overall, our docking results are similar to an experimental biological study where only 6-8 HIV-1 derived peptides were identified as restricted to HLA class I alleles [48]. In that study, Ziegler et al. infected CD4$^+$ T cells with HIV-1 and measured HLA class I (HLA-A*02:01/*02:01, B*27:05/*40:01, C*02:02/*03:04) repertoire, suggesting that these molecules present a small set of epitopes derived from the HIV-1 proteome at variable relative quantities [48].

While previous research on HIV-1 vaccine candidates has predominantly focused on the protective HLA-B alleles, it is noteworthy that HIV-nef attachment selectively downregulates the cell surface expression of both HLA-A and B molecules [49,50]. This downregulation phenomenon facilitates immune evasion through CTL escape by virally infected cells. Consequently, HLA-C-restricted CTL responses remain intact to facilitate the recognition and destruction of HIV-infected cells. Most crucially, the HLA-C*03:02 cytoplasmic tail lacks both tyrosine and aspartate, which are the targets of Nef-dependent downregulation of HLA cellular surface expression. Instead, HLA-C*03:02 has Leu321 and Val328 in the cytoplasmic tail [49]. Therefore, compensatory mechanisms enhance HLA-C cell surface expression, favorably explaining the role of HLA-C-restricted CTL responses in HIV-1 control [51]. Furthermore, HLA-C alleles lacking a binding site for microRNA-148a in the 3′ untranslated region of their messenger RNA exhibit a compensated high surface expression, potentially influencing immune recognition and responsiveness [52]. Interestingly, the HLA-C*03:02 allele demonstrates strong linkage disequilibrium with a C variant located 35kb upstream of the HLA-C gene. The presence of the -35C allele is strongly associated with increased cell surface expression of HLA-C molecules, potentially providing a mechanistic explanation for the observed impact of HLA-C*03:02 on HIV-1 control [51,53]. In this study, we find that GY9-induced IFN-γ responses were not shared with other HLA-C, -A, or -B alleles (Table S5); this would suggest that clade-specific GY9 HLA-C*03:02-restricted responses are highly allele-specific. This restricted binding specificity of the GY9 epitope is predicted to play a crucial role in determining immune responses following HIV-1 infection and may have implications for vaccine design and understanding the individual variation in immune recognition.

Limitations of the study

Our study has some notable limitations. Full or partial viral sequencing and CD4$^+$ T cell counts were not performed on all participants in our study. Viral sequencing could have yielded valuable insights into the degree of conservancy within the GY9 sequence among individuals positive or negative for the HLA-C*03:02 allele, enabling empirical assessment of epitope dominance in HIV-infected individuals. Our use of consensus and primary strain sequences for epitope prediction may potentially overlook naturally occurring epitopes in the studied population. A comprehensive analysis of primary strain sequences is crucial to identify conserved epitopes capable of eliciting robust and broad immune responses. Nonetheless, a recent study by Bugembe et al. demonstrated that the same computational tools used here identify 95% of experimentally mapped HIV-1 clade A and D epitopes [31]. Furthermore, measurements of CD4$^+$ T cell levels would have provided a baseline assessment of immunocompetence, a factor known to influence immune responses to HIV-1 epitopes [54]. In the future, investigations employing well-characterized study populations, incorporating advanced immunopeptidomics techniques, intracellular cytokine flow cytometry, and tetramer staining assays will be essential to build upon our current findings and overcome the methodological limitations observed in our study [30,55]. These approaches hold the potential to deepen our understanding of the immunological responses and contribute valuable insights to the field of immunology.

4. Materials and Methods

4.1. Patient recruitment

We used stored PBMC samples from 25 previously recruited participants in the parent study: the Collaborative African Genomics Network (CAfGEN). The details of participant recruitment have been described in detail elsewhere [11,56]. The clinical characteristics of patients before and after treatment are presented in Table S5. We selected all 13 participants expressing the HLA-C*03:02 allele and 12 controls that are HLA-C*03:02^{-ve}.

4.2. HLA-C*03:02 homology modeling and validation

The 3D structure model of HLA-C*03:02 was predicted using SWISS-MODEL (https://swissmodel.expasy.org/), starting with the 366 amino acids full-length protein sequence downloaded from the IMGT/HLA database (https://www.ebi.ac.uk/ipd/imgt/hla/). An optimal template to model the HLA-C*03:02 protein was selected based on PDB ID: 5w6a.2 HLA-C*06:02 with a sequence identity >90%, query coverage ≥70% and X-ray resolution at ≤2Å. The constructed model underwent comprehensive validation assessments in two distinct stereochemical and spatial analysis domains. The stereochemical analysis of parameters, including bond length, torsion angle, and rotational angle, within the model was evaluated using online tools servers SAVES (https://saves.mbi.ucla.edu/) and Pro-Q scores (https://proq.bioinfo.se/cgi-bin/ProQ/ProQ.cgi). The Ramachandran plot confirmed stereochemical quality (https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/). The spatial features of the model based on the 3D conformation were analyzed using the Verify 3D (https://saves.mbi.ucla.edu/) and ProSA scores (https://prosa.services.came.sbg.ac.at/prosa.php). The model's overall quality was determined from the ProTSAV score (http://www.scfbio-iitd.res.in/software/proteomics/protsav.jsp).

4.3. HIV-1 ligand prediction and preparation for docking

HIV-1 ligands (8-14mer) predicted to bind HLA-C*03:02 were determined using NetMHCpan-4.1b server (https://services.healthtech.dtu.dk/services/NetMHCpan-4.1/) and supplementary epitopes with Motif Scan (https://myhits.sib.swiss/cgi-bin/motif_scan). We used full-length HIV-1 subtype A/A1 and C consensus sequences retrieved from the HIV-1 Sequence Database (https://www.hiv.lanl.gov/content/index) representative of the predominant circulating clades in Uganda and Botswana, respectively [57]. From the available ligands, we selected a subset of those predicted as strong or weak binders by NetMHCpan-4 or Motif Scan. Additionally, these ligands were predicted to undergo proteasomal cleavage according to NetChop v3.1 (https://services.healthtech.dtu.dk/services/NetChop-3.1/). For the docking experiment, the 3D structure of the selected linear peptides was predicted using the PEP-FOLD3 server (https://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::PEP-FOLD3) followed by an energy minimization step using the minimize structure module in Chimera [58]. Briefly, essential hydrogens were added, and Gastiger charges were assigned to ligand residues using the Amber ff14SB force field. The 3D structures achieved convergence after 100 steps of steepest descent followed by 1000 steps of conjugate gradient [58].

4.4. Molecular docking protocol and analysis

The model HLA-C*03:02 structure was used for docking HIV-1 ligands with DINC, a parallelized meta-docking method for the incremental docking of large ligands. Some modifications were adopted to the default DINC protocol [59]. The grid box of 50 x 40 x 72 xyz points with a grid spacing of 0.375Å was generated and centered at 11.95 x 57.95 x -6.34 around the six binding pockets using AutoDock Tools [60]. To maximize the docking accuracy, the vina exhaustiveness was set to 8, and the number of binding modes generated at each round of incremental docking was set at 40. An additional round of docking was performed using the whole ligand with full flexibility to get a larger docking sampling. The predicted ligand poses were rescored using Convex-PL, shown to achieve >80% accuracy in identifying the best binders [61]. Molecular visualization with UCSF ChimeraX was used to identify and analyze the intermolecular (pHLA) interactions [58,62]. The selection of top-ranked ligand poses was guided by several rigorous criteria, including a Convex-PL score ≥7, a DINC binding score ≤ −7.0 kcal/mol, a minimum of 6 strong hydrogen bonds formed with the pHLA complex, and an RMSD ≤1.7Å compared to the native ligand (PDB ID: 5w6a.2) of the C*0602 allele.

4.5. Molecular dynamics simulation protocol and analysis

MD simulations were performed on the top-ranked ligands using GROMACSv2020.3 software under the CHARMM36 all-atom force field [63,64]. The receptor-ligand coordinates generated during molecular docking were utilized to reconstruct protein-ligand complexes using Chimera. All hydrogen molecules were removed from the final structure. We used the Avogadro program to add hydrogens to ligands and the CHARMM General Force Field (CGenFF) program to generate ligand parameters and topologies [65,66]. The resultant HLA-C*03:02-unbound and HLA-C*03:02-ligand complex were solvated in the center of a cubic unit cell of the volume of 10000nm^3 with ~31,000 molecules of TIP3-point water. We allowed a minimum distance of 1nm between the box boundary and the complex. The system was neutralized with the addition of 10 Na+ ions. The system was subjected to energy minimization using the steepest descent method with a maximum force constraint of 10kJ/mol. Position restraints were applied on both the ligand and HLA-C*03:02 receptor. The system temperature and pressure were equilibrated at 300K using the modified Berendsen thermostats coupling method and at 4.5x10^{-5}bar^{-1} using the Berendsen coupling barostats method, respectively, for 1000ps. All relaxed systems were subjected to MD simulations for 200ns using periodic boundary conditions without ligand-protein restraints. The stability of the complexes was examined by analyzing changes in the root mean square deviation (RMSD) and hydrogen bonds network using GROMACS functions hbond and rms respectively [63].

The binding free energy (denoted as $\Delta G_{bind} = \Delta G_{complex} - \Delta G_{receptor} - \Delta G_{peptide}$) was calculated using the molecular mechanics (MM) with Poisson–Boltzmann (PB) and surface area solvation method implemented in the gmx_MMPBSA program [35]. The critical residues at the interface of pHLA binding (within 5Å of the ligand) were determined by performing computational alanine scanning (CAS) experiments on the ligand and HLA-C*03:02. The resultant binding free energy due to the mutant residue was calculated by comparing the wild-type ($\Delta G_{wild\text{-}type}$) and mutant ($\Delta G_{mutant}$) complexes, as denoted by the equation: $\Delta\Delta G_{bind} = \Delta G_{wild\text{-}type} - \Delta G_{mutant}$.

4.6. HIV-1 epitope conservancy analysis

To assess the positional conservancy of the candidate epitopes at the individual residue level, we used the AL2CO sequence conservation analysis server (http://prodata.swmed.edu/al2co/). Specifically, we utilized an alignment file generated from African representative HIV-1 clades A, C, D, and K and their recombinant sequences deposited in the LANL HIV-1 Sequence Database (https://www.hiv.lanl.gov/content/index) for calculating the conservancy scores.

4.7. HIV-1 genotyping

The participants' genomic DNA was extracted from whole blood with the PaxGene DNA blood kit (Qiagen) as previously described. A three-round nested PCR assay was performed targeting the HIV-1 proviral DNA Gag-Pol region (the third round nested PCR is to add Illumina-specific adaptor sequences) [67,68]. The final PCR product was purified using the Agencourt AMPure XP magnetic beads (Beckman Coulter). The purified PCR was used for library preparation using the Nextera XT DNA Library Preparation Kit (Illumina) (indexing was done with the IDT for Illumina DNA/RNA UD Indexes Set A) according to the manufacturer's protocol. Equimolar concentrations of all samples were pooled and sequenced on an Illumina MiSeq instrument (Illumina) using the paired-end (2x300bp) method with the MiSeq-v3 reagent kit (Illumina). The read quality of the generated files was determined using FastQC, and the low-quality sequences were trimmed using Trimmomatic. The resultant reads were aligned/mapped to HIV-1 reference (RefSeq: NC_001802.1) using the BWA to generate viral contigs. HIV-1 subtyping was done using the REGA-v3 HIV-1 Subtyping Tool, [69] and any refractory sequences/samples were resolved using the RIP tool (https://www.hiv.lanl.gov/content/sequence/RIP/RIP.html) or HIV-1 blast (https://www.hiv.lanl.gov/content/sequence/BASIC_BLAST/basic_blast.html) followed by phylogenetic analysis with PhyML (https://www.hiv.lanl.gov/content/sequence/PHYML/interface.html).

4.8. HIV-1-specific IFN-γ and IL-2 dual ELISpot assay

HIV-1 specific HLA-C*03:02-restricted $CD8^+$ T cell responses were evaluated using a dual ELISpot assay. HIV-1 peptides were synthesized using the Fmoc (fluoren-9-ylmethoxycarbonyl) means of solid-phase peptide synthesis technology and the purity was confirmed using high-pressure liquid chromatography (Bio-Synthesis, Inc). Peptides were diluted to a final concentration of 10μg/ml. PBMCs were isolated by density gradient centrifugation from EDTA whole blood and cryopreserved. Frozen PBMCs were thawed, and viability was confirmed by trypan blue, then rested overnight before plating. We evaluated the secretion of IFN-γ and/or IL-2 by PBMC using the ELSP5710/5810 AID iSpot FluoroSpot kit (AID Autoimmun Diagnostika) according to the manufacturer's instructions. Briefly, 96-well plates pre-coated with both IFN-γ and IL-2 monoclonal antibodies were incubated with 100μl of 2×10^5 viable cells and 100μl of peptide solution per well at 37°C in humidified 5% CO_2 for 40 hrs. Media alone was used as a negative control (NC), and pokeweed as a positive control. Plates were washed and stained with biotinylated anti-human IL-2 and anti-human IFN-γ FITC. IFN-γ and IL-2 production was quantified using an AID iSpot EliSpot/FluoroSpot Reader (AID Autoimmun Diagnostika) and expressed as spot-forming cells (SFC) per million PBMC after subtraction of background spots from NC.

5. Conclusions

In conclusion, we have used an immunoinformatics approach to identify an HLA-C*03:02-restricted epitope, eliciting T cell-specific responses, suggesting that the GY9 epitope plays a significant role in HLA-C*03:02-mediated HIV-1 control among children. This study supports the hypothesis that an effective HIV-1 vaccine should be clade-specific; therefore, efforts for a global vaccine may not be feasible. And as such, more focus should be placed on identifying possible epitopes mapped among all clades, especially those restricted to protective HLA-C alleles. Finally, our study expands upon prior studies by providing evidence supporting the notion that the HIV-1 matrix protein p17 represents a promising epitope candidate for developing a vaccine against HIV/AIDS.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Figure S1: Model building and validation; Figure S2: Modeling validation; Figure S3: The molecular docking representation and RMSD analysis of C*0302 with docked epitopes GL9 (A, B and G), VY8 (C&D), WF9 (E) and KN9 (F and H); Table S1: List of peptides classified as strong or weak binders; Table S2: Stereochemical and spatial analysis of the HLA-C*03:02 model by different computational tools; Table S3: Hydrogen bond occupancy of C*03:02 residues making contact with peptides; Table S4: Calculated MM-PBSA binding energy difference; Table S5: Participant characteristics and IFN-γ response.

Author Contributions: Conceptualization, Samuel Kyobe,; Formal analysis, Samuel Kyobe,; Funding acquisition, Samuel Kyobe,, Sununguko Mpoloka, Moses Lutaakome Joloba, Dithan Kiragga, Graeme Mardon, Adeodata Kekitiinwa, Mogomotsi Matshaba, Neil Hanchard, Jacqueline Kyosiimire and David Robinson; Investigation, Savannah Mwesigwa, Joel Kabali, Grace Kisitu, Eric Katagirya, Marion Amujal, Misaki Wayengera, Fahim Yiga, Graeme Mardon, Neil Hanchard and Jacqueline Kyosiimire; Methodology, Samuel Kyobe,, Gyaviira Nkurunungi, Moses Egesa, Joel Kabali, Fahim Yiga, Graeme Mardon, Neil Hanchard, Jacqueline Kyosiimire and David Robinson; Project administration, Gaone Retshabile; Resources, Grace Kisitu and Graeme Mardon; Supervision, Hakim Sendagire, Sununguko Mpoloka, Graeme Mardon, Adeodata Kekitiinwa, Mogomotsi Matshaba, Neil Hanchard, Jacqueline Kyosiimire and David Robinson; Validation, Samuel Kyobe,; Visualization, Samuel Kyobe,; Writing – original draft, Samuel Kyobe,; Writing – review & editing, Savannah Mwesigwa, Gyaviira Nkurunungi, Gaone Retshabile, Moses Egesa, Joel Kabali, Grace Kisitu, Eric Katagirya, Marion Amujal, Misaki Wayengera, Fred Katabazi, Edgar Kigozi, Busisiwe Mlotshwa, Lesedi Williams, Hakim Sendagire, Sununguko Mpoloka, Dithan Kiragga, Graeme Mardon, Mogomotsi Matshaba, Neil Hanchard, Jacqueline Kyosiimire and David Robinson.

Funding: The project described was supported by Award Number, U54AI110398 administered by the National Institute of Allergy and Infectious Disease (NIAID), Eunice Kennedy Shriver National Institute Of Child Health & Human Development (NICHD), and National Human Genome Research Institute (NHGRI) as part of the NIH Common Fund H3Africa Initiative. This work was also supported by the Makerere University-Uganda Virus Research Institute Centre of Excellence for Infection and Immunity Research and Training (MUII). MUII is supported through the DELTAS Africa Initiative (Grant no. 107743). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS), Alliance for Accelerating Excellence in Science in Africa (AESA), and supported by the New Partnership for Africa's Development Planning and

Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust (Grant no. 107743) and the UK Government. The government of Uganda funded this work through Makerere University Research and Innovations Fund (grant number: MAK-RIF).

Institutional Review Board Statement: This study was approved by the School of Biomedical Sciences Institutional Review Board (IRB), Uganda National Council for Science and Technology, University of Botswana IRB, Botswana Health Research and Development Committee, and the Baylor College of Medicine IRB.

Informed Consent Statement: All participants provided written informed consent.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: We would like to acknowledge the following individuals for their contributions on behalf of the CAfGEN Consortium: Edward D. Pettitt, Marape Marape, and Bhekumusa Lukhele, who were participating investigators. We acknowledge Nasinghe Emmanuel, John Mukisa, Gaseene Sebetso, Thembela Mavuso, Bheki Ntshangase, Buhle Dlamini, Abhilash Sathyamoorthi, Bathusi Mathuba, Yves Mafulu, Keboletse Mokete, Lesego Ketumile, Kennedy Sichone, Keofentse Mathuba, LeToya Balebetse, Muambi Muyaya, Nancy Zwane, Nicholas Muriithi, Sibongile Mumanga, Thabo Diphoko, Thobile Jele, Marion Amujal, Ronald Oceng, and Thato Regonamanye. We also acknowledge Alison Eliot's contribution, Damalie Nakanjako, Victoria Bukirwa, Joshua Mandre, and Moses Kiiza from the MUIIplus UVRI/MRC/LSHTM AIDS Research Unit.

Conflicts of Interest: The authors declare no competing interests.

References

1. McLaren, P.J.; Carrington, M. The Impact of Host Genetic Variation on Infection with HIV-1. *Nat. Immunol.* **2015**, *16*, 577–583, doi:10.1038/ni.3147.
2. Cohen, O.J.; Paolucci, S.; Bende, S.M.; Daucher, M.; Moriuchi, H.; Moriuchi, M.; Cicala, C.; Davey, R.T.; Baird, B.; Fauci, A.S. CXCR4 and CCR5 Genetic Polymorphisms in Long-Term Nonprogressive Human Immunodeficiency Virus Infection: Lack of Association with Mutations Other than CCR5-Δ32. *J. Virol.* **1998**, *72*, 6215–6217, doi:10.1128/jvi.72.7.6215-6217.1998.
3. Walker, B.D.; Pereyra, F.; Jia, X.; McLaren, P.J.; Telenti, A.; Bakker, P.I.W. de; De Bakker, P.I.W.; Walker, B.D. The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation. *Science (80-.).* **2010**, *330*, 1551–1557, doi:10.1126/science.1195271.
4. Boudreau, J.E.; Hsu, K.C. Natural Killer Cell Education in Human Health and Disease. *Curr. Opin. Immunol.* **2018**, *50*, 102–111, doi:10.1016/j.coi.2017.11.003.
5. Kaseke, C.; Tano-Menka, R.; Senjobe, F.; Gaiha, G.D. The Emerging Role for CTL Epitope Specificity in HIV Cure Efforts. *J. Infect. Dis.* **2021**, *223*, S32–S37, doi:10.1093/infdis/jiaa333.
6. Pymm, P.; Tenzer, S.; Wee, E.; Weimershaus, M.; Burgevin, A.; Kollnberger, S.; Gerstoft, J.; Josephs, T.M.; Ladell, K.; McLaren, J.E.; et al. Epitope Length Variants Balance Protective Immune Responses and Viral Escape in HIV-1 Infection. *Cell Rep.* **2022**, *38*, doi:10.1016/j.celrep.2022.110449.
7. Kutsch, O.; Vey, T.; Kerkau, T.; Hünig, T.; Schimpl, A. HIV Type 1 Abrogates TAP-Mediated Transport of Antigenic Peptides Presented by MHC Class I. *AIDS Res. Hum. Retroviruses* **2002**, *18*, 1319–1325, doi:10.1089/088922202320886361.
8. Li, L.; Bouvier, M. Structures of HLA-A*1101 Complexed with Immunodominant Nonamer and Decamer HIV-1 Epitopes Clearly Reveal the Presence of a Middle, Secondary Anchor Residue. *J. Immunol.* **2004**, *172*, 6175–6184, doi:10.4049/jimmunol.172.10.6175.
9. Carlson, J.M.; Brumme, C.J.; Martin, E.; Listgarten, J.; Brockman, M.A.; Le, A.Q.; Chui, C.K.S.; Cotton, L.A.; Knapp, D.J.H.F.; Riddler, S.A.; et al. Correlates of Protective Cellular Immunity Revealed by Analysis of Population-Level Immune Escape Pathways in HIV-1. *J. Virol.* **2012**, *86*, 13202–13216, doi:10.1128/jvi.01998-12.
10. Stewart-Jones, G.B.E.; Gillespie, G.; Overton, I.M.; Kaul, R.; Roche, P.; McMichael, A.J.; Rowland-Jones, S.; Jones, E.Y. Structures of Three HIV-1 HLA-B*5703-Peptide Complexes and Identification of Related HLAs Potentially Associated with Long-Term Nonprogression. *J. Immunol.* **2005**, *175*, 2459–2468, doi:10.4049/jimmunol.175.4.2459.
11. Kyobe, S.; Mwesigwa, S.; Kisitu, G.P.; Farirai, J.; Katagirya, E.; Mirembe, A.N.; Ketumile, L.; Wayengera, M.; Katabazi, F.A.; Kigozi, E.; et al. Exome Sequencing Reveals a Putative Role for HLA-C*03:02 in Control of HIV-1 in African Pediatric Populations. *Front. Genet.* **2021**, *12*, 1–14, doi:10.3389/fgene.2021.720213.
12. Leslie, A.; Matthews, P.C.; Listgarten, J.; Carlson, J.M.; Kadie, C.; Ndung'u, T.; Brander, C.; Coovadia, H.; Walker, B.D.; Heckerman, D.; et al. Additive Contribution of HLA Class I Alleles in the Immune Control of HIV-1 Infection. *J. Virol.* **2010**, *84*, 9879–9888, doi:10.1128/jvi.00320-10.

13. Gijsbers, E.F.; Anton Feenstra, K.; Van Nuenen, A.C.; Navis, M.; Heringa, J.; Schuitemaker, H.; Kootstra, N.A. HIV-1 Replication Fitness of HLA-B*57/58:01 CTL Escape Variants Is Restored by the Accumulation of Compensatory Mutations in Gag. *PLoS One* **2013**, *8*, e81235, doi:10.1371/journal.pone.0081235.
14. Borghans, J.A.M.; Mølgaard, A.; de Boer, R.J.; Keşmir, C. HLA Alleles Associated with Slow Progression to AIDS Truly Prefer to Present HIV-1 P24. *PLoS One* **2007**, *2*, e920, doi:10.1371/journal.pone.0000920.
15. Chen, H.; Piechocka-Trocha, A.; Miura, T.; Brockman, M.A.; Julg, B.D.; Baker, B.M.; Rothchild, A.C.; Block, B.L.; Schneidewind, A.; Koibuchi, T.; et al. Differential Neutralization of Human Immunodeficiency Virus (HIV) Replication in Autologous CD4 T Cells by HIV-Specific Cytotoxic T Lymphocytes. *J. Virol.* **2009**, *83*, 3138–3149, doi:10.1128/jvi.02073-08.
16. Murakoshi, H.; Zou, C.; Kuse, N.; Akahoshi, T.; Chikata, T.; Gatanaga, H.; Oka, S.; Hanke, T.; Takiguchi, M. CD8+ T Cells Specific for Conserved, Cross-Reactive Gag Epitopes with Strong Ability to Suppress HIV-1 Replication. *Retrovirology* **2018**, *15*, 1–14, doi:10.1186/s12977-018-0429-y.
17. Adland, E.; Carlson, J.M.; Paioni, P.; Kløverpris, H.; Shapiro, R.; Ogwu, A.; Riddell, L.; Luzzi, G.; Chen, F.; Balachandran, T.; et al. Nef-Specific CD8+ T Cell Responses Contribute to HIV-1 Immune Control. *PLoS One* **2013**, *8*, 1–10, doi:10.1371/journal.pone.0073117.
18. Altfeld, M.; Addo, M.M.; Eldridge, R.L.; Yu, X.G.; Thomas, S.; Khatri, A.; Strick, D.; Phillips, M.N.; Cohen, G.B.; Islam, S.A.; et al. Vpr Is Preferentially Targeted by CTL During HIV-1 Infection. *J. Immunol.* **2001**, *167*, 2743–2752, doi:10.4049/jimmunol.167.5.2743.
19. Ferrando-Martínez, S.; Casazza, J.P.; Leal, M.; Machmach, K.; Muñoz-Fernández, M.Á.; Viciana, P.; Koup, R.A.; Ruiz-Mateos, E. Differential Gag-Specific Polyfunctional T Cell Maturation Patterns in HIV-1 Elite Controllers. *J. Virol.* **2012**, *86*, 3667–3674, doi:10.1128/jvi.07034-11.
20. Falivene, J.; Ghiglione, Y.; Laufer, N.; Socías, M.E.; Holgado, M.P.; Ruiz, M.J.; Maeto, C.; Figueroa, M.I.; Giavedoni, L.D.; Cahn, P.; et al. Th17 and Th17/Treg Ratio at Early HIV Infection Associate with Protective HIV-Specific CD8+ T-Cell Responses and Disease Progression. *Sci. Rep.* **2015**, *5*, 11511, doi:10.1038/srep11511.
21. Boulet, S.; Song, R.; Kamya, P.; Bruneau, J.; Shoukry, N.H.; Tsoukas, C.M.; Bernard, N.F. HIV Protective KIR3DL1 and HLA-B Genotypes Influence NK Cell Function Following Stimulation with HLA-Devoid Cells. *J. Immunol.* **2010**, *184*, 2057–2064, doi:10.4049/jimmunol.0902621.
22. Zappacosta, F.; Borrego, F.; Brooks, A.G.; Parker, K.C.; Coligan, J.E. Peptides Isolated from HLA-Cw*0304 Confer Different Degrees of Protection from Natural Killer Cell-Mediated Lysis. *Proc. Natl. Acad. Sci.* **1997**, *94*, 6313–6318, doi:10.1073/pnas.94.12.6313.
23. Hraber, P.; Seaman, M.S.; Bailer, R.T.; Mascola, J.R.; Montefiori, D.C.; Korber, B.T. Prevalence of Broadly Neutralizing Antibody Responses during Chronic HIV-1 Infection. *Aids* **2014**, *28*, 163–169, doi:10.1097/QAD.0000000000000106.
24. Payne, R.P.; Branch, S.; Kloverpris, H.; Matthews, P.C.; Koofhethile, C.K.; Strong, T.; Adland, E.; Leitman, E.; Frater, J.; Ndung'u, T.; et al. Differential Escape Patterns within the Dominant HLA-B*57:03-Restricted HIV Gag Epitope Reflect Distinct Clade-Specific Functional Constraints. *J. Virol.* **2014**, *88*, 4668–4678, doi:10.1128/jvi.03303-13.
25. Brackenridge, S.; Evans, E.J.; Toebes, M.; Goonetilleke, N.; Liu, M.K.P.; di Gleria, K.; Schumacher, T.N.; Davis, S.J.; McMichael, A.J.; Gillespie, G.M. An Early HIV Mutation within an HLA-B*57-Restricted T Cell Epitope Abrogates Binding to the Killer Inhibitory Receptor 3DL1. *J. Virol.* **2011**, *85*, 5415–5422, doi:10.1128/jvi.00238-11.
26. Yang, Y.; Sun, W.; Guo, J.; Zhao, G.; Sun, S.; Yu, H.; Guo, Y.; Li, J.; Jin, X.; Du, L.; et al. In Silico Design of a DNA-Based HIV-1 Multi-Epitope Vaccine for Chinese Populations. *Hum. Vaccines Immunother.* **2015**, *11*, 795–805, doi:10.1080/21645515.2015.1012017.
27. Haynes, B.F.; McElrath, M.J. Progress in HIV-1 Vaccine Development. *Curr. Opin. HIV AIDS* **2013**, *8*, 326–332, doi:10.1097/COH.0b013e328361d178.
28. Song, H.; Giorgi, E.E.; Ganusov, V. V.; Cai, F.; Athreya, G.; Yoon, H.; Carja, O.; Hora, B.; Hraber, P.; Romero-Severson, E.; et al. Tracking HIV-1 Recombination to Resolve Its Contribution to HIV-1 Evolution in Natural Infection. *Nat. Commun.* **2018**, *9*, 1928, doi:10.1038/s41467-018-04217-5.
29. Hraber, P.; Korber, B.T.; Lapedes, A.S.; Bailer, R.T.; Seaman, M.S.; Gao, H.; Greene, K.M.; McCutchan, F.; Williamson, C.; Kim, J.H.; et al. Impact of Clade, Geography, and Age of the Epidemic on HIV-1 Neutralization by Antibodies. *J. Virol.* **2014**, *88*, 12623–12643, doi:10.1128/jvi.01705-14.
30. Chikata, T.; Paes, W.; Akahoshi, T.; Partridge, T.; Murakoshi, H.; Gatanaga, H.; Ternette, N.; Oka, S.; Borrow, P.; Takiguchi, M. Identification of Immunodominant HIV-1 Epitopes Presented by HLA-C*12:02, a Protective Allele, Using an Immunopeptidomics Approach. *J. Virol.* **2019**, *93*, doi:10.1128/JVI.00634-19.
31. Bugembe, D.L.; Ekii, A.O.; Ndembi, N.; Serwanga, J.; Kaleebu, P.; Pala, P. Computational MHC-I Epitope Predictor Identifies 95% of Experimentally Mapped HIV-1 Clade A and D Epitopes in a Ugandan Cohort. *BMC Infect. Dis.* **2020**, *20*, 172, doi:10.1186/s12879-020-4876-4.
32. Hsin, J.; Arkhipov, A.; Yin, Y.; Stone, J.E.; Schulten, K. Using VMD: An Introductory Tutorial. *Curr. Protoc. Bioinforma.* **2008**, *24*, doi:10.1002/0471250953.bi0507s24.

33. Knapp, B.; van der Merwe, P.A.; Dushek, O.; Deane, C.M. MHC Binding Affects the Dynamics of Different T-Cell Receptors in Different Ways. *PLOS Comput. Biol.* **2019**, *15*, e1007338, doi:10.1371/journal.pcbi.1007338.
34. Li, X.; Singh, N.K.; Collins, D.R.; Ng, R.; Zhang, A.; Lamothe-Molina, P.A.; Shahinian, P.; Xu, S.; Tan, K.; Piechocka-Trocha, A.; et al. Molecular Basis of Differential HLA Class I-Restricted T Cell Recognition of a Highly Networked HIV Peptide. *Nat. Commun.* **2023**, *14*, 2929, doi:10.1038/s41467-023-38573-8.
35. Valdés-Tresanco, M.S.; Valdés-Tresanco, M.E.; Valiente, P.A.; Moreno, E. Gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. *J. Chem. Theory Comput.* **2021**, *17*, 6281–6291, doi:https://doi.org/10.1021/acs.jctc.1c00645.
36. Ndongala, M.L.; Kamya, P.; Boulet, S.; Peretz, Y.; Rouleau, D.; Tremblay, C.; Leblanc, R.; Côté, P.; Baril, J.; Thomas, R.; et al. Changes in Function of HIV-Specific T-Cell Responses with Increasing Time from Infection. *Viral Immunol.* **2010**, *23*, 159–168, doi:10.1089/vim.2009.0084.
37. Excler, J.L.; Robb, M.L.; Kim, J.H. Prospects for a Globally Effective HIV-1 Vaccine. *Vaccine* **2015**, *33*, D4–D12, doi:10.1016/j.vaccine.2015.03.059.
38. Korber, B.; Fischer, W. T Cell-Based Strategies for HIV-1 Vaccines. *Hum. Vaccin. Immunother.* **2020**, *16*, 713–722, doi:10.1080/21645515.2019.1666957.
39. Shin, S.Y. Recent Update in HIV Vaccine Development. *Clin. Exp. Vaccine Res.* **2016**, *5*, 6, doi:10.7774/cevr.2016.5.1.6.
40. Zipeto, D.; Beretta, A. HLA-C and HIV-1: Friends or Foes? *Retrovirology* **2012**, *9*, 39, doi:10.1186/1742-4690-9-39.
41. Buranapraditkun, S.; Hempel, U.; Pitakpolrat, P.; Allgaier, R.L.; Thantivorasit, P.; Lorenzen, S.-I.; Sirivichayakul, S.; Hildebrand, W.H.; Altfeld, M.; Brander, C.; et al. A Novel Immunodominant CD8+ T Cell Response Restricted by a Common HLA-C Allele Targets a Conserved Region of Gag HIV-1 Clade CRF01_AE Infected Thais. *PLoS One* **2011**, *6*, e23603, doi:10.1371/journal.pone.0023603.
42. Hashimoto, M.; Akahoshi, T.; Murakoshi, H.; Ishizuka, N.; Oka, S.; Takiguchi, M. CTL Recognition of HIV-1-Infected Cells via Cross-Recognition of Multiple Overlapping Peptides from a Single 11-Mer Pol Sequence. *Eur. J. Immunol.* **2012**, *42*, 2621–2631, doi:10.1002/eji.201242483.
43. Perdomo-Celis, F.; Taborda, N.A.; Rugeles, M.T. CD8+ T-Cell Response to HIV Infection in the Era of Antiretroviral Therapy. *Front. Immunol.* **2019**, *10*, doi:10.3389/fimmu.2019.01896.
44. Addo, M.M.; Yu, X.G.; Rathod, A.; Cohen, D.; Eldridge, R.L.; Strick, D.; Johnston, M.N.; Corcoran, C.; Wurcel, A.G.; Fitzpatrick, C.A.; et al. Comprehensive Epitope Analysis of Human Immunodeficiency Virus Type 1 (HIV-1)-Specific T-Cell Responses Directed against the Entire Expressed HIV-1 Genome Demonstrate Broadly Directed Responses, but No Correlation to Viral Load. *J. Virol.* **2003**, *77*, 2081–2092, doi:10.1128/JVI.77.3.2081-2092.2003.
45. Rosás-Umbert, M.; Gunst, J.D.; Pahus, M.H.; Olesen, R.; Schleimann, M.; Denton, P.W.; Ramos, V.; Ward, A.; Kinloch, N.N.; Copertino, D.C.; et al. Administration of Broadly Neutralizing Anti-HIV-1 Antibodies at ART Initiation Maintains Long-Term CD8+ T Cell Immunity. *Nat. Commun.* **2022**, *13*, 6473, doi:10.1038/s41467-022-34171-2.
46. Joglekar, A. V.; Liu, Z.; Weber, J.K.; Ouyang, Y.; Jeppson, J.D.; Noh, W.J.; Lamothe-Molina, P.A.; Chen, H.; Kang, S. gu; Bethune, M.T.; et al. T Cell Receptors for the HIV KK10 Epitope from Patients with Differential Immunologic Control Are Functionally Indistinguishable. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 1877–1882, doi:10.1073/pnas.1718659115.
47. Tassiopoulos, K.; Landay, A.; Collier, A.C.; Connick, E.; Deeks, S.G.; Hunt, P.; Lewis, D.E.; Wilson, C.; Bosch, R. CD28-Negative CD4+ and CD8+ T Cells in Antiretroviral Therapy–Naive HIV-Infected Adults Enrolled in Adult Clinical Trials Group Studies. *J. Infect. Dis.* **2012**, *205*, 1730–1738, doi:10.1093/infdis/jis260.
48. Ziegler, M.C.; Nelde, A.; Weber, J.K.; Schreitmüller, C.M.; Martrus, G.; Huynh, T.; Bunders, M.J.; Lunemann, S.; Stevanovic, S.; Zhou, R.; et al. HIV-1 Induced Changes in HLA-C*03:04-Presented Peptide Repertoires Lead to Reduced Engagement of Inhibitory Natural Killer Cell Receptors. *AIDS* **2020**, *34*, 1713–1723, doi:10.1097/QAD.0000000000002596.
49. Cohen, G.B.; Gandhi, R.T.; Davis, D.M.; Mandelboim, O.; Chen, B.K.; Strominger, J.L.; Baltimore, D. The Selective Downregulation of Class I Major Histocompatibility Complex Proteins by HIV-1 Protects HIV-Infected Cells from NK Cells. *Immunity* **1999**, *10*, 661–671, doi:10.1016/S1074-7613(00)80065-5.
50. Mann, J.K.; Byakwaga, H.; Kuang, X.T.; Le, A.Q.; Brumme, C.J.; Mwimanzi, P.; Omarjee, S.; Martin, E.; Lee, G.Q.; Baraki, B.; et al. Ability of HIV-1 Nef to Downregulate CD4 and HLA Class I Differs among Viral Subtypes. *Retrovirology* **2013**, *10*, 100, doi:10.1186/1742-4690-10-100.
51. Fellay, J.; Shianna, K. V; Ge, D.; Colombo, S.; Ledergerber, B.; Weale, M.; Zhang, K.; Gumbs, C.; Castagna, A.; Cossarizza, A.; et al. A Whole-Genome Association Study of Major Determinants for Host Control of HIV-1. *Science (80-.).* **2007**, *317*, 944–947, doi:10.1126/science.1143767.
52. Kulkarni, S.; Savan, R.; Qi, Y.; Gao, X.; Yuki, Y.; Bass, S.E.; Martin, M.P.; Hunt, P.; Deeks, S.G.; Telenti, A.; et al. Differential MicroRNA Regulation of HLA-C Expression and Its Association with HIV Control. *Nature* **2011**, *472*, 495–498, doi:10.1038/nature09914.

53. Thomas, R.; Apps, R.; Qi, Y.; Gao, X.; Male, V.; O'hUigin, C.; O'Connor, G.; Ge, D.; Fellay, J.; Martin, J.N.; et al. HLA-C Cell Surface Expression and Control of HIV/AIDS Correlate with a Variant Upstream of HLA-C. *Nat. Genet.* **2009**, *41*, 1290–1294, doi:10.1038/ng.486.
54. Koita, O.A.; Dabitao, D.; Mahamadou, I.; Tall, M.; Dao, S.; Tounkara, A.; Guiteye, H.; Noumsi, C.; Thiero, O.; Kone, M.; et al. Confirmation of Immunogenic Consensus Sequence HIV-1 T-Cell Epitopes in Bamako, Mali and Providence, Rhode Island. *Hum. Vaccin.* **2006**, *2*, 119–128, doi:10.4161/hv.2869.
55. Goepfert, P.A.; Bansal, A.; Edwards, B.H.; Ritter, G.D.; Tellez, I.; McPherson, S.A.; Sabbaj, S.; Mulligan, M.J. A Significant Number of Human Immunodeficiency Virus Epitope-Specific Cytotoxic T Lymphocytes Detected by Tetramer Binding Do Not Produce Gamma Interferon. *J. Virol.* **2000**, *74*, 10249–10255, doi:10.1128/JVI.74.21.10249-10255.2000.
56. Retshabile, G.; Mlotshwa, B.C.; Williams, L.; Mwesigwa, S.; Mboowa, G.; Huang, Z.; Rustagi, N.; Swaminathan, S.; Katagirya, E.; Kyobe, S.; et al. Whole-Exome Sequencing Reveals Uncaptured Variation and Distinct Ancestry in the Southern African Population of Botswana. *Am. J. Hum. Genet.* **2018**, *102*, 731–743, doi:10.1016/j.ajhg.2018.03.010.
57. Bbosa, N.; Kaleebu, P.; Ssemwanga, D. HIV Subtype Diversity Worldwide. *Curr. Opin. HIV AIDS* **2019**, *14*, 153–160, doi:10.1097/COH.0000000000000534.
58. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612, doi:10.1002/jcc.20084.
59. Antunes, D.A.; Moll, M.; Devaurs, D.; Jackson, K.R.; Lizée, G.; Kavraki, L.E. DINC 2.0: A New Protein-Peptide Docking Webserver Using an Incremental Approach. *Cancer Res.* **2017**, *77*, e55–e57, doi:10.1158/0008-5472.CAN-17-0511.
60. El-Hachem, N.; Haibe-Kains, B.; Khalil, A.; Kobeissy, F.H.; Nemer, G. AutoDock and AutoDockTools for Protein-Ligand Docking: Beta-Site Amyloid Precursor Protein Cleaving Enzyme 1(BACE1) as a Case Study. In *Methods in Molecular Biology*; 2017; pp. 391–403.
61. Kadukova, M.; Grudinin, S. Convex-PL: A Novel Knowledge-Based Potential for Protein-Ligand Interactions Deduced from Structural Databases Using Convex Optimization. *J. Comput. Aided. Mol. Des.* **2017**, *31*, 943–958, doi:10.1007/s10822-017-0068-8.
62. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Meng, E.C.; Couch, G.S.; Croll, T.I.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers. *Protein Sci.* **2021**, *30*, 70–82, doi:10.1002/pro.3943.
63. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25, doi:10.1016/j.softx.2015.06.001.
64. Best, R.B.; Zhu, X.; Shim, J.; Lopes, P.E.M.; Mittal, J.; Feig, M.; MacKerell, A.D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ 1 and χ 2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273, doi:10.1021/ct300400x.
65. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2009**, NA-NA, doi:10.1002/jcc.21367.
66. Hanwell, M.D.; Curtis, D.E.; Lonie, D.C.; Vandermeersch, T.; Zurek, E.; Hutchison, G.R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminform.* **2012**, *4*, 17, doi:10.1186/1758-2946-4-17.
67. Lapointe, H.R.; Dong, W.; Lee, G.Q.; Bangsberg, D.R.; Martin, J.N.; Mocello, A.R.; Boum, Y.; Karakas, A.; Kirkby, D.; Poon, A.F.Y.; et al. HIV Drug Resistance Testing by High-Multiplex "Wide" Sequencing on the MiSeq Instrument. *Antimicrob. Agents Chemother.* **2015**, *59*, 6824–6833, doi:10.1128/AAC.01490-15.
68. Tachbele, E.; Kyobe, S.; Katabazi, F.A.; Kigozi, E.; Mwesigwa, S.; Joloba, M.; Messele, A.; Amogne, W.; Legesse, M.; Pieper, R.; et al. Genetic Diversity and Acquired Drug Resistance Mutations Detected by Deep Sequencing in Virologic Failures among Antiretroviral Treatment Experienced Human Immunodeficiency Virus-1 Patients in a Pastoralist Region of Ethiopia. *Infect. Drug Resist.* **2021**, *Volume 14*, 4833–4847, doi:10.2147/IDR.S337485.
69. Pineda-Peña, A.-C.; Faria, N.R.; Imbrechts, S.; Libin, P.; Abecasis, A.B.; Deforche, K.; Gómez-López, A.; Camacho, R.J.; de Oliveira, T.; Vandamme, A.-M. Automated Subtyping of HIV-1 Genetic Sequences for Clinical and Surveillance Purposes: Performance Evaluation of the New REGA Version 3 and Seven Other Tools. *Infect. Genet. Evol.* **2013**, *19*, 337–348, doi:10.1016/j.meegid.2013.04.032.