Virus Research 273 (2019) 197763

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Effect of diversity in gp41 membrane proximal external region of primary HIV-1 Indian subtype C sequences on interaction with broadly neutralizing antibodies 4E10 and 10E8

Jyoti Sutar[a], Varsha Padwal[a], Archana Sonawani[b], Vidya Nagar[c], Priya Patil[c], Bhalachandra Kulkarni[d], Nitin Hingankar[e], Suprit Deshpande[e], Susan Idicula-Thomas[b], Dhanashree Jagtap[d], Jayanta Bhattacharya[e], Atmaram Bandivdekar[a,*], Vainav Patel[a,*]

[a] Department of Biochemistry, National Institute for Research in Reproductive Health (ICMR-NIRRH), Parel, Mumbai, India
[b] ICMR Biomedical Informatics Centre, National Institute for Research in Reproductive Health (ICMR-NIRRH), Parel, Mumbai, India
[c] Department of Medicine, Grant Government Medical College, Byculla, Mumbai, India
[d] Department of Structural Biology, National Institute for Research in Reproductive Health (ICMR-NIRRH), Parel, Mumbai, India
[e] HIV Vaccine Translational Research Laboratory, Translational Health Science and Technology Institute, NCR Biotech Science Cluster, Faridabad, Haryana, India

ARTICLE INFO

Keywords:
HIV-1 clade C
MPER
India
Deep sequencing
Broadly neutralizing antibodies
10E8

ABSTRACT

Human Immunodeficiency Virus-1 Clade C (HIV-1C) dominates the AIDS epidemic in India, afflicting 2.1 million individuals within the country and more than 15 million people worldwide. Membrane proximal external region (MPER) is an attractive target for broadly neutralizing antibody (bNAb) based therapies. However, information on MPER sequence diversity from India is meagre due to limited sampling of primary viral sequences. In the present study, we examined the variation in MPER of HIV-1C from 24 individuals in Mumbai, India by high throughput sequencing of uncultured viral sequences. Deep sequencing of MPER (662-683; HXB2 envelope amino acid numbering) allowed quantification of intra-individual variation up to 65% at positions 662, 665, 668, 674 and 677 within this region. These variable positions included contact sites targeted by bNAbs 2F5, Z13e1, 4E10 as well as 10E8. Both major and minor epitope variants i.e. 'haplotypes' were generated for each sample dataset. A total of 23, 34 and 25 unique epitope haplotypes could be identified for bNAbs 2F5, Z13e1 and 4E10/10E8 respectively. Further analysis of 4E10 and 10E8 epitopes from our dataset and meta-analysis of previously reported HIV-1 sequences from India revealed 26 epitopes (7 India-specific), heretofore untested for neutralization sensitivity. Peptide-Ab docking predicted 13 of these to be non-binding to 10E8. ELISA, Surface Plasmon Resonance and peptide inhibition of HIV-1 neutralization assays were then performed which validated predicted weak/non-binding interactions for peptides corresponding to six of these epitopes. These results highlight the under-representation of 10E8 non-binding HIV-1C MPER sequences from India. Our study thus underscores the need for increased surveillance of primary circulating envelope sequences for development of efficacious bNAb-based interventions in India.

1. Introduction

As per *UNAIDS* 2018 estimates, 36.9 million people are living with HIV/AIDS (PLHIV) worldwide (UNAIDS, 2018). Presently, approximately 2.1 million people in India are infected with clade C, the predominant clade of HIV-1 that causes 46.6% infections worldwide (Bhattacharya, 2018; Hemelaar et al., 2019; NACO, 2017). In addition to the inherent lack of proofreading activity of the reverse transcriptase enzyme, HIV uses various strategies such as, base substitution, recombination and accumulation of insertions and deletions (indels) for generation of viral variants, or 'Quasispecies' (Korber et al., 2001). The resultant antigenic variability facilitates continual escape from immune as well as antiviral drug pressure (Jadhav et al., 2011; Mcmichael et al., 2009). A successful curative strategy for HIV infection has thus remained elusive in spite of multiple attempts to induce protective immunity (Ajbani et al., 2015; Alter and Barouch, 2018; Patel et al., 2013; Shapiro, 2019).

HIV envelope glycoprotein is a heterodimer of non-covalently associated gp120 (Surface unit) and gp41 (transmembrane unit). The process of viral ingress into target cells is initiated by gp120 through

* Corresponding authors.
E-mail addresses: batmaram@gmail.com (A. Bandivdekar), patelv@nirrh.res.in (V. Patel).

https://doi.org/10.1016/j.virusres.2019.197763
Received 22 July 2019; Received in revised form 19 September 2019; Accepted 20 September 2019
Available online 22 September 2019
0168-1702/

binding with CD4 and a co-receptor, followed by fusion and entry through gp41 (Blumenthal et al., 2012). While enabling immune evasion through sequence variation, HIV envelope remains a 'site of vulnerability' due to structural constraints in functionally important conserved domains (Kwong et al., 2011). Antibodies targeting these domains with the ability to potently neutralize a broad range of cross-clade viruses, also known as broadly neutralizing antibodies (bNAbs) have been at the center of anti-HIV therapeutic efforts including vaccine development (Kumar et al., 2018; McCoy, 2018). In fact, as of 2017, 90 bNAbs had been reported, of which 12 are already being clinically evaluated (Caskey et al., 2019; Possas et al., 2018) as adjuncts/alternatives to anti-retroviral therapy.

The ectodomain region of gp41 has been well studied across clades in attempts to design bNAb based vaccines (Burton et al., 2012a; Salzwedel et al., 1999). Previously described bNAbs against Membrane proximal external region (MPER) such as 4E10, 2F5, Z13e1 etc. have been shown to be either lacking breadth or being highly autoreactive (Binley et al., 2004; Buchacher et al., 1994; Haynes et al., 2005; Nelson et al., 2007; Stiegler et al., 2001; Zwick et al., 2001). A study in 2012 reported the isolation of an MPER specific antibody 10E8 which was shown to exhibit broad neutralization breadth (~98%) as well as high potency (Huang et al., 2012a). The discovery of 10E8 bNAb has re-invigorated interest in MPER as a target for bNAbs and in the effort to discover 10E8 like bNAbs. Subsequently, multiple studies have been carried out to analyze various aspects of this antibody in order to further its development for potential therapeutic interventions (Georgiev et al., 2014; Irimia et al., 2017; Pegu et al., 2014; Rujas et al., 2015).

Sequences isolated from India for the gp41 region have previously been reported to cluster distinctly from sequences from Africa, another major geographical region for subtype C epidemic (Agnihotri et al., 2006). Los Alamos National Laboratory HIV sequence database (LANL-HIV) contains 290 sequences reported from India that fully/partially cover gp41 region. Furthermore, in the period from 1990 to 2018, an average of only 9 sequences (Median 5) have been reported each year. Of these, 286 sequences consist of complete MPER domain. The present study was undertaken to extend the diversity documentation observed in MPER from India. High throughput sequencing based variation analysis was performed on circulating viral sequences/proviruses isolated from 24 HIV-1 clade C infected individuals at different stages of disease progression from Mumbai, India. The influence of diversity in MPER, identified through deep sequencing as well as meta-analysis, on interaction with bNAbs 10E8 and 4E10, was also assessed *in silico* and *in vitro*.

2. Results

2.1. Clinical features of participants

All the clinical specimens were collected from chronically HIV-1 clade C infected individuals (Table 1). Of the 24, 16 individuals were receiving ART while 8 were ART naïve. There were no statistically significant differences in the CD4 counts ($p = 0.9404$) and viral loads ($p = 0.2344$) of the two groups (unpaired *t*-test with Welch's correction). However, variances of both groups were significantly different due to some of the ART receiving individuals suspected to be failing their therapy (CD4 count: $p = 0.0178$, Viral load: $p = 0.0002$). Detailed description of clinical features has been provided in Supplementary file 1. Blood derived high throughput sequencing (HTS) data for MPER was obtained from all individuals. Matched datasets of both plasma circulating virions and proviruses were obtained from 1 individual from each of the ART naïve (S1) and ART receiving (S2) categories. While MPER HTS data was obtained from only circulating plasma viruses for S3 and S4, all other datasets (S5-S24) were generated from proviral DNA. To account for errors introduced during different protocols followed for amplification and sequencing, technical replicates of S5, S6 and S8 were used for threshold calculation and 5.87% was set as the threshold for consideration of variation. Consensus sequences generated from each NGS dataset were confirmed to be from HIV-1 subtype C with Recombinant Identification Program (RIP) hosted on LANL HIV database (Siepel et al., 1995).

Table 1
Study particpant characteristics.

Designation	Age	Sex	CD4 count (cells/mm^3)	Viral load (copies/mL)	ART status
S1	8	M	966	45034	AN
S2	40	F	185	26967	AR
S3	8	M	864	146462	AN
S4	43	F	427	68856	AN
S5	40	F	440	2749	AN
S6	45	M	864	Undetectable	AR
S7	9	M	1324	Undetectable	AR
S8	35	M	230	101032	AN
S9	40	M	162	304046	AR
S10	40	M	372	184723	AR
S11	55	M	124	34062	AR
S12	48	F	269	7293	AR
S13	36	F	179	29194	AR
S14	48	M	195	< 34	AR
S15	51	M	194	97	AR
S16	45	M	185	848	AR
S17	45	F	202	132788	AR
S18	42	F	2585	Undetectable	AR
S19	37	F	1061	Undetectable	AR
S20	43	M	714	< 34	AR
S21	45	F	424	< 34	AR
S22	28	M	534	197054	AN
S23	45	M	612	6585	AN
S24	43	M	334	861484	AN

AN: ART naïve, AR: ART receiving.

2.2. MPER Heatmap and haplotype generation

MPER domain contains epitopes for four well studied broadly neutralizing antibodies: 2F5 (662-667: HXB2 envelope numbering), Z13e1 (666-677), 4E10 (671-683) and 10E8 (671-683) (Burton et al., 2012a). CATNAP (Compile, Analyze and Tally NAb Panels) is a comprehensive platform/web server that provides integrated datasets of neutralization (IC50 and IC80 values), pseudoviral sequences as well as all relevant information from all published studies on broadly neutralizing antibodies (Yoon et al., 2015). Data for the key epitope contact residues involved in the interaction of these antibodies was retrieved through CATNAP webserver of the LANL immunology database (Yoon et al., 2015) and literature survey (Bryson et al., 2009; Huang et al., 2012a; Nelson et al., 2007; Song et al., 2009). Position-wise amino acid variation data specific for this domain was used for generation of heat maps (Fig. 1A). In spite of being highly conserved, MPER was observed to have up to 65% variation at several residues. As expected, in both matched datasets, as well as in the cumulative datasets (Marked plasma_all and provirus_all), circulating viral RNA had less variation than proviral DNA, latter being the archival nucleic acid. Positions 663, 669, 673, 675, 678, 679 and 681 were most conserved, while positions 662, 665, 668, 674 and 677 demonstrated variation. Position 674 was observed to be most variable across datasets. To assess the amino acid variation at contact sites in the bNAb epitopes a 'position weight matrix (PWM)' was created from all the NGS datasets pooled together. Sequence logos were generated based on these PWMs for 2F5 (Fig. 1B), Z13e1 (Fig. 1C) and 4E10/10E8 (Fig. 1D) epitopes. Epitope for 2F5 was observed to be conserved with predominant replacement of the DKW core with DSW. In case of the Z13e1 epitope, both the heat map and the sequence logos indicated amino acid positions 671 and 674 to be variable in the dataset. Epitope contact sites reported to be important for 4E10 mediated neutralization are positions 671, 672, 673 and 676 (Song et al., 2009; Yoon et al., 2015). As indicated by the heat map and

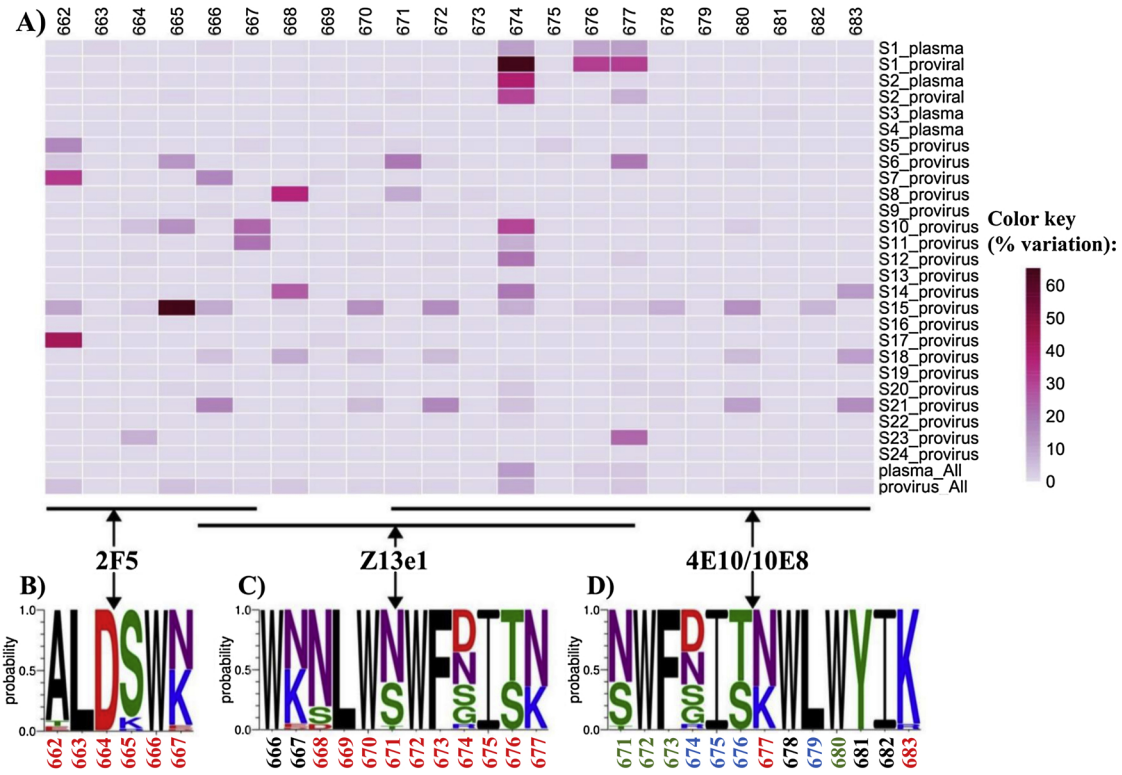

Fig. 1. Variation profile associated with bNAb epitopes in MPER domain.
A) Heat maps were generated for the MPER (HXB2 numbering: 662-683, x axis) based on the high throughput sequencing data wherein each pixel represents 1 amino acid position. Each pixel has been colored based on the amino acid variability at that position from light to dark color as depicted in the color key. Each row (y axis) represents dataset from 1 clinical specimen with last two rows representing cumulative data from plasma virions and proviruses across individuals. Epitopes for following broadly neutralizing antibodies have been highlighted 2F5 (662-667), Z13e1 (666-677), 4E10 (671-683), 10E8 (671- 683). Sequence logos depicting cumulative amino acids vs probability of their presence in the specific epitope position were generated for epitopes of bNAbs 2F5 (B), Z13e1 (C) and 4E10/10E8 (D) respectively. Contact sites for the epitopes have been indicated with position names highlighted in red color (residues important for 4E10: blue, 10E8: red, both 4E10 and 10E8: green) (Width: Double column, Color).

sequence logos, these positions demonstrated variation up to 20% intra-individually but < 10% inter-individually. Similar observations were made about 10E8 contact residues (672, 673, 680, 683) being relatively conserved inter-individually. Since minor variants are capable of altering the outcomes of therapeutic interventions, impact of this variation was further assessed. Analysis of conformational effects of position specific variation on epitope structure and consequently binding to bNAbs was undertaken through generation of haplotypes for each of the epitopes from NGS data using a custom haplotype analysis script, HaploCount (described in detail in materials and methods). A total of 38 haplotypes were constructed for 2F5 with 23 unique epitope sequences (Supplementary file 2). Upon comparison with LANL CATNAP database, 3 haplotypes sequences were found to be untested while 35 matched tested pseudovirus sequences. In case of Z13e1, 49 epitope haplotypes were constructed (34 unique). Upon comparison with LANL CATNAP database only 2 epitopes matched Z13e1 sensitive pseudoviruses, 15 matched resistant pseudoviruses and 32 were untested i.e. they did not match any tested pseudovirus sequences (Supplementary file 2). A total of 45 4E10/10E8 epitope haplotypes were generated in the present analysis, 25 of which were found to be unique (Table 2). In the matched datasets with sequencing performed for both plasma derived RNA and proviral DNA, as expected, RNA haplotypes were found to be subsets of the respective proviral DNA. Deep (1500-30000x) sequencing revealed 6 haplotype sequences at frequencies of 8.1–91% (indicated in bold as Test1-5 and Test24) that have not been tested for 10E8 neutralization sensitivity. While Epitope haplotypes Test1-5 have never been reported in the database as originating from India, all except Test2 have been reported in the LANL database from HIV-1 clades including but not limited to, clades A1, B, C, D, F and G. Further survey of sequences reported from India in LANL database revealed 20 epitope sequences that have not been tested for 10E8 neutralization (indicated as Test6-26 in Table 3). Evaluation of these sequences could prove to be of significant value in predicting prevalence of bNAb non binding primary viral sequences. Therefore, an *in silico* docking analysis was performed to predict the binding ability of the untested haplotypes to 10E8 as well as 4E10.

2.3. Docking analysis

Docking analysis was performed with two different algorithms viz., ZDOCK-RDOCK and Rosetta framework for 10E8 and with Rosetta framework for 4E10.

2.3.1. ZDOCK-RDOCK

The docking algorithm and the parameters were validated by re-docking the structure of antibody 10E8 Fab with the epitope (residues 671-683) of gp41 (P chain). Redocking correctly reproduced the bound crystal structure (PDB: 4G6F) with an RMSD of 0.17 Å. The performance of the docking algorithm was assessed using the experimentally determined IC50 values reported in the CATNAP webserver and the computed RDOCK energies for the 14 positive and 5 negative controls. RDOCK energy of −6 kcal/mol was used as an empirically-fixed threshold to distinguish between potential binders and non-binders (Fig. 2A). The effect of the residue change/s on antibody binding was predicted as expected for 11 of the 14 positive controls and 3 of the 5 negative controls. Therefore, the prediction sensitivity of Zdock

Table 2
Haplotype distribution.

	Dataset	Haplotype	Frequency (%)	No. of reads	Predicted interaction with 10E8	Matching pseudovirus ID
1	S1_Plasma	NWFSITKWLWYIK*	70.06	557	Sensitive	16055_2_3
2		NWFNISNWLWYIK*	7.3	58	Sensitive	1170887_08
3	S1_provirus	NWFGITKWLWYIK	28.14	1032	Sensitive	CNE17
4		NWFSITKWLWYIK*	24.46	897	Sensitive	16055_2_3
5		NWFNISNWLWYIK*	23.73	870	Sensitive	1170887_08
6	S2_Plasma	NWFDITNWLWYIK#	51.31	1746	Sensitive	247_23
7		NWFNITNWLWYIK#	30.5	1038	Sensitive	25710_2_43
8	S2_Provirus	NWFDITNWLWYIK#	48.76	4505	Sensitive	247_23
9		NWFNITNWLWYIK#`	24.81	2292	Sensitive	25710_2_43
10		NWFDITKWLWYIK	6.66	615	Sensitive	45_01DG5
11	S3_Plasma	SWFDISNWLWYIK	80.98	6841	Sensitive	BB201_B42
12	S4_Plasma	SWFDISNWLWYIK	79.41	3652	Sensitive	BB201_B42
13	S5_Provirus	SWFNITNWLWYIK	75.57	6284	Sensitive	CA327_D2_2
14	S6_Provirus	SWFNITNWLWYIK	63.36	4664	Sensitive	CA327_D2_2
15		NWFNITKWLWYIK	15.31	1127	Sensitive	25710_2_43
16	S7_Provirus	SWFNITNWLWYIK	78.43	4113	Sensitive	CA327_D2_2
17	S8_Provirus	TWFDITNWLWYIK	72.06	10398	Sensitive	CAAN5342
18		NWFDITNWLWYIK	7.61	1098	Sensitive	247_23
19	S9_Provirus	SWFDISNWLWYIK	97.97	22821	Sensitive	BB201_B42
20	S10_Provirus	NWFSITNWLWYIK	65.71	5510	Sensitive	1394_C9G1
21		NWFNITNWLWYIK	29.36	2462	Sensitive	235_47
22	S11_Provirus	**NWFRITKWLWYIK**	**90.51**	**20017**	**Untested**	**NA**
23		NWFGITKWLWYIK	8.08	1786	Sensitive	CNE17
24	S12_Provirus	SWFDISKWLWYIK	76.09	14260	Sensitive	CH181_12
25		**SWFEISKWLWYIK**	**18.52**	**3470**	**Untested**	**NA**
26	S13_Provirus	NWFSITKWLWYIK	78.94	15476	Sensitive	16055_2_3
27		NWFGITKWLWYIK	19.94	3909	Sensitive	CNE17
28	S14_Provirus	NWFSISNWLWYIK	63.21	13075	Sensitive	20803520
29		NWFNISNWLWYIK	13.29	2749	Sensitive	1170887_08
30		NWFSISNWLWYIR	6.87	1421	Sensitive	20965238
31	S15_Provirus	NWFGITKWLWYIK	87.67	6018	Sensitive	CNE17
32		**NWFGITKWLWYVK**	**8.13**	**558**	**Untested**	**NA**
33	S16_Provirus	NWFNISNWLWYIK	98.9	16088	Sensitive	1170887_08
34	S17_Provirus	NWFSITKWLWYIK	98.58	21578	Sensitive	16055_2_3
35	S18_Provirus	SWFNISNWLWYIK	79.06	16117	Sensitive	703010217_B6
36		NWFDITKWLWYIK	11.99	2444	Sensitive	0439_V5_C1
37	S19_Provirus	NWFNITNWLWYIK	97.84	14751	Sensitive	235_47
38	S20_Provirus	NWFDITKWLWYIK	95.18	13515	Sensitive	0439_V5_C1
39	S21_Provirus	SWFDISNWLWYIR	92.99	9099	Sensitive	3415_V1_C1
40	S22_Provirus	NWFGISNWLWYIK	96.53	13217	Sensitive	BS208_B1
41	S23_Provirus	NWFDISKWLWYIK	75.38	12407	Sensitive	211_9
42		**NWFDISSWLWYIK**	**15.53**	**2556**	**Untested**	**NA**
43		NWFDISNWLWYIK	7.84	1290	Sensitive	0330_V4_C3
44	S24_Provirus	**SWFGITNWLWYIK**	**90.36**	**16697**	**Untested**	**NA**
45		**SWFGITNWLWYIR**	**7.78**	**1437**	**Untested**	**NA**

No. of reads column refers to actual number of contributing filtered reads that cover the entire epitope. Epitope haplotypes untested in neutralization have been highlighted in bold font: Epitope haplotypes detected from the same individual RNA and proviral have been indicated with '*' and '#' notations.

approach was found to be 78.57%, specificity was 60% and accuracy was 73.68%. The RDOCK energies computed for the 21 test haplotypes are shown in Table 3.

2.3.2. Rosetta framework

Rosetta framework 2016.13.58602_bundle was used for protein peptide docking analysis of 10E8 and 4E10 with 64 control haplotypes and 26 test haplotypes. Binding energy per unit area was used as a parameter for this analysis along with the binding energy of the complex, to ensure that the interface of the docked pose included high quality molecular contacts instead of many low-quality contacts across the interface. Support vector machine (SVM)-based training was performed for 30 control haplotypes (27 binders and 3 non binders). The predictions based on this training dataset were then extended to test classification of all 64 known haplotypes. Predictions made for 58 of the 59 positive controls of 10E8 (Fig. 2B) and all 60 positive controls of 4E10 (Fig. 2C) were as expected. Furthermore, predictions of 3 of the 5 negative controls for 10E8 as well as 2 out of 4 for 4E10 were also as expected. Therefore, Rosetta approach for 10E8 was found to be 98.33% sensitive, 60% specific and 95.31% accurate, while that for 4E10 was 96.7% sensitive, 50% specific and 96.87% accurate. The binding energies computed for the 26 neutralization untested haplotypes are shown in Table 3.

2.3.3. Docking prediction for test epitope haplotypes

Overall 10 out of 26 haplotypes were predicted to be sensitive to both the bNAbs by all 3 approaches. Out of 21 epitope haplotype sequences analyzed by Zdock approach, 9 were found to be non-binding to 10E8, whereas Rosetta approach predicted 6 out of 26 10E8 epitopes to be non-binding. Epitope haplotypes Test20, and Test24 (Table 3) were predicted to be non-binding to 10E8 by both approaches used. Four out of 26 haplotypes were also reported to be non-binding to 4E10 by Rosetta. Interestingly, of the 16 non-binding epitope haplotypes reported, 13 were found to be reported earlier from countries with HIV-1 clade C epidemic including India, while 7 of them have been reported exclusively from India (Table 4).

Further *in-vitro* validation was performed for six apparently non-binding epitopes (labelled E5-E10) through ELISA, Surface plasmon resonance (SPR) and peptide inhibition assay of HIV-1 neutralization. Epitope sequences known to be binding to 10E8 and 4E10 (E1, E2), non-binding to both (E3) and differential binding (E4: binding to 4E10 but not 10E8) were chosen as the control peptides.

Table 3
In silico docking analysis.

	Epitope sequence	10E8		4E10
		Rosetta-dG_separated (REU)	RDOCK energy (kcal/mol)	Rosetta-dG_separated (REU)
Test1	NWFDISSWLWYIK	−21.191	NP	−45.031
Test2	NWFGITKWLWYVK	**−18.588**	NP	−43.894
Test3	NWFRITKWLWYIK	−23.688	NP	−44.947
Test4	SWFEISKWLWYIK	−20.95	NP	**−39.224**
Test5	SWFGITNWLWYIK	−22.882	NP	−42.632
Test6	DWFDISNWLWYIK	**−15.545**	−14.17	−42.17
Test7	KWFSITKWLWYIK	−23.827	**ND**	−42.826
Test8	NWFDITKWLGYIK	**−17.527**	−14.93	**−38.84**
Test9	NWFDITKWLRYIK	−19.778	**ND**	−40.841
Test10	NWFDITSWLWYIK	−22.153	−14.71	−48.122
Test11	NWFGISNWLWYIR	−22.892	**ND**	−42.603
Test12	NWFGITNWLWHIK	−22.949	−14.22	−47.363
Test13	NWFNITNWLWCIK	−21.377	−13.46	−47.155
Test14	NWFSIPNWLWYIK	**−18.151**	−13.48	−44.3
Test15	NWFSISRWLWYIK	−21.096	−10.09	−43.746
Test16	SWFDIAKWLWYIR	−23.364	**ND**	−41.209
Test17	SWFDISNWPWYIR	**−17.548**	**ND**	−41.823
Test18	SWFDITKWLWYIR	−23.833	−6.01	−44.688
Test19	SWFDITRWLWYIK	−23.137	−8.57	−48.135
Test20	SWFNITQWLWYIK	**−19.271**	**ND**	−45.744
Test21	SWFSISKWLWYIK	−20.586	−9.99	**−39.565**
Test22	SWLNITNWLWYIK	−21.058	−7.42	**−39.989**
Test23	TWFGISKWLWYIK	−21.487	−11.54	−41.032
Test24	SWFGITNWLWYIR	−23.289	**−5.5**	−42.736
Test25	DWFNISNWLWYIK	−20.669	**ND**	−43.541
Test26	NWFCISNWLWYIK	−20.838	**ND**	−46.031

Epitopes predicted to be potentially non-binding are highlighted with bold font.
Key: - NP: analysis not performed, ND: No docked pose obtained.

2.4. Peptide binding analysis by Enzyme linked immunosorbent assay

In order to evaluate the predicted interaction of the test epitope peptides, ELISA was performed with both 10E8 and 4E10 antibodies. Multiple dilutions of bNAbs 10E8 (concentration range:10–0.125 μg/mL) and 4E10 (concentration range:0.2-0.001 μg/mL) were assessed for binding to peptides E1-E10. Antibody concentrations were selected to obtain a range of interaction from minimal to saturation for control E1. Specificity of the peptides was established through binding with gp120 directed (unrelated) Ab VRC01 wherein no nonspecific interaction was observed for controls E1-E4 (Supplementary file 3). While test peptides E5 and E6 demonstrated very little binding at Ab concentration greater than 2 μg/mL of 10E8, E7 demonstrated no binding at all (Fig. 3A). Similarly, peptides E8 and E10 were found to be weak binders to 10E8 with no binding observed for peptide E9 (Fig. 3B). In agreement with the docking predictions, E5 and E6 demonstrated binding to 4E10 similar to that of the known binding controls E1, E2 and E4 (Fig. 3C). However, E7 demonstrated very little binding only at concentrations greater than 0.02 μg/mL. Furthermore, binding comparable to positive controls could also be observed for E8 and E10 with no binding observed for peptide E9 (Fig. 3D).

2.5. bNAb-peptide binding kinetic analysis through SPR

SPR experiments were performed with the bNAbs immobilized onto a Biacore CM5 chip. To understand the binding kinetics of MPER peptides to bNAbs the peptides were flowed at concentrations 0.03125–8 μM over bNAbs 10E8, 4E10 and 2F5. The overall binding observed was highest (RU) for 2F5 compared to 4E10 and 10E8, consistent with previously published reports (Huang et al., 2012b). All the test epitopes displayed no binding to 10E8 except E5 and E6 (Fig. 4A). While, E5 and E6 bound readily to 4E10, E9 was observed to be non-binding (Fig. 4B). Epitopes E7, E8 and E10 also bound to 4E10 Ab. Only E1 (MPER peptide HXB2: 662-664) with an intact 2F5 epitope demonstrated binding to the 2F5 antibody (Fig. 4C). Binding constants obtained following kinetic analysis revealed that while E6 did bind to 10E8 antibody, its dissociation rate was much higher than the control epitope E2 (Table 5). Similarly, low association and high dissociation rates were observed for epitopes E7 and E10 (Table 5). Detailed binding kinetics sensograms with fitted curves as per Langmuir (1:1) interaction model have been provided in Supplementary file 4. Having studied the interaction kinetics of the epitope peptides with bNAbs 10E8 and 4E10, we next tested if observed interactions were reproducible in the context of competition with MPER expressed on a virion. This was achieved through a modified HIV-1 neutralization assay as follows.

2.6. Peptide inhibition assay of HIV-1 neutralization

Pseudoviruses 25710-2.43 (tier 1b/2) and 16055-2.3 (tier 2) were utilized for this assay. As per the CATNAP database, pseudovirus 25710-2.43 was reported to be more sensitive to 10E8 neutralization (IC50: 0.045, IC80:0.21 μg/mL) compared to 16055-2.3 (IC50:0.82,

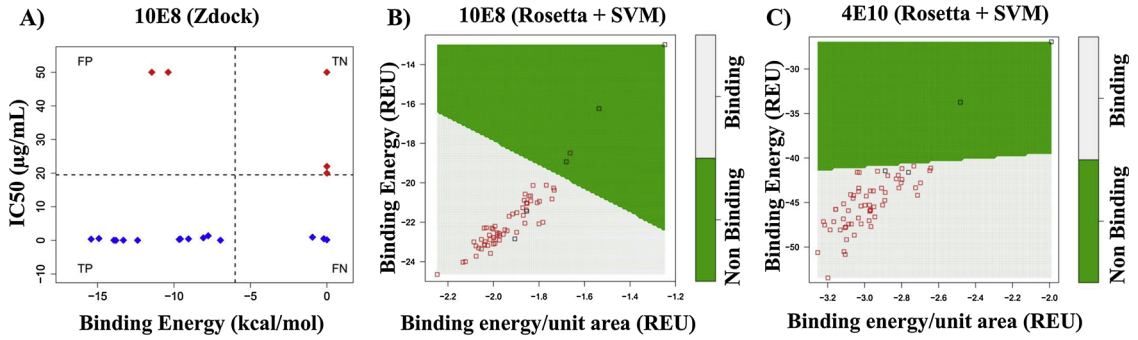

Fig. 2. Binding prediction through Antibody-peptide docking of predicted MPER haplotypes to bNAbs 10E8 and 4E10 using training models.
Training datasets were assessed for their bNAb binding prediction ability by docking experimentally validated MPER sequences to 10E8 (PDB: 4G6F) and 4E10 (PDB: 2FX7) crystal structures.
(A) ZDOCK-RDOCK for 10E8: Experimentally determined IC50 values of 14 positive and 5 negative epitope sequences plotted against their computed RDOCK energies. RDOCK energy of -6 kcal/mol was used as an empirically-fixed threshold to distinguish between potential binders and non-binders, Key: FP- False positive, FN- False negative, TN- True negative, TP- True Positive, positive control epitopes have been indicated with diamond symbols and negative control epitopes have been indicated with filled circle symbols. (B) ROSETTA-SVM for 10E8: Rosetta binding energies (REU) of 59 positive and 5 negative epitope sequences were plotted against their binding energy/unit area (REU). An SVM model was trained to classify binders and non-binders. (C) ROSETTA-SVM for 4E10: Rosetta binding energies (REU) of 60 positive and 4 negative epitope sequences were plotted against their binding energy/unit area (REU). An SVM model was trained to classify binders and non-binders. Red squares indicate known binders while black squares indicate binder in both SVM models. Furthermore, green area indicates predicted plot points of non-binders while white area indicates plot points of predicted binders. (Width: Double columns, Color).

Table 4
Information on epitopes in LANL HIV Database.

No.	Epitope Designation	Epitope sequence	GenBank Accession	Country
1	Test2 (E5)	NWFGITKWLWYVK	Not reported	NA
2	Test4	SWFEISKWLWYIK	AY347730	Turkey
3	Test6	DWFDISNWLWYIK	DQ381976, MF500691	India, Malawi
4	Test7	KWFSITKWLWYIK	DQ398881	India
5	Test8	NWFDITKWLGYIK	AY492999, EU760891	Kenya, India
6	Test9 (E6)	NWFDITKWLRYIK	EU622010	India
7	Test11	NWFGISNWLWYIR	DQ367260, KX069224, KF770281, KF770274, KY229347	India, Botswana, Brazil, Zambia
8	Test14	NWFSIPNWLWYIK	DQ367258	India
9	Test16	SWFDIAKWLWYIR	EU622001	India
10	Test17 (E7)	SWFDISNWPWYIR	JN400531	India
11	Test20 (E8)	SWFNITQWLWYIK	AY935236, HM204581, HQ595866, JX219189	India, South Africa, Malawi, Rwanda
12	Test21	SWFSISKWLWYIK	AF405098, DQ866338, EU857654, HQ596127, HQ707924	South Africa, Zambia, India, Malawi, Zimbabwe
13	Test22	SWLNITNWLWYIK	EU622003	India
14	Test24	SWFGITNWLWYIR	KC186730	Malawi
15	Test25 (E9)	DWFNISNWLWYIK	KC156151, EU166687, HM638816, HQ595759	India, Zambia, Malawi, SA
16	Test26 (E10)	NWFCISNWLWYIK	DQ367253	India

In cases of multiple GenBank entries for an identical epitope sequence, maximum 5 have been enlisted in the table.

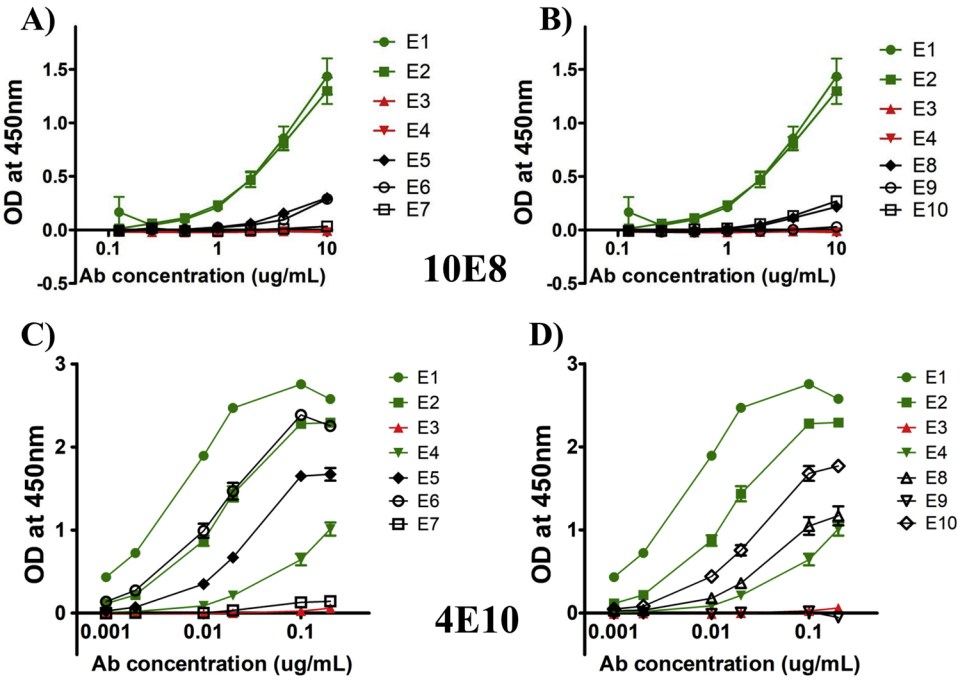

Fig. 3. Binding analysis of epitope test peptides with bNAbs by ELISA.
Binding ability of untested MPER variants (E5-E10) to bNAbs 10E8 (A, B) and 4E10 (C, D) was assessed with ELISA for bNAb concentration range 0.125–10 μg/mL of 10E8 and 0.001-0.2 μg/mL of 4E10. Peptides E1 and E2 were used as positive controls while Peptide E3 was used as a negative control for both bNAbs. Peptide E4 was used as a differential control that has higher affinity for 4E10 than 10E8. (Width: 1.5 columns, Color).

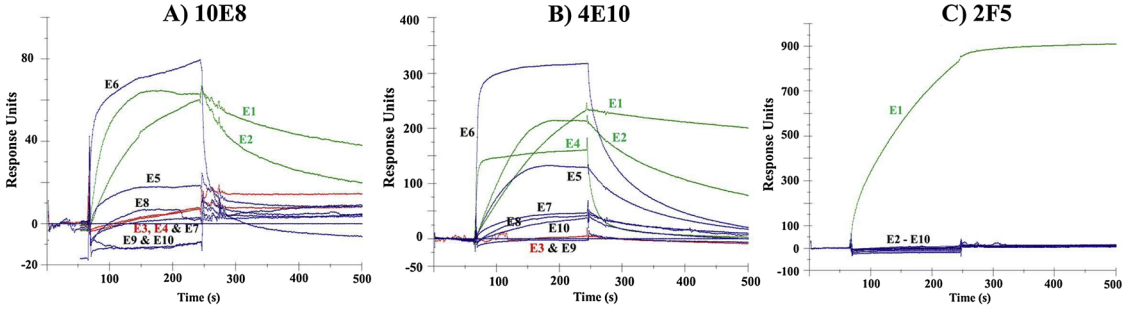

Fig. 4. Surface Plasmon resonance analysis of binding kinetics of untested MPER variant haplotypes.
Binding analysis of MPER variants (E5-E10) were tested with bNAbs 10E8, 4E10 and 2F5 along with control peptides E1-E4 as described previously. All the plots depict an overlay of Peptides (0.5 μM) injected over (A) 10E8, (B) 4E10 and (C) 2F5 immobilized on a CM5 sensor chip with association and dissociation times of 3 min. and 5 min. respectively. (Width: Double columns, Color).

IC80: 3.16 μg/mL). Similarly, pseudovirus 25710-2.43 was also reported to be easily neutralized by 4E10 (IC50:0.54, IC80:3.42 μg/mL) compared to 16055-2.3 (IC50:1.7, IC80:27.36 μg/mL). In the present assay, in case of Ab-peptide binding, percent neutralization is expected to drop in a dose-dependent manner, whereas little to no change is expected in percent neutralization in case of non-binding peptides.

Table 5
SPR analysis.

	k_a (1/Ms)		k_d (1/s)		K_D (M)		Fold K_D	
	10E8	4E10	10E8	4E10	10E8	4E10	10E8	4E10
E1	1.23E+04	1.33E+04	4.81E-03	8.57E-04	3.91E-07	6.43E-08	7.4	2.5
E2	1.22E+05	1.04E+05	6.42E-03	2.68E-03	5.26E-08	2.58E-08	1	1
E3	NB	NB	NB	NB	NB	NB	NA	NA
E4	NB	1.15E+05	NB	1.14E-02	NB	9.94E-08	NA	3.9
E5	3.13E+05	5.88E+04	8.98E-03	7.11E-03	2.87E-08	1.21E-07	0.5	4.7
E6	2.48E+04	5.10E+05	9.95E-03	9.15E-03	4.02E-07	1.79E-08	7.6	0.7
E7	NB	2.51E+04	NB	0.0126	NB	5.04E-07	NA	19.5
E8	NB	6.90E+04	NB	7.23E-03	NB	1.05E-07	NA	4.1
E9	NB	NB	NB	NB	NB	NB	NA	NA
E10	NB	11500	NB	6.52E-03	NB	5.64E-07	NA	21.9

Fold KD values were computed w.r.t K_D of E2. NB: Non-binding, NA: Not applicable.

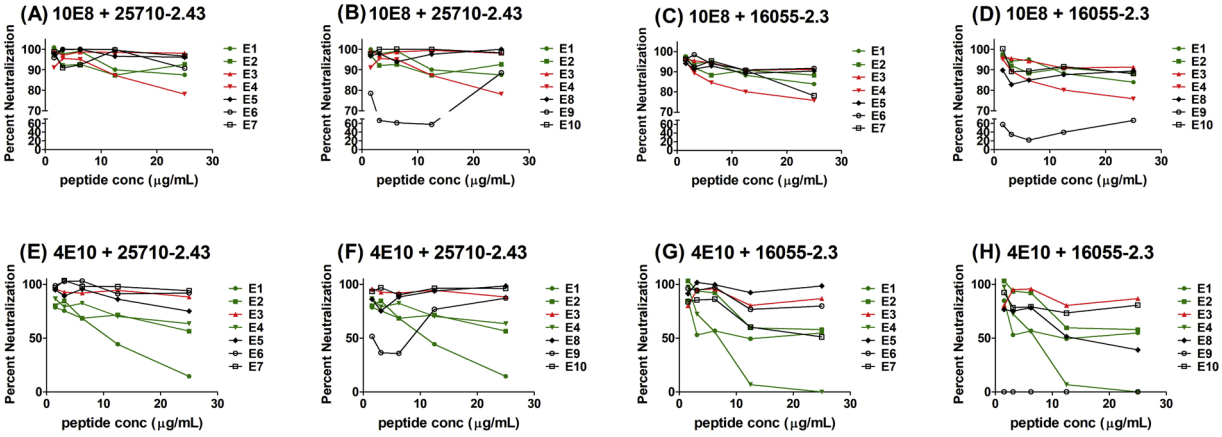

Fig. 5. Peptide inhibition assay of HIV-1 neutralization.
To assess if the MPER variant peptides (E5-E6) can interfere with HIV-1 neutralization, peptide inhibition assay of viral neutralization was performed for peptides E1-E10, with two pseudoviruses 25710-2.43 (HIV-1 clade C, neutralization tier 1b) and 16055-2.3 (HIV-1 clade C, neutralization tier 2). Percent neutralization changes were plotted with peptide concentration increasing from 1.5625 μg/mL to 25 μg/mL for bNAb 10E8 with pseudovirus 25710-2.43 (A, B) and 16055-2.3 (C, D). Similarly, results were also plotted for bNAb 4E10 with pseudovirus 25710-2.43 (E, F) and 16055-2.3 (G, H). Percent neutralization values were calculated for every concentration of each of the peptides with each bNAb- pseudovirus combination as follows: Percent neutralization = 100-(test-CC)/(VC-CC)*100, Where, CC is luminescence (RLU) of Cell control (no peptide, pseudovirus and antibody), VC is luminescence (RLU) of Virus control with pseudovirus and cells (no peptide, antibody), Test is luminescence (RLU) of a test well containing a specific peptide concentration, antibody, pseudovirus and cells. (Width: Double columns, Color).

Overall effect of peptide binding was much more pronounced in case of 4E10 compared to 10E8 (Fig. 5). For 10E8 Ab, peptides E5, E7, E8 and E10 demonstrated little to no binding i.e. little to no change in neutralization of 25710-2.43, with E6 binding only at high concentration (Fig. 5A and B). In case of peptide E9, atypical behavior was observed wherein, the peptide seemed to have completely inhibited neutralization, i.e. high affinity binding, however, this effect was not dose dependent and therefore was deemed inconclusive (Fig. 5B). Of note, control peptide E4, which demonstrated little to no binding in ELISA as well as SPR, was observed to significantly inhibit the neutralization of 25710-2.43. This may possibly be due to the high peptide concentration used in this assay compared to the earlier experiments matching the IC50 value of ≥20 μg/mL reported for this epitope. In the case of pseudovirus 16055-2.3, peptide binding was not observed for E5, E6 and E8 (Fig. 5C and D). Similar to E4, E7 demonstrated considerable inhibition at high peptide concentration in discordance with ELISA and SPR. E10 demonstrated inhibition of neutralization at lower concentrations but the effect was lost with increasing peptide concentration. E9, demonstrated atypical interaction similar to that with 25710-2.43 (Fig. 5D).

In case of the 4E10 antibody, no binding/neutralization inhibition was observed for E6, E7, E8 and E10 while low binding was observed for E5 with pseudovirus 25710-2.43 (Fig. 5E and F). Atypical binding was observed for E9 with 4E10 similar to 10E8 (Fig. 5F). In the assay with the pseudovirus 16055-2.3, no inhibition was observed for peptides E5, E6 and E10 (Fig. 5F and G). Neutralization inhibition was observed for peptides E7, E8 with atypical dose non-dependent inhibition by E9. The results for E5 and E10 were discordant with ELISA and SPR carried out with these peptides.

Taken together, for the 10E8 bNAb, data from 3 qualitatively distinct *in vitro* assays supported the *in silico* predictions of weak/non-binding of the selected haplotypes. Differential binding was also observed for these haplotypes with the bNAb 4E10. Table 6 (10E8) and Table 7 (4E10) summarize the concordant results obtained by the above described *in vitro* methods.

3. Discussion

HIV sequence analysis studies are instructive of emergence of viral diversity, evolution within a host as well as over a population and epidemiological fitness of the virus (Liao et al., 2013; Moradigaravand et al., 2014; Payne et al., 2014; Sangeda et al., 2013). The present study sought to document the variation in the MPER of gp41 from a cohort of HIV-1C infected individuals from Mumbai, India as well as other reported data for MPER from India. Furthermore, the impact of this variation was evaluated on binding to broadly neutralizing antibodies 4E10 and 10E8.

Indian HIV epidemic is believed to have arisen from African viral

Table 6
Summary of bNAb 10E8 assays.

Peptide	*In silico* docking with bNAb 10E8	ELISA	SPR	Peptide Inhibition of HIV-1 neutralization	
				25710-2.43	16055-2.3
E1 (C)	D	+ +	+ +	+ +	+ +
E2 (C)	D	+ +	+ +	+ +	+ +
E3 (C)	N	–	–	–	–
E5	N	+	+	–	–
E6	N	+	+ +	+ +	–
E8	N	+	–	–	–
E9	N	–	–	IC	IC
E10	N	+	–	–	+

N: no *in silico* binding, D: *in silico* binding, + +: binding, +: low binding, −: no binding, IC: inconclusive, (C): Control peptides.

Table 7
Summary of bNAb 4E10 assays.

Peptide	*In silico* docking with bNAb 4E10	ELISA	SPR	Peptide Inhibition of HIV-1 neutralization	
				25710-2.43	16055-2.3
E1 (C)	D	+ +	+ +	+ +	+ +
E2 (C)	D	+ +	+ +	+ +	+ +
E3 (C)	N	–	–	–	–
E4 (C)	D	+ +	+ +	+ +	+ +
E5	D	+ +	+ +	+	–
E7	D	+	+	–	+ +
E8	D	+ +	+	–	+ +
E9	D	–	–	IC	IC

N: no *in silico* binding, D: *in silico* binding, + +: binding, +: low binding, −: no binding, IC: inconclusive, (C): Control peptides.

strains, yet, previously published data demonstrates distinct phylogenetic clustering of African and Indian gp41 sequences (Agnihotri et al., 2006; Neogi et al., 2012). HIV1C sequences remain poorly represented in the LANL HIV database, a prime resource for HIV research. As of March 2019, amongst sequences for any HIV-1 genomic region (N = 782,325), the database contains 16.77% sequences (N = 131,248) reported for clade C. Of the total gp41 sequence data (N = 73,389), only 0.4% (N = 294) has been sampled from India. Therefore, it may be hypothesized that the diversity in gp41 region of the Indian epidemic may possibly be underreported and thereby understudied. In the study described herein, a total of 24 HIV-1C chronically infected individuals were recruited and MPER variation analysis was performed for blood derived primary viral sequences (plasma circulating/proviral) with high throughput sequencing. For a robust analysis, technical replicates of three datasets sequenced by two different amplification/sequencing protocols were used for calculation of minimum threshold for viral variation. While the MPER was observed to be largely conserved, some of the amino acid positions displayed variation on both intra and interindividual levels, which included the contact sites of four well characterized broadly neutralizing antibody (bNAb) epitopes in MPER, viz., 2F5, Z13e1, 4E10 and 10E8. In spite of the sensitivity afforded by high throughput sequencing in studying amino acid variation, their interaction and cumulative impact on the epitope structure cannot be predicted with simple 'position-wise' variation analyses (Henn et al., 2012). To address this caveat, a custom haplotype analysis script was developed and successfully applied in the present study. The applicability of this script can be extended to study variation in any stretch of DNA sequence (length of sequence of interest < read length) in which proximate residues are expected to interact and where position specific variation data may not be enough to derive biologically relevant conclusions. The applications include continuous bNAb epitopes, CD4/CD8 T-cell epitopes and active sites of enzymes. All the generated epitope sequences analyzed were predicted to be 2F5 resistant and to have moderate predicted sensitivity to Z13e1 as reported previously (Gray et al., 2006; Kulkarni et al., 2009; Nelson et al., 2007). The epitope haplotype analysis also generated multiple untested 4E10 and 10E8 epitopes for subsequent study. Upon comparison with the epitope sequences of viruses with reported neutralization data, we found 6 epitope haplotypes with no available 10E8 neutralization data. With extension of this query to all the sequences reported from India, 20 more 10E8 untested epitope sequences were identified. 10E8 has been considered a promising candidate for bNAb based interventions including as a component of multi-bNAb combinations as well as engineered bispecific antibodies (Padte et al., 2018; Wagh et al., 2018, 2016). Therefore, evaluation of binding of untested epitope sequences to 10E8 and 4E10 may further inform the phenotypic implications of genotypic diversity in the MPER region in India. To achieve this, an *in silico* docking analysis was performed with two independent molecular docking approaches in an attempt to predict binding ability of the 4E10/10E8 haplotypes generated. Thirteen out of the 26 epitopes were predicted to be non-binding to 10E8 with 3 more predicted to be non-binding to 4E10. Interestingly, 7 of these non-binding sequences were found to be exclusively reported from India. From the 13 epitopes predicted to be non-binding for 10E8, 2, predicted by both docking approaches and 4 more chosen randomly were selected for experimental validation.

The *in vitro* analysis for validation of *in silico* docking observations with 10E8 and 4E10 bNAbs was performed by a three-pronged approach: i) An end-point ELISA assay with variable concentrations of bNAbs, ii) A real time SPR based kinetic binding analysis with variable concentration of peptides and iii) A peptide inhibition assay of HIV-1 neutralization. Both ELISA and SPR results were concordant on weak/no-binding to 10E8 for 5 out of the 6 epitopes (all except E6). For E6, lower association rate (k_a) coupled with higher dissociation rate (k_d) resulting in 7.6 times higher K_D could explain this contrariety. ELISA and SPR results were also consistent in case of 4E10 for peptides E5, E6, E7 and E9. SPR was able to resolve the apparently strong interaction observed in the end-point ELISA in terms of low affinity interaction observed for E8 and E10. The peptide inhibition of neutralization was the most challenging of the three assays as MPER binding affinity of 10E8 has been reported to be significantly lower than 4E10 in spite of higher neutralization potency compared to the latter (Huang et al., 2012a). Furthermore, our choice of pseudoviruses included a member of Tier 2 category of pseudoviruses with a closed conformation that is reported to be closest to the circulating HIV strains in the population and are thus considered important in neutralization assessment (Montefiori et al., 2018). We employed two distinct pseudoviruses reported from India for the neutralization assay, 25710-2.43, a tier 1b/2 pseudovirus and 16055-2.3, a tier 2 virus with the expectation that the inhibition effect, if any, of the peptides would be more pronounced in case of 16055-2.3 compared to 25710-2.43. The extent of peptide inhibition was lower in the case of 10E8 (upto 20%) compared to 4E10 (upto 100%) probably due to lower binding affinity of 10E8 (Huang et al., 2012b). Peptides E5, E8 and E10 demonstrated no inhibition of neutralization by 10E8 in case of both the pseudoviruses tested. While E6 demonstrated dose dependent inhibition only in case of 25710-2.43, inhibition of neutralization was also observed for E7 at high peptide concentration similar to that observed for E4, but only in case of 16055-2.3. The results for peptide E9 were inconclusive as no dose dependent 10E8 neutralization inhibition could be clearly discerned. Thus, for bNAb 10E8, clearly discordant results were obtained in cases of peptide E4 and E7, wherein no interaction was observed in ELISA and SPR but strong interaction was seen with the neutralization assay. Little to no inhibition of 4E10 neutralization was observed for peptide epitopes E5, E6 and E10. Neutralization inhibition by epitopes E7 and E8 was observed in the assay with 16055-2.3 but not with 25710-2.43. We believe this may have been caused by the vast difference between the IC50 and IC80 values of 16055-2.3 (IC50:1.7, IC80:27.36 μg/mL) compared to

25710-2.43 (IC50:0.54, IC80:3.42 μg/mL). For bNAb 4E10, clear interaction was observed for peptides E5 and E10 with ELISA and SPR but no interaction was detected with the neutralization assay. The reasons for the discordant results are unclear and we plan to explore these further in future studies. Overall, for both bNAbs, when results from all 3 assays were compared, we noted that there appeared to be some discordance (2 out of 10 peptides for each bNAb) between the peptide inhibition of HIV-1 neutralization assay compared to ELISA and SPR which were highly concordant. In summation, by studying the different facets of epitope-bNAB interaction through ELISA, SPR and neutralization assay, we found that the epitope haplotypes predicted to be non-binding through docking analyses were indeed weak binders or could not bind 10E8 antibody at all. Differential binding of these peptides was also observed with bNAb 4E10.

High throughput sequencing, single cell analyses, sensitive and precise neutralization assays have provided an unprecedented ability to perform in-depth study of viral evolution as well as to rapidly discover and optimize broadly neutralizing antibodies (Burton et al., 2012b). Yet, predicting viral phenotype, its exact structure and its interaction with antiviral drugs or bNAbs from only sequence data is still a challenging task. Efforts are being made to address this impediment through implementation of molecular docking and machine learning algorithms such as decision trees, neural networks, random forests and Bayesian networks (Hepler et al., 2014). Following a similar strategy, we have described presence of 10E8 weakly binding/non-binding variants in HIV strains from the Indian population which remains vastly underreported. While we included stringent technical controls to account for errors introduced during processing of the sample, we could not rule out presence of these variants in non-replicating proviral DNA. Nonetheless, entries of identical epitope sequences reported in the database as major variants detected through traditional sanger sequencing rules out this possibility. As noted previously, 10E8 interaction predictions suffer from low specificity due to lack of enough data on resistant epitope sequences (Hepler et al., 2014). Similarly, while our docking prediction sensitivity and accuracy were higher, the specificity was low. Increased number of nonbinding sequences as described in the present study will perhaps strengthen the prediction abilities. Interestingly, 7 out of the 16 non-binding sequences were reported exclusively from India as early as 1995 suggesting that their frequency in the Indian population maybe much higher. Indeed, epitope E8 from our study is present at a frequency of 0.004% in global HIV-1C sequences but at a frequency of 1.36% in reported sequences from India. Additionally, the analyses herein were restricted to the MPER peptides and therefore, interactions and influence of gp120 as well as gp41 cytoplasmic tail were not addressed which may possibly be significant (Bricault et al., 2019).

While the analyses described herein were restricted to 10E8 epitope in the MPER region, they hold true for other viral regions, not only for India with the third largest HIV epidemic in the world, but also clade C overall (Bhattacharya, 2018; Kumar et al., 2018). Such lack of information pertaining to population specific diversity/genetic drift may present a major obstacle in development of universal intervention strategies. Regular surveillance of HIV molecular clones circulating in the population needs to be undertaken to address this concern.

4. Conclusion

Integrative approaches addressing viral variation are essential to inform interventions for control and/or long-term disease management of HIV infection. High throughput sequencing technologies and *in silico* tools, used as demonstrated in the present analyses, could uncover critical population-specific viral diversity. The data described herein pertaining to the underreporting of 10E8 non-binding gp41 sequences from India, highlight the need for regular national surveillance of circulating molecular clones.

5. Materials and methods

5.1. Study participants

5.1.1. Ethics statement and collection of specimens

HIV-1 clade C infected individuals were recruited from J.J. group of Hospitals, Mumbai following approval of ethics committees of both the participating Institutes (NIRRH and Grant Government Medical College). The informed consent forms were provided and duly signed by all study participants prior to recruitment. In the case of minors/children, duly signed informed consent was obtained from their next of kin/caretakers/guardian, as per the participating institutes' ethics committee guidelines.

A total of 24 HIV-1 infected individuals (S1-S24) without any documented coinfections at sampling were recruited. Participants selected for the study were of the ages ranging from 8 to 55 years and chronically infected with documented period of infection ranging from a few months to 16 years. CD4 counts for each subject were estimated through flow cytometry on a BD FACS Calibur cytometer (Becton Dickinson, San Jose, CA). The total nucleic acids were isolated from plasma samples using MagNA pure automated Nucleic Acids isolation system (Roche) and Real time PCR was performed for viral load estimation with COBAS® Taqman HIV-1 V2.0, Roche as per manufacturer's instructions.

5.2. Processing of samples

5.2.1. Plasma and PBMC separation

10 ml of blood was collected from each subject into EDTA vacutainers (BD Biosciences,). Blood plasma was collected following centrifugation at 700*g* for 10 min and the aliquots were stored at −80 °C until further use. Peripheral Blood Mononuclear Cells (PBMCs) were isolated from each sample by continuous density gradient centrifugation (Kanof et al., 1996). Briefly, blood samples were diluted 1:1 with RPMI-1640 (Himedia laboratories Ltd., India) and overlayed on Hisep (Himedia laboratories Ltd., India) in 3:1 proportion and centrifuged at 700*g* for 20 min at room temperature. PBMCs obtained were given two washes with RPMI 1640 and $\sim 5 \times 10^6$ cells were suspended in 200 μL RPMI-1640. Isolated PBMCs were further processed for extraction of genomic DNA.

5.2.2. Preparation of proviral DNA and viral RNA derived cDNA

Genomic DNA was extracted form PBMCs by using QIAamp blood DNA mini kit (Qiagen, Valencia, CA) as per manufacturer's instruction. The stored plasma samples were centrifuged at 1500*g* for 15 min at room temperature and viral RNA was isolated using QIAamp Viral RNA mini kit (Qiagen, Valencia, CA) as per manufacturer's instructions. Both DNA and RNA were checked for purity and quantitated spectrophotometrically.

Plasma isolated RNA from samples S1, S2, S3 and S4 was primed with 'MSR5' oligo (5'-GCACTCAAGGCAAGCTTTATTGAGGCT-3') proximal to 3' end of the HIV RNA genome (HXB2: 9605-9632). The extracted RNA (~1 μg) was reverse transcribed in a total volume of 20 μL with PrimeScript 1st strand cDNA synthesis kit (TaKaRa). The RNA, 500μM of dNTPs mix and 0.4μM MSR5 primer were incubated at 65 °C for 5 min, followed by addition of 1X RT buffer, 20U of RNase Inhibitor, 200U of PrimeScript RTase and RNase free H_2O. The reaction mixture was further incubated at 50 °C for 1 h followed by an inactivation step at 70 °C for 15 min.

5.2.3. PCR amplification of gp41 gene

Extracted genomic DNA and cDNA were used as templates for amplification of gp41 gene using two approaches that employ a nested PCR method. Primer sequences for both protocols have been used as described by previously (Nadai et al., 2008). Primer information is provided in Supplementary file 5. For participants S1 and S2, amplicons

were generated from both plasma derived circulating viral RNA (designated as S1_1 and S2_1) as well as PBMC derived proviral DNA (designated as S1_2 and S2_2). For participants S3 and S4, amplicons were generated from plasma derived circulating viral RNA only whereas, for participants S5-S24 the amplicons were generated only from PBMC derived proviral DNA.

Templates S1-S8 were amplified with protocol 1 wherein gp41 region amplified by a nested PCR consisting of a 1 st round PCR (3631 bp amplicon) followed by the second round PCR covering gp41 region (950 bp amplicon, HXB2: 7859-8812). Amplification protocol 2 consisted of near full-length genome (NFLG) amplification which was employed for templates S9-S24. NFLG was amplified as reported previously (Nadai et al., 2008) with 3 overlapping fragments covering gag, pol and env regions of HIV-1 genome (HXB2: 769-9089). Equal volumes of PCR products amplified in three independent reactions were pooled to partially mitigate amplification bias and purified with Nucleospin PCR gel purification kit (Machery Nagel) per manufacturer's instructions. Detailed description of PCR primers and condition has been provided in Supplementary file 5.

5.3. Next generation sequencing (NGS) of PCR products

Next generation sequencing of PCR products (S1-S8) was performed commercially (SciGenom Labs, Kochi, India). Library preparation was performed as per Truseq sample preparation protocol V2. Sequencing was carried out on the Illumina Miseq V2 platform to obtain 'paired read' data (read length 151 bases) in 'FASTQ' format. The sequence data have been deposited with links to BioProject accession number PRJNA243844. (NCBI-SRA: SRP040990). Next generation sequencing of PCR products (S9-S24) was also performed commercially (Interpretomics India Pvt. Ltd., Bengaluru, India). The DNA libraries were prepared from PCR amplicons using TrueSeq Nano DNA library preparation kit. Sequencing was carried out on Illumina Hiseq platform to obtain 'paired read' data (read length 101 bases) in 'FASTQ' format. The sequence data have been deposited with links to BioProject accession number PRJNA493619. (NCBI-SRA: SRP162802). Templates from S5, S6, and S8 were amplified and sequenced by both amplification sequencing protocols for amplification and sequencing bias analyses.

5.4. NGS data analysis

5.4.1. Read alignment and variant calling

Consensus sequences were generated for each sample using VICUNA v1.3 with default parameter settings (Yang et al., 2012). Quality filtration of the reads was performed with Trimmomatic v0.32 (Bolger et al., 2014). The reads were aligned to their respective consensus sequences using Mosaik v2.2.3 to obtain Binary alignments/map (BAM) files (Lee et al., 2014). Bamtools v2.3.0 was used for removal of aligned singletons with unmapped partners along with any unaligned reads from the BAM files (Barnett et al., 2011). Further processing involving sorting, indexing and removal of PCR duplicates was performed using Picard tools v1.95 (http://broadinstitute.github.io/picard/). Indel realignment was carried out with Genome analysis toolkit v3.7 (McKenna et al., 2010). Samtools v1.3.1 (Li et al., 2009) and custom bash/awk scripts were used to generate detailed statistics of the FASTQ and BAM files and to perform coverage analysis. All the NGS datasets had coverage ranging from 1500x to 30,000×. BAM files were visualized at various analysis stages with Interactive Genomics Viewer (IGV) (Robinson et al., 2011; Thorvaldsdóttir et al., 2013). Consensus sequences generated by 'Vicuna' for each of the individuals were deposited in the GenBank (Accession numbers MN462559-MN462584).

5.4.2. Variant calling and analysis

Variant calling was performed with V-Phaser 2.0 and the output was analyzed with V profiler to generate nucleotide, codon frequency tables (Henn et al., 2012; Yang et al., 2013). Codon frequency tables were converted to amino acid frequency tables after accounting for codon redundancy. Variation/Heat at every position was calculated as (100 – percent frequency of the major variant). Position- wise sequence logos were prepared using the tool Weblogo 3 (Crooks et al., 2004; Schneider and Stephens, 1990). Amino acid frequencies derived from this analysis were used to generate epitope-wise heat maps using the 'pheatmap' package of R statistical computing software (v3.4.0) and R studio v1.0.143 (R Core Team, 2018; Raivo Kolde, 2018; RStudio Team, 2015). Amino acid frequency tables generated from all of the datasets were also pooled together and used for generation of position weight matrices for epitope regions of bNAbs 2F5, Z13e1, 4E10 and 10E8 and subsequently used for generation of Sequence logos.

5.4.3. Variant threshold calculation

As templates generated from individuals S5, S6 and S8 were amplified and sequenced by both protocols (2 datasets per template), data generated from them was used to calculate a threshold frequency for acceptance of a variant to address amplification as well as sequencing bias. Amino acid frequency at each position was compared between the two datasets from the same template. An average percent variation difference per position was calculated. The difference obtained was 6.31%, 7.07%, and 4.22% respectively, average of which, 5.87% was set as a minimum acceptance frequency for an amino acid variant.

5.4.4. Epitope haplotype analysis with 'Haplocount'

Haplotypes were generated for epitopes of antibodies 2F5 (HXB2 numbering 662-667), Z13e1 (666-677), 4E10 and 10E8 (671-683) using Vprofiler. As Vprofiler was unable to generate haplotypes for datasets containing indels, a custom bash script, HaploCount, was written in 'bash' to generate epitope haplotypes and their corresponding frequencies. The haplotype analysis script accepts a BAM file as an input. An epitope haplotype is considered, provided it is present in a single contiguous read sequence. Briefly, aligned forward and reverse reads are drawn from the BAM file using samtools. Reads are filtered with fastx toolkit v0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit/) based on a threshold of phred quality of 30 for > 95% bases in a read. Reads are translated in all reading frames with 'transeq' tool of the EMBOSS package (Rice et al., 2000) and the correct reading frames are selected with HMMER v3.1b2 (https://github.com/EddyRivasLab/hmmer) based on a training dataset of 584 HIV-1 clade C gp41 sequences. The translated reads thus selected are aligned with Muscle v3.8.31 (Edgar, 2004) and the epitopes are extracted using 'extractalign' tool of the EMBOSS package. Finally, frequencies of epitopes obtained are calculated following filtration as per the pre-calculated threshold of 5.87% as described previously. Upon comparison with Vprofiler, at 5.87% threshold, both scripts demonstrated a concurrence of ~87% in terms of number of haplotypes identified with comparable frequencies. Epitope haplotypes thus generated were further analyzed for their binding affinities to the 10E8 and 4E10 antibodies with *in silico* docking strategies. The haplotype analysis script- HaploCount.sh is hosted on GitHub (https://github.com/jyotiS92/HaploCount).

5.5. Protein-peptide docking

The crystal structures of antibody 10E8 Fab complexed with HIV-1 gp41 peptide (PDB: 4G6F) as well as antibody 4E10 Fab complexed with HIV-1 gp41 peptide (PDB: 2FX7) were retrieved from Protein Data Bank. Protein-peptide docking was carried out with two algorithms. A dataset of positive control haplotypes (Antibody sensitive Pseudovirus epitope sequences reported by LANL CATNAP database) and negative control haplotypes (Antibody resistant Pseudovirus epitope sequences reported by LANL CATNAP database) were used for prediction training of both the algorithms as follows:

5.5.1. ZDOCK-RDOCK

The 3D structures of the epitope variants viz., 14 positive controls including Indian sequences, 5 negative control sequences and 21 test haplotypes were built, energy minimized and further docked to FAB region of 10E8 using ZDOCK algorithm of Accelrys Discovery

studio 3.5 (Li et al., 2003). For each of the peptides, ZDOCK generated 10 clusters; each containing several docked poses ranked according to their ZDOCK and ZRANK scores (Pierce and Weng, 2007). The docked poses were further refined using RDOCK, wherein electrostatics and de-solvation energies were computed and the poses were re-ranked according to the sum of these two energy terms after several rounds of CHARMM minimization (Rong Chen, 2003). The pose with the most favorable RMSD and RDOCK energy was selected for each epitope variant.

5.5.2. Rosetta modelling

Rosetta framework 2016.13.58602_bundle (Leaver-Fay et al., 2011) was used for protein peptide docking analysis of 59 positive control haplotypes, 5 negative control haplotypes and 26 test haplotypes as defined previously for the antibody 10E8. Similar protocol was followed for docking of 60 positive haplotypes and 4 negative haplotypes for the antibody. The crystal structures were prepared for use in Rosetta following an all-atom refinement with all heavy atom constraints applied through Rosetta Relax protocol (Conway et al., 2014; Nivón et al., 2013; Tyka et al., 2011). The amino acids in the co-crystallized peptide were replaced with those of predicted epitope haplotypes using FixBB protocol. As this protocol does not move the atoms of peptide backbone, the Relax protocol was used again on each of the structures for energy re-stabilization. Further, FlexPepDock refinement protocol of Rosetta was applied to perform molecular docking through iterative optimization of peptide backbone and its rigid body orientation relative to the antibody FAb along with side chain optimization (Raveh et al., 2010). For each of the controls/test haplotypes, 100 decoys were generated with low-res pre-optimization and 100 decoys were generated without this flag. Top 10 models were selected for each of the categories by model score (total 20) and further processed for interface analysis. Rosetta Interface analyzer was used to calculate the binding energy (dG separated) and the binding energy per unit area (dG separated/dSASA_int * 100) (Lewis and Kuhlman, 2011) following separation and repacking of both the subunits of the model. The models were re-ranked as per the model score and the binding energy of the top scoring model was used as the binding energy value for the haplotype.

5.5.3. Classification of binding versus non-binding haplotypes with support vector machine

Using the 'binding energy' and 'binding energy per unit buried surface area' values obtained for the training datasets through Rosetta docking protocol, SVM models were built by implementation of SVM training algorithm from the R package 'e1071' (Dimitriadou et al., 2019). SVM algorithm with linear kernel was implemented for the classification of test epitopes with a constraint violation cost of 10.

5.5.4. Docking algorithm attributes

Binding predictions for the training data sets by both algorithms were categorized as i) true positives (TP)- neutralization sensitive epitopes predicted to be binding ii) true negatives (TN)- neutralization resistant epitopes predicted to be non-binding iii) false negative (FN)- neutralization sensitive epitopes predicted to be non-binding and iv) false positive (FP)- neutralization resistant epitopes predicted to be binding. Sensitivity for both algorithms was calculated as TP/(TP + FN), specificity as TN/(TN + FP) and accuracy as (TP + TN)/(TP + FP + TN + FN).

Training models generated as described above were then used for classification of the 26 test haplotypes as binders and non-binders to bNAbs 10E8 and 4E10.

5.6. In vitro binding analysis of novel MPER epitope haplotypes

In vitro validation of MPER epitope haplotypes predicted to be non-binding to bNAbs 10E8 and 4E10 was performed with ELISA, SPR and peptide inhibition assay of HIV-1 neutralization. Peptides E1 (26mer MPER peptide), E2 (NWFSITKWLWYIK), E3 (NLLDISHWLGYIR) and E4(NWFDISNWLRYIQ) with published neutralization IC50 values were used as internal controls. Test epitope peptides (E5-E10) were synthesized commercially with > 90% purity (ABclonal Biotechnology Co., Ltd)

5.6.1. Peptide binding analysis with enzyme-linked immunosorbent assay (ELISA)

Peptides (100 µL, 5 µg/mL) were coated onto 96 well flat-bottomed ELISA plates (Maxisorp, Nunc) in coating buffer (0.1 M carbonate-bicarbonate buffer, pH 9.6) and incubated overnight at 4 °C. Plates were blocked with 5% gelatin (gelatin from cold water fish skin, Sigma) in 0.1 M PBS for 1 h at 37 °C. Plates were further incubated with 100 µL of bNAbs 10E8 (concentration range (µg/mL): 10, 4, 2, 1, 0.5, 0.25 and 0.125) and 4E10 (concentration range (µg/mL):0.2, 0.1, 0.02, 0.01, 0.002, 0.001) in 1% BSA (Fisher scientific) in 0.1 M PBS and incubated for 1.5 h at 37 °C. Subsequently, 100 µL of horse radish peroxidase conjugated goat antihuman secondary IgG antibody (1:2000, Sigma) was added to each well and incubated at 37 °C for 1 h. Plates were washed thrice with 0.05% tween-20 in 0.1 M PBS between each step. Reaction was developed at RT with 3,3′,5,5′-tetramethyl benzidine (TMB) substrate (Bangalore genei) for 20 min and stopped with 4 N H_2SO_4. The absorbance value at 450 nm (OD450) was read using ELISA reader (BioTek). A similar ELISA assay was also performed with bNAb VRC01 (concentration range (µg/mL): 10, 1, 0.1, 0.01, 0.001 and 0.0001) for peptides E1-E4 with CN54-gp120 protein as the positive control (Morikawa et al., 1990; Wang et al., 1995; Wu et al., 2010).

5.6.2. Peptide binding kinetic analysis with surface plasmon resonance (SPR)

Surface plasmon resonance based kinetic binding analysis was performed on Biacore 3000 instrument (GE Healthcare). Broadly neutralizing antibodies 10E8, 4E10 and 2F5 were covalently coupled to a CM5 sensor chips by amine coupling at final densities of ~14000 Response Units (RU). Peptides were serially diluted (2-fold) from 8µM to 0.03125µM in Biacore HBSEP buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.1% P20) and evaluated at a 30 µl/min flow rate with permitted association and dissociation times of 3 and 5 min respectively. Binding analysis was performed with BiaEvaluation software (GE Healthcare).

5.6.3. Preparation of Env-pseudotyped viruses

Env-pseudotyped viruses were prepared as described previously (Patil et al., 2016). Briefly, 293 T cells were co-transfected with an HIV-1 backbone plasmid lacking a functional envelope (pSG3ΔEnv) and an HIV-1 envelope expressing plasmid (25710-2.43 or 16055-2.3) using FuGENE6 transfection kit (Promega). Culture supernatants containing the Env- pseudotyped viruses were collected 48 h post-transfection and stored at −80 °C until further use. Infectivity titres of the pseudoviruses thus produced were assessed with a luciferase-based assay wherein TZM-bl cells (10^5cells/ml) were infected in presence of 25 µg/mL of DEAE-dextran in 96-well microtiter plates. Virus titres were determined by addition of Britelite plus luciferase substrate (PerkinElmer) to the assay plate followed by luciferase activity measurement based on relative luminescence unit (RLU) using a Victor X2 luminometer (PerkinElmer)

5.6.4. Peptide inhibition assay of HIV-1 neutralization

Peptide inhibition assay of HIV-1 neutralization was performed for 2 HIV-1 clade C pseudoviruses: 25710-2.43 (tier 1b/2) and 16055-2.3 (tier 2). Serially diluted peptides (25 µL, concentrations - 25 µg/mL

-1.5625 μg/mL, 2-fold dilution) were incubated with 25 μL of bNAbs 10E8 (conc: 5 μg/mL) and 4E10 (conc: 10 μg/mL) in 96 well /cell culture plates and incubated at 37 °C for 30 min. in CO_2 incubator under humidified condition. Pseudotyped viruses (50 μL, ~100,000RLU) were added to each well and further incubated at 37 °C for 30 min. in CO_2 incubator under humidified condition. Subsequently, 1×10^4 TZM-bl cells were added to each well with 25 μg/mL of DEAE-dextran and neutralization assay was performed as described previously (Deshpande et al., 2016). The plates were incubated for 48 h and the degree of virus neutralization was assessed by measuring the luminescence (RLU). All the neutralization assays were performed in duplicates. Percent neutralization values of each peptide concentration were plotted using GraphPad Prism version 5.00 for Windows (GraphPad Software, San Diego California USA, www.graphpad.com).

Author contributions

JS, SIT, DJ, AB, JB and VP2 designed the study. JS, VP1, VN, PP recruited the study participants and collected the clinical specimens. JS, VP1, AS, BK, NH, SD performed the experiments. JS, AS and SD analyzed the data. JS and VP2 prepared the manuscript. All authors read and approved the final manuscript.

Funding

Funding for this study was provided through the Ramalingaswami fellowship received by Dr. Vainav Patel (DBT, India) and intramural funding provided by ICMR, India to Dr. Vainav Patel and Dr. Atmaram Bandivdekar. Ms. Jyoti Sutar was supported by research fellowship provided by the Lady Tata Memorial Trust (India). The HIV-1 pseudotyped virus neutralization assay carried out at THSTI was supported by the funding support from DBT National Bioscience Award grant (BT/HRD/NBA/34/0112012-13 (iv)) to Dr Jayanta Bhattacharya. The funding agencies had no role in design, execution as well as analysis of this study.

Declaration of Competing Interest

None declared.

Acknowledgements

The authors would like to thank all the study participants.

The following reagents were obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: 1) Anti-HIV-1 gp41 Monoclonal (10E8), from Dr. Mark Connors. 2) Anti-HIV-1 gp41 Monoclonal (4E10) from Polymun Scientific 3) Anti-HIV-1 gp41 Monoclonal (2F5) from Polymun Scientific (cat# 1475) 4) Anti-HIV-1 gp120 Monoclonal (VRC01), from Dr. John Mascola (cat# 12033) 5) HIV-1 CN54 gp120 Recombinant Protein from DAIDS, NIAID. The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH from Drs. John C. Kappes and Xiaoyun Wu: pSG3Δenv.; 25710-2.43 and 16055-2.3 plasmid DNA from Drs. R. Paranjape, S. Kulkarni and D. Montefiori.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.virusres.2019.197763.

References

Agnihotri, K.D., Tripathy, S.P., Jere, A.P., Kale, S.M., Paranjape, R.S., 2006. Molecular analysis of gp41 sequences of HIV type 1 subtype C from India. J. Acquir. Immune Defic. Syndr. 41, 345–351. https://doi.org/10.1097/01.qai.0000209898.67007.1a.

Ajbani, S.P., Velhal, S.M., Kadam, R.B., Patel, V.V., Bandivdekar, A.H., 2015. Immunogenicity of Semliki Forest virus based self-amplifying RNA expressing Indian HIV-1C genes in mice. Int. J. Biol. Macromol. 81, 794–802. https://doi.org/10.1016/j.ijbiomac.2015.09.010.

Alter, G., Barouch, D., 2018. Immune correlate-guided HIV vaccine design. Cell Host Microbe 24, 25–33. https://doi.org/10.1016/j.chom.2018.06.012.

Barnett, D.W., Garrison, E.K., Quinlan, A.R., Strmberg, M.P., Marth, G.T., 2011. Bamtools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics 27, 1691–1692. https://doi.org/10.1093/bioinformatics/btr174.

Bhattacharya, J., 2018. HIV prevention & treatment strategies - current challenges & future prospects. Indian J. Med. Res. 148, 671–674. https://doi.org/10.4103/0971-5916.252150.

Binley, J.M., Wrin, T., Korber, B., Zwick, M.B., Wang, M., Chappey, C., Stiegler, G., Kunert, R., Zolla-Pazner, S., Katinger, H., Petropoulos, C.J., Burton, D.R., 2004. Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. J. Virol. 78, 13232–13252. https://doi.org/10.1128/JVI.78.23.13232-13252.2004.

Blumenthal, R., Durell, S., Viard, M., 2012. HIV entry and envelope glycoprotein-mediated fusion. J. Biol. Chem. 287, 40841–40849. https://doi.org/10.1074/jbc.R112.406272.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Bricault, C.A., Yusim, K., Seaman, M.S., Yoon, H., Theiler, J., Giorgi, E.E., Wagh, K., Theiler, M., Hraber, P., Macke, J.P., Kreider, E.F., Learn, G.H., Hahn, B.H., Scheid, J.F., Kovacs, J.M., Shields, J.L., Lavine, C.L., Ghantous, F., Rist, M., Bayne, M.G., Neubauer, G.H., McMahan, K., Peng, H., Chéneau, C., Jones, J.J., Zeng, J., Ochsenbauer, C., Nkolola, J.P., Stephenson, K.E., Chen, B., Gnanakaran, S., Bonsignori, M., Williams, L.D., Haynes, B.F., Doria-Rose, N., Mascola, J.R., Montefiori, D.C., Barouch, D.H., Korber, B., 2019. HIV-1 neutralizing antibody signatures and application to epitope-targeted vaccine design. Cell Host Microbe 25https://doi.org/10.1016/j.chom.2018.12.001. 59–72.e8.

Bryson, S., Julien, J.-P., Hynes, R.C., Pai, E.F., 2009. Crystallographic definition of the epitope promiscuity of the broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2F5: vaccine design implications. J. Virol. 83, 11862–11875. https://doi.org/10.1128/JVI.01604-09.

Buchacher, A., Predl, R., Strutzenberger, K., Steinfellner, W., Trkola, A., Purtscher, M., Gruber, G., Tauer, C., Steindl, F., Jungbauer, A., Katinger, H., 1994. Generation of human monoclonal antibodies against HIV-1 proteins; electrofusion and epstein-barr virus transformation for peripheral blood lymphocyte immortalization. AIDS Res. Hum. Retroviruses 10, 359–369. https://doi.org/10.1089/aid.1994.10.359.

Burton, D.R., Ahmed, R., Barouch, D.H., Butera, S.T., Crotty, S., Godzik, A., Kaufmann, D.E., McElrath, M.J., Nussenzweig, M.C., Pulendran, B., Scanlan, C.N., Schief, W.R., Silvestri, G., Streeck, H., Walker, B.D., Walker, L.M., Ward, A.B., Wilson, I.a., Wyatt, R., 2012a. A blueprint for HIV vaccine discovery. Cell Host Microbe 12, 396–407. https://doi.org/10.1016/j.chom.2012.09.008.

Burton, D.R., Ahmed, R., Barouch, D.H., Butera, S.T., Crotty, S., Godzik, A., Kaufmann, D.E., McElrath, M.J., Nussenzweig, M.C., Pulendran, B., Scanlan, C.N., Schief, W.R., Silvestri, G., Streeck, H., Walker, B.D., Walker, L.M., Ward, A.B., Wilson, I.a., Wyatt, R., 2012b. A blueprint for HIV vaccine discovery. Cell Host Microbe 12, 396–407. https://doi.org/10.1016/j.chom.2012.09.008.

Caskey, M., Klein, F., Nussenzweig, M.C., 2019. Broadly neutralizing anti-HIV-1 monoclonal antibodies in the clinic. Nat. Med. https://doi.org/10.1038/s41591-019-0412-8.

Conway, P., Tyka, M.D., DiMaio, F., Konerding, D.E., Baker, D., 2014. Relaxation of backbone bond geometry improves protein energy landscape modeling. Protein Sci. 23, 47–55. https://doi.org/10.1002/pro.2389.

Crooks, G.E., Hon, G., Chandonia, J., Brenner, S.E., 2004. WebLogo : a sequence logo generator. Genome Res. 14, 1188–1190. https://doi.org/10.1101/gr.849004.1.

Deshpande, S., Patil, S., Kumar, R., Shrivastava, T., Srikrishnan, A.K., Murugavel, K.G., Koff, W.C., Chakrabarti, B.K., Bhattacharya, J., 2016. Association of mutations in V3/C3 domain with enhanced sensitivity of HIV-1 clade C primary envelopes to autologous broadly neutralizing plasma antibodies. Retrovirology 13, 1–8. https://doi.org/10.1186/s12977-016-0273-x.

Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D., Weingessel, A., Leisch, M.F., 2019. e1071: Misc Functions of the Department of Statistics, Probability Theory Group, (Formerly: E1071), R Software Package. Available at http://cran.rproject.org/web/packages/e1071/index.html. TU Wien. R package version 1.7-1. .

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. https://doi.org/10.1093/nar/gkh340.

Georgiev, I.S., Rudicell, R.S., Saunders, K.O., Shi, W., Kirys, T., McKee, K., O'Dell, S., Chuang, G.-Y., Yang, Z.-Y., Ofek, G., Connors, M., Mascola, J.R., Nabel, G.J., Kwong, P.D., 2014. Antibodies VRC01 and 10E8 neutralize HIV-1 with high breadth and potency even with Ig-framework regions substantially reverted to germline. J. Immunol. 192, 1100–1106. https://doi.org/10.4049/jimmunol.1302515.

Gray, E.S., Meyers, T., Gray, G., Montefiori, D.C., Morris, L., 2006. Insensitivity of paediatric HIV-1 subtype C viruses to broadly neutralising monoclonal antibodies raised against subtype B. PLoS Med. 3, e255. https://doi.org/10.1371/journal.pmed.0030255.

Haynes, B.F., Fleming, J., St Clair, E.W., Katinger, H., Stiegler, G., Kunert, R., Robinson, J., Scearce, R.M., Plonk, K., Staats, H.F., Ortel, T.L., Liao, H.-X., Alam, S.M., 2005. Cardiolipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. Science (80-.) 308, 1906–1908. https://doi.org/10.1126/science.1111781.

Hemelaar, J., Elangovan, R., Yun, J., Dickson-Tetteh, L., Fleminger, I., Kirtley, S., Williams, B., Gouws-Williams, E., Ghys, P.D., Abimiku, A.G., Agwale, S., Archibald, C., Avidor, B., Barbás, M.G., Barre-Sinoussi, F., Barugahare, B., Belabbes, E.H.,

Bertagnolio, S., Birx, D., Bobkov, A.F., Brandful, J., Bredell, H., Brennan, C.A., Brooks, J., Bruckova, M., Buonaguro, L., Buonaguro, F., Buttò, S., Buve, A., Campbell, M., Carr, J., Carrera, A., Carrillo, M.G., Celum, C., Chaplin, B., Charles, M., Chatzidimitriou, D., Chen, Z., Chijiwa, K., Cooper, D., Cunningham, P., Dagnra, A., de Gascun, C.F., Del Amo, J., Delgado, E., Dietrich, U., Dwyer, D., Ellenberger, D., Ensoli, B., Essex, M., Gao, F., Fleury, H., Fonjungo, P.N., Foulongne, V., Gadkari, D.A., García, F., Garsia, R., Gershy-Damet, G.M., Glynn, J.R., Goodall, R., Grossman, Z., Lindenmeyer-Guimarães, M., Hahn, B., Hamers, R.L., Hamouda, O., Handema, R., He, X., Herbeck, J., Ho, D.D., Holguin, A., Hosseinipour, M., Hunt, G., Ito, M., Hadj, Bel, Kacem, M.A., Kahle, E., Kaleebu, P.K., Kalish, M., Kamarulzaman, A., Kang, C., Kanki, P., Karamov, E., Karasi, J.C., Kayitenkore, K., Kelleher, T., Kitayaporn, D., Kostrikis, L.G., Kucherer, C., Lara, C., Leitner, T., Liitsola, K., Lingappa, J., Linka, M., Lorenzana de Rivera, I., Lukashov, V., Maayan, S., Mayr, L., McCutchan, F., Meda, N., Menu, E., Mhalu, F., Mloka, D., Mokili, J.L., Montes, B., Mor, O., Morgado, M., Mosha, F., Moussi, A., Mullins, J., Najera, R., Nasr, M., Ndembi, N., Neilson, J.R., Nerurkar, V.R., Neuhann, F., Nolte, C., Novitsky, V., Nyambi, P., Ofner, M., Paladin, F.J., Papa, A., Pape, J., Parkin, N., Parry, C., Peeters, M., Pelletier, A., Pérez-Álvarez, L., Pillay, D., Pinto, A., Quang, T.D., Rademeyer, C., Raikanikoda, F., Rayfield, M.A., Reynes, J.M., Rinke de Wit, T., Robbins, K.E., Rolland, M., Rousseau, C., Salazar-Gonzales, J., Salem, H., Salminen, M., Salomon, H., Sandstrom, P., Santiago, M.L., Sarr, A.D., Schroeder, B., Segondy, M., Selhorst, P., Sempala, S., Servais, J., Shaik, A., Shao, Y., Slim, A., Soares, M.A., Songok, E., Stewart, D., Stokes, J., Subbarao, S., Sutthent, R., Takehisa, J., Tanuri, A., Tee, K.K., Thapa, K., Thomson, M., Tran, T., Urassa, W., Ushijima, H., van de Perre, P., van der Groen, G., van Laethem, K., van Oosterhout, J., van Sighem, A., van Wijngaerden, E., Vandamme, A.M., Vercauteren, J., Vidal, N., Wallace, L., Williamson, C., Wolday, D., Xu, J., Yang, C., Zhang, L., Zhang, R., 2019. Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. Lancet Infect. Dis. 19, 143–155. https://doi.org/10.1016/S1473-3099(18)30647-9.

Henn, M.R., Boutwell, C.L., Charlebois, P., Lennon, N.J., Power, K.A., Macalalad, A.R., Berlin, A.M., Malboeuf, C.M., Ryan, E.M., Gnerre, S., Zody, M.C., Erlich, R.L., Green, L.M., Berical, A., Wang, Y., Casali, M., Streeck, H., Bloom, A.K., Dudek, T., Tully, D., Newman, R., Axten, K.L., Gladden, A.D., Battis, L., Kemper, M., Zeng, Q., Shea, T.P., Zedlack, C., Gasser, O., Brander, C., Hess, C., Gu, H.F., Brumme, Z.L., Brumme, C.J., Bazner, S., Rychert, J., Tinsley, J.P., Ken, H., Rosenberg, E.S., Pereyra, F., Levin, J.Z., Young, S.K., Jessen, H., Birren, B.W., Walker, B.D., Allen, T.M., 2012. Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. PLoS Pathog. 8. https://doi.org/10.1371/journal.ppat.1002529.

Hepler, N.L., Scheffler, K., Weaver, S., Murrell, B., Richman, D.D., Burton, D.R., Poignard, P., Smith, D.M., Kosakovsky Pond, S.L., 2014. IDEPI: Rapid Prediction of HIV-1 Antibody Epitopes and Other Phenotypic Features from Sequence Data Using a Flexible Machine Learning Platform. PLoS Comput. Biol. 10, e1003842. https://doi.org/10.1371/journal.pcbi.1003842.

Huang, J., Ofek, G., Laub, L., Louder, M.K., Doria-Rose, N.a, Longo, N.S., Imamichi, H., Bailer, R.T., Chakrabarti, B., Sharma, S.K., Alam, S.M., Wang, T., Yang, Y., Zhang, B., Migueles, S.a, Wyatt, R., Haynes, B.F., Kwong, P.D., Mascola, J.R., Connors, M., 2012a. Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. Nature 491, 406–412. https://doi.org/10.1038/nature11544.

Huang, J., Ofek, G., Laub, L., Louder, M.K., Doria-Rose, N.a, Longo, N.S., Imamichi, H., Bailer, R.T., Chakrabarti, B., Sharma, S.K., Alam, S.M., Wang, T., Yang, Y., Zhang, B., Migueles, S.a, Wyatt, R., Haynes, B.F., Kwong, P.D., Mascola, J.R., Connors, M., 2012b. Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. Nature 491, 406–412. https://doi.org/10.1038/nature11544.

Irimia, A., Serra, A.M., Sarkar, A., Jacak, R., Kalyuzhniy, O., Sok, D., Saye-Francisco, K.L., Schiffner, T., Tingle, R., Kubitz, M., Adachi, Y., Stanfield, R.L., Deller, M.C., Burton, D.R., Schief, W.R., Wilson, I.A., 2017. Lipid interactions and angle of approach to the HIV-1 viral membrane of broadly neutralizing antibody 10E8: Insights for vaccine and therapeutic design. PLoS Pathog. 13, e1006212. https://doi.org/10.1371/journal.ppat.1006212.

Jadhav, S.K., Velhal, S.M., Deshpande, A., Maitra, A., Chinnaraj, S., Bandivdekar, A.H., 2011. Characterization of Human Immunodeficiency Virus (HIV1C) Variants in Peripheral Blood Mononuclear Cells and Spermatozoa. J. Med. Virol. 767, 760–767. https://doi.org/10.1002/jmv.

Korber, B., Gaschen, B., Yusim, K., Thakallapally, R., Kesmir, C., Detours, V., 2001. Evolutionary and immunological implications of contemporary HIV-1 variation. Br. Med. Bull. 58, 19–42.

Kulkarni, S.S., Lapedes, A., Tang, H., Gnanakaran, S., Daniels, M.G., Zhang, M., Bhattacharya, T., Li, M., Polonis, V.R., McCutchan, F.E., Morris, L., Ellenberger, D., Butera, S.T., Bollinger, R.C., Korber, B.T., Paranjape, R.S., Montefiori, D.C., 2009. Highly complex neutralization determinants on a monophyletic lineage of newly transmitted subtype C HIV-1 Env clones from India. Virology 385, 505–520. https://doi.org/10.1016/j.virol.2008.12.032.

Kumar, R., Qureshi, H., Deshpande, S., Bhattacharya, J., 2018. Broadly neutralizing antibodies in HIV-1 treatment and prevention. Ther. Adv. Vaccines Immunother. 6, 61–68. https://doi.org/10.1177/2515135518800689.

Kwong, P.D., Mascola, J.R., Nabel, G.J., 2011. Rational design of vaccines to elicit broadly neutralizing antibodies to HIV-1. Cold Spring Harb. Perspect. Med. https://doi.org/10.1101/cshperspect.a007278.

Leaver-Fay, A., Tyka, M., Lewis, S.M., Lange, O.F., Thompson, J., Jacak, R., Kaufman, K.W., Renfrew, P.D., Smith, C.A., Sheffler, W., Davis, I.W., Cooper, S., Treuille, A., Mandell, D.J., Richter, F., Ban, Y.-E.A., Fleishman, S.J., Corn, J.E., Kim, D.E., Lyskov, S., Berrondo, M., Mentzer, S., Popović, Z., Havranek, J.J., Karanicolas, J., Das, R., Meiler, J., Kortemme, T., Gray, J.J., Kuhlman, B., Baker, D., Bradley, P., 2011. Rosetta3. pp. 545–574. https://doi.org/10.1016/B978-0-12-381270-4.00019-6.

Lee, W.P., Stromberg, M.P., Ward, A., Stewart, C., Garrison, E.P., Marth, G.T., 2014. MOSAIK: a hash-based algorithm for accurate next-generation sequencing short-read mapping. PLoS One 9. https://doi.org/10.1371/journal.pone.0090581.

Lewis, S.M., Kuhlman, B.A., 2011. Anchored design of protein-protein interfaces. PLoS One 6, e20872. https://doi.org/10.1371/journal.pone.0020872.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. 1000 Genome project data processing subgroup, 1000 genome project data processing, 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Li, L., Chen, R., Weng, Z., 2003. RDOCK: refinement of rigid-body protein docking predictions. Proteins 53, 693–707. https://doi.org/10.1002/prot.10460.

Liao, H.-X., Rebecca Lynch, T., Zhou, O., Gao, F., Alam, S.M., Boyd, S.D., Fire, A.Z., Roskin, K.M., Schramm, C.A., Zhang, Z., Zhu, J., Shapiro, L., Program, N.C.S., Mullikin, J.C., Gnanakaran, S., Mascola, J.R., Haynes, B.F., 2013. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. Nature 496, 469–476.

Kanof, Marjorie E., Smith, Phillip D., H.Z, 1996. Isolation of whole mononuclear cells from peripheral blood and cord blood. Curr. Protoc. Immunol. 1–7.

McCoy, L.E., 2018. The expanding array of HIV broadly neutralizing antibodies. Retrovirology 15, 70. https://doi.org/10.1186/s12977-018-0453-y.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, Ma, 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. https://doi.org/10.1101/gr.107524.110.

Mcmichael, A.J., Borrow, P., Tomaras, G.D., Goonetilleke, N., Haynes, B.F., 2009. The immune response during acute HIV-1 infection: clues for vaccine development. Nat. Rev. Immunol. 10, 11–23. https://doi.org/10.1038/nri2674.

Montefiori, D.C., Roederer, M., Morris, L., Seaman, M.S., 2018. Neutralization tiers of HIV-1. Curr. Opin. HIV AIDS 13, 128–136. https://doi.org/10.1097/COH.0000000000000442.

Moradigaravand, D., Kouyos, R., Hinkley, T., Haddad, M., Petropoulos, C.J., Engelstädter, J., Bonhoeffer, S., 2014. Recombination accelerates adaptation on a large-scale empirical fitness landscape in HIV-1. PLoS Genet. 10. https://doi.org/10.1371/journal.pgen.1004439.

Morikawa, Y., Overton, H.A., Moore, J.P., Wilkinson, A.J., Brady, R.L., Lewis, S.J., Jones, I.M., 1990. Expression of HIV-1 gp120 and human soluble CD4 by recombinant baculoviruses and their interaction in vitro. AIDS Res. Hum. Retroviruses 6, 765–773. https://doi.org/10.1089/aid.1990.6.765.

NACO, 2017. National Strategic Plan for HIV/AIDS and STI 2017-2024, National AIDS Control Organization. Government of India, pp. 168.

Nadai, Y., Eyzaguirre, L.M., Constantine, N.T., Sill, A.M., Cleghorn, F., Blattner, W.a, Carr, J.K., 2008. Protocol for nearly full-length sequencing of HIV-1 RNA from plasma. PLoS One 3, e1420. https://doi.org/10.1371/journal.pone.0001420.

Nelson, J.D., Brunel, F.M., Jensen, R., Crooks, T., Cardoso, R.M.F., Wang, M., Hessell, A., Wilson, I.a, Binley, J.M., Dawson, P.E., Dennis, R., Zwick, M.B., Crooks, E.T., Burton, D.R., Cardoso, R.M.F., Wang, M., Hessell, A., Wilson, Ia, Binley, J.M., Dawson, P.E., Burton, D.R., Zwick, M.B., 2007. An affinity-enhanced neutralizing antibody against the membrane-proximal external region of human immunodeficiency virus type 1 gp41 recognizes an epitope between those of 2F5 and 4E10. J. Virol. 81, 4033–4043. https://doi.org/10.1128/JVI.02588-06.

Neogi, U., Bontell, I., Shet, A., de Costa, A., Gupta, S., Diwan, V., Laishram, R.S., Wanchu, A., Ranga, U., Banerjea, A.C., Sönnerborg, A., 2012. Molecular epidemiology of HIV-1 subtypes in India: origin and evolutionary history of the predominant subtype C. PLoS One 7. https://doi.org/10.1371/journal.pone.0039819.

Nivón, L.G., Moretti, R., Baker, D., 2013. A pareto-optimal refinement method for protein design scaffolds. PLoS One 8, e59004. https://doi.org/10.1371/journal.pone.0059004.

Padte, N.N., Yu, J., Huang, Y., Ho, D.D., 2018. Engineering multi-specific antibodies against HIV-1. Retrovirology 15. https://doi.org/10.1186/S12977-018-0439-9.

Patel, V., Jalah, R., Kulkarni, V., Valentin, A., Rosati, M., Alicea, C., von Gegerfelt, A., Huang, W., Guan, Y., Keele, B.F., Bess, J.W., Piatak, M., Lifson, J.D., Williams, W.T., Shen, X., Tomaras, G.D., Amara, R.R., Robinson, H.L., Johnson, W., Broderick, K.E., Sardesai, N.Y., Venzon, D.J., Hirsch, V.M., Felber, B.K., Pavlakis, G.N., 2013. DNA and virus particle vaccination protects against acquisition and confers control of viremia upon heterologous simian immunodeficiency virus challenge. Proc. Natl. Acad. Sci. U. S. A. 110, 2975–2980. https://doi.org/10.1073/pnas.1215393110.

Patil, S., Kumar, R., Deshpande, S., Samal, S., Shrivastava, T., Boliar, S., Bansal, M., Chaudhary, N.K., Srikrishnan, A.K., Murugavel, K.G., Solomon, S., Simek, M., Koff, W.C., Goyal, R., Chakrabarti, B.K., Bhattacharya, J., 2016. Conformational epitope-specific broadly neutralizing plasma antibodies obtained from an HIV-1 clade C-infected elite neutralizer mediate autologous virus escape through mutations in the V1 loop. J. Virol. 90, 3446–3457. https://doi.org/10.1128/jvi.03090-15.

Payne, R., Muenchhoff, M., Mann, J., Roberts, H.E., Matthews, P., Adland, E., Hempenstall, A., Huang, K.-H., Brockman, M., Brumme, Z., Sinclair, M., Miura, T., Frater, J., Essex, M., Shapiro, R., Walker, B.D., Ndung'u, T., McLean, A.R., Carlson, J.M., Goulder, P.J.R., 2014. Impact of HLA-driven HIV adaptation on virulence in populations of high HIV seroprevalence. Proc. Natl. Acad. Sci. 5393–5400. https://doi.org/10.1073/pnas.1413339111.

Pegu, A., Yang, Z.Y., Boyington, J.C., Wu, L., Ko, S.Y., Schmidt, S.D., McKee, K., Kong, W.P., Shi, W., Chen, X., Todd, J.P., Letvin, N.L., Huang, J., Nason, M.C., Hoxie, J.A., Kwong, P.D., Connors, M., Rao, S.S., Mascola, J.R., Nabel, G.J., 2014. Neutralizing antibodies to HIV-1 envelope protect more effectively in vivo than those to the CD4 receptor. Sci. Transl. Med. 6. https://doi.org/10.1126/scitranslmed.3008992.

Pierce, B., Weng, Z., 2007. ZRANK: reranking protein docking predictions with an optimized energy function. Proteins Struct. Funct. Genet. 1086, 1078–1086. https://doi.org/10.1002/prot.

Possas, C., de S. Antunes, A.M., Lins Mendes, F.M., Veloso, V., Martins, R.M., Homma, A.,

2018. HIV cure: global overview of bNAbs' patents and related scientific publications. Expert Opin. Ther. Pat. 28, 551–560. https://doi.org/10.1080/13543776.2018.1495708.
R Core Team, 2018. R: a Language and Environment for Statistical Computing.
Kolde, Raivo, 2018. pheatmap: Pretty Heatmaps.
Raveh, B., London, N., Schueler-Furman, O., 2010. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. Proteins Struct. Funct. Bioinforma. 78, 2029–2040. https://doi.org/10.1002/prot.22716.
Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: the European molecular biology open software suite. Trends Genet. 16, 276–277.
Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011. Integrative genomics viewer. Nat. Biotechnol. 29, 24–26. https://doi.org/10.1038/nbt0111-24.
Rong Chen, Z.W., 2003. A novel shape complementarity scoring function for protein-Protein docking. Proteins Struct. Funct. Genet. 408, 397–408.
RStudio Team, 2015. RStudio: Integrated Development for R.
Rujas, E., Gulzar, N., Morante, K., Tsumoto, K., Scott, J.K., Nieva, J.L., Caaveiro, J.M.M., 2015. Structural and thermodynamic basis of epitope binding by neutralizing and nonneutralizing forms of the Anti-HIV-1 antibody 4E10. J. Virol. 89, 11975–11989. https://doi.org/10.1128/JVI.01793-15.
Salzwedel, K., West, J.T., Hunter, E., 1999. A conserved tryptophan-rich motif in the membrane-proximal region of the human immunodeficiency virus type 1 gp41 ectodomain is important for Env-mediated fusion and virus infectivity. J. Virol. 73, 2469–2480. https://doi.org/0022-538X/99/$04.00+0.
Sangeda, R.Z., Theys, K., Beheydt, G., Rhee, S.Y., Deforche, K., Vercauteren, J., Libin, P., Imbrechts, S., Grossman, Z., Camacho, R.J., Van Laethem, K., Pironti, A., Zazzi, M., Sönnerborg, A., Incardona, F., De Luca, A., Torti, C., Ruiz, L., Van de Vijver, D.aM.C., Shafer, R.W., Bruzzone, B., Van Wijngaerden, E., Vandamme, A.M., 2013. HIV-1 fitness landscape models for indinavir treatment pressure using observed evolution in longitudinal sequence data are predictive for treatment failure. Infect. Genet. Evol. 19, 349–360. https://doi.org/10.1016/j.meegid.2013.03.014.
Schneider, T.D., Stephens, R.M., 1990. Sequence logos: a new way to display consensus sequences. Nucleic Acids Res. 18, 6097–6100. https://doi.org/10.1093/nar/18.20.6097.
Shapiro, S.Z., 2019. Lessons for general vaccinology research from attempts to develop an HIV vaccine. Vaccine. https://doi.org/10.1016/j.vaccine.2019.04.005.
Siepel, A.C., Halpern, A.L., Macken, C., Korber, B.T.M., 1995. A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. AIDS Res. Hum. Retroviruses 11, 1413–1416. https://doi.org/10.1089/aid.1995.11.1413.
Song, L., Sun, Z.-Y.J., Coleman, K.E., Zwick, M.B., Gach, J.S., Wang, J., Reinherz, E.L., Wagner, G., Kim, M., 2009. Broadly neutralizing anti-HIV-1 antibodies disrupt a hinge-related function of gp41 at the membrane interface. Proc. Natl. Acad. Sci. U. S. A. 106, 9057–9062. https://doi.org/10.1073/pnas.0901474106.
Stiegler, G., Kunert, R., Purtscher, M., Wolbank, S., Voglauer, R., Steindl, F., Katinger, H., 2001. A potent cross-clade neutralizing human monoclonal antibody against a novel epitope on gp41 of human immunodeficiency virus type 1. AIDS Res. Hum. Retroviruses 17, 1757–1765. https://doi.org/10.1089/08892220152741450.
Thorvaldsdóttir, H., Robinson, J.T., Mesirov, J.P., 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. 14, 178–192. https://doi.org/10.1093/bib/bbs017.
Tyka, M.D., Keedy, D.A., André, I., Dimaio, F., Song, Y., Richardson, D.C., Richardson, J.S., Baker, D., 2011. Alternate states of proteins revealed by detailed energy landscape mapping. J. Mol. Biol. 405, 607–618. https://doi.org/10.1016/j.jmb.2010.11.008.
UNAIDS, 2018. The Joint United Nations Programme on HIV/AIDS (UNAIDS) Data 2018. UNAIDS 1–376.
Wagh, K., Bhattacharya, T., Williamson, C., Robles, A., Bayne, M., Garrity, J., Rist, M., Rademeyer, C., Yoon, H., Lapedes, A., Gao, H., Greene, K., Louder, M.K., Kong, R., Karim, S.A., Burton, D.R., Barouch, D.H., Nussenzweig, M.C., Mascola, J.R., Morris, L., Montefiori, D.C., Korber, B., Seaman, M.S., 2016. Optimal combinations of broadly neutralizing antibodies for prevention and treatment of HIV-1 clade C infection. PLoS Pathog. 12, e1005520. https://doi.org/10.1371/journal.ppat.1005520.
Wagh, K., Seaman, M.S., Zingg, M., Fitzsimons, T., Barouch, D.H., Burton, D.R., Connors, M., Ho, D.D., Mascola, J.R., Nussenzweig, M.C., Ravetch, J., Gautam, R., Martin, M.A., Montefiori, D.C., Korber, B., 2018. Potential of conventional & bispecific broadly neutralizing antibodies for prevention of HIV-1 subtype A, C & D infections. PLoS Pathog. 14, e1006860. https://doi.org/10.1371/journal.ppat.1006860.
Wang, Y.H., Davies, A.H., Jones, I.M., 1995. Expression and purification of glutathione S-transferase-tagged HIV-1 gp120: no evidence of an interaction with CD26. Virology 208, 142–146. https://doi.org/10.1006/viro.1995.1137.
Wu, X., Yang, Z.-Y., Li, Y., Hogerkorp, C.-M., Schief, W.R., Seaman, M.S., Zhou, T., Schmidt, S.D., Wu, L., Xu, L., Longo, N.S., McKee, K., O'Dell, S., Louder, M.K., Wycuff, D.L., Feng, Y., Nason, M., Doria-Rose, N., Connors, M., Kwong, P.D., Roederer, M., Wyatt, R.T., Nabel, G.J., Mascola, J.R., 2010. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science 329, 856–861. https://doi.org/10.1126/science.1187659.
Yang, X., Charlebois, P., Gnerre, S., Coole, M.G., Lennon, N.J., Levin, J.Z., Qu, J., Ryan, E.M., Zody, M.C., Henn, M.R., 2012. De novo assembly of highly diverse viral populations. BMC Genomics 13, 475. https://doi.org/10.1186/1471-2164-13-475.
Yang, X., Charlebois, P., Macalalad, A., Henn, M.R., Zody, M.C., 2013. V-Phaser 2: variant inference for viral populations. BMC Genomics 14, 674. https://doi.org/10.1186/1471-2164-14-674.
Yoon, H., Macke, J., West, A.P., Foley, B., Bjorkman, P.J., Korber, B., Yusim, K., Yusim, K., 2015. CATNAP: a tool to compile, analyze and tally neutralizing antibody panels. Nucleic Acids Res. 43, W213–9. https://doi.org/10.1093/nar/gkv404.
Zwick, M.B., Labrijn, A.F., Wang, M., Saphire, E.O., Binley, J.M., Moore, P., Stiegler, G., Katinger, H., Dennis, R., Parren, P.W.H.I., Spenlehauer, C., Moore, J.P., Burton, D.R., 2001. Broadly neutralizing antibodies targeted to the membrane-proximal external region of human immunodeficiency virus type 1 glycoprotein gp41. J. Virol. 75, 10892–10905. https://doi.org/10.1128/JVI.75.22.10892.