BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 11(2)

Neutralization function affected by single amino acid replacement in the HIV-1 antibody targets

Johnson Christdas[1], Prabu Manoharan[2] & Shakila Harshavardhan[1]*

[1]Department of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai, Tamil Nadu, India, 625021; [2]Centre of Excellence in Bioinformatics, Madurai Kamaraj University, Madurai, Tamil Nadu, India, 625021; Shakila Harshavardhan – Email: mohanshakila@yahoo.com; *Corresponding author

Received December 01, 2014; Accepted January 08, 2015; Published February 28, 2015

Abstract:
The viral envelope glycoproteins are essential for entry into their host cells and studied extensively for designing vaccines. We hypothesize that the glycosylation on the HIV-1 viral envelope glycoprotein 41(gp41) at critical residues offers viral escape from the specific immune surveillant neutralizing antibodies Z13, 4E10 and 10E8 targeted to their linear epitopes in the Membrane Proximal External Region (MPER). The glycosylation occurring on the 50th residue (Asparagine) contained in the target (NWFNIT) can mask itself to be inaccessible for these neutralizing antibodies. The glycosylation rate of the epitopes which are shared by the Z13, 4E10 and 10E8 neutralizing antibodies of HIV-1 were predicited *in silico*. We analyzed the reliable frequency of glycosylation on the HIV-1 envelope gp41 using prediction tools to unravel the plausibility of the glycosylation by a mannose at 50th residue in the 59 amino acid long HIV-gp41 trimer (PDBID: 2M7W and 2LP7). It is evident that the glycosylation by a mannose that masks these targets is possible only when the 50th amino-acid is N (Asparagine, Asn) which is not possible when N is mutated to D (Aspartatic acid, Asp). The additive advantage for the retrovirus is its error-prone reverse transcriptase which can choose to copy these survivable mutants with Asn N-50 that can be glycosylated as explained by the Copy-choice model. So the glycan shields varying in their intensity and patterns have to be essentially studied to understand the viral escape strategies that will give a way forward towards a successful vaccine that can elicit a neutralizing antibody response to confer protection.

Keywords: N- linked glycosylation, HIV gp41, Immune escape, Neutralizing antibodies, Immunogens.

Background:
The viral envelope proteins are used as conventional vaccine targets. Apart from the primary and the other structures of proteins, glycosylation is an essential feature of any protein that gives its complete functionality. This occurs as a part of post-translational modifications. The N-Linked glycosylation of the protein is a remarkable feature occurring by the programs in the chaperones, which determines the trafficking of the protein either to be secreted or localized to its target **[1].**

The varied glycosylation pattern and degree can modulate the protein function. This feature is found in most taxa as their taxonomic character **[2].** The alteration in the glycosylation patterns of the immunoglobulins ultimately impairs the viral entry and replication influencing the susceptibility to the viral infections **[3].** These fore-running reports on glycosylation suggest the need for enhancing knowledge on glycosylation of viral envelope proteins for vaccine design. High throughput analysis of glycans is evolving for meeting the diagnostic and therapeutic requirements. Genetics associated with this phenomenon are also being investigated across populations as it may have a clinical impact in handling the diseases **[4].**

Glycosylation on the viral envelope glycoproteins affects their interaction with the receptors on the host cells through which they enter into the cells. The enveloped viruses can evade the immune system by exploiting the glycans as shields that covers the antibody targets. The glycosylation on the envelope

proteins of HIV-1 and Human *Influenza* Hemagglutinin, helps them in escaping from the host immune recognition as they are the sole targets of the humoral immune response. Moreover, engineering the N-linked glycosylation sites of Hepatitis C Virus (HCV) envelope glycoproteins alters the T cell response **[5].** It has been already shown that the differences in the N-linked glycosylation of HIV and Ebola virus glycoproteins modulate the DC-SIGN (Dendritic Cell Specific Intercellular adhesion molecule Grabbing Non-integrin) and its receptor **[6].** The envelope glycoproteins of HIV-1 not only feature antigenicity but also the structure, biology, tropism and the course of the viral evolution **[7].** The biology of the virus starts with the entry into the host cell which is possible only with these glycoproteins **[8].** Beyond this, they also offer resistance for the virus against the innate components of the host immune system **[9].** For example, the glycoproteins of the *Influenza* Virus that causes flu enables its firm adherence onto the mucous membrane, prevents it from being eliminated by the ciliated epithelial cells. The sites of glycosylation on the envelope glycoproteins of West Nile virus can be mutated for developing alternative forms of vaccines that will be more efficient **[10].** The prophylactic vaccines for the viruses mentioned above (HIV, H. Influenza, HCV, Ebola) are yet to be achieved. Hence, a computational study that focuses on glycosylation of envelope protein, would pave way for efficient vaccine designing strategies.

The HIV-1 Env is a heavily glycosylated trimeric protein having three identical surface gp120 molecules, each noncovalently associated with a transmembrane gp41 molecule **[11].** Most of the HIV-1 glycoproteins are N-glycosylated in the regions that are putative targets for the vaccine design and it is proven in the SIV – gp41 **[12].** The deglycosylation of these targets had shown increase in neutralization **[13].** With the expanding and advanced knowledge and technology, the field of medicine is still impaired for about three decades with its struggle for a successful vaccine that can address to the epidemiological status of HIV-1 infection that produces at least 6500 new infections daily **[14].** The pathogen either as a whole or a part had been used as a successful vaccine in the conventional preventive strategies. But such a vaccine for HIV-1 is not yet achieved. The viral diversity with its other immune evasion strategies are being investigated. The glycosylation pattern on the viral targets being a strategy for the viral evasion must be studied extensively for designing a successful vaccine **[15].** The Z13, 4E10 and 10E8 neutralizing antibodies of HIV-1 are putatively targeting the regions, SLWNWFNITN NWFDIT and NWFDISNWLWYIK **[16]** found in the transmembranous HIV-gp41. By employing computational methods, these neutralizing antibody targets were investigated in this paper for their glycosylation patterns and the mutations that favor such patterns.

Methodology:

Sequence Analysis

FASTA format of six amino acid sequences, three chains of two trimers, HIV-1 envelope glycoproteins i.e., 2M7W and 2LP7 from PDB containing the target for 4E10, 10E8 and Z13 antibodies were downloaded. 2M7W and 2LP7 are trimers and each chain (Chain A, B, C) constituting the trimer was taken as a single sequence and was aligned using ClustalW feature of Bioedit v7.0.9 and was stored as FASTA file. It was then uploaded into the N Glycosite, an online tool available at http://www.hiv.lanl.gov/content/sequence/GLYCOSITE/glycosite.html **[17].** A typical glycosylation site will have N-X-S/T and a potential site for glycosylation shown in **Figure 1.** The glycosylation sites found in the NWFDIT, NWFDISNWLWYIK and SLWNWFNITN targeted by 4E10, 10E8 and Z13 antibodies respectively were studied using this online tool. The D (Asp) was substituted for N (Asn) at the 50th position that contains the targets for these anitbodies. The glycosylation at position 50 with either N (Asn) or D (Asp) were investigated using the above mentioned online tool. The glycosylation pattern for the same dataset was also checked for the glycosylation pattern using another such tool NetNGlycServer 1.0 available at http://www.cbs.dtu.dk /services/NetNGlyc/ to check the reliability of the results.

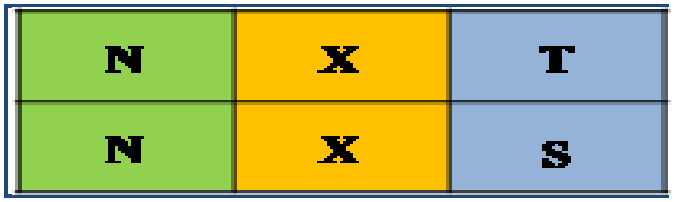

Figure 1: The potential glycosylation found in any protein must have the above sequence where N is Asparagine; X is any amino acid except Proline; S is Serine and T is Threonine.

Protein and Ligand Preparation for Docking

The PDB of the HIV –gp41 homotrimer structure was hydrogen adjusted at neutral pH. The mutant structure of this protein was prepared by substituting 50th amino acid (Asn) glycosylation site on chains A, B and C to aspartic acid. This was achieved with the use of SPDB viewer version 4.1 software **[18].** The 3D structure of glycan (mannose) was retrieved from pubchem database. The ligands protonation state was assigned at neutral pH and ligand structure optimization was carried out using biopolymer module of InsightII software. The energy optimization of protein and ligand was carried out using 1000 steps of steepest descent followed by 2000 steps of conjugate gradient energy minimization. This energy optimized protein and ligand structure was used for covalent docking.

Covalent Docking

The covalent docking of mannose to HIV-gp41 (unmutated and mutated) homotrimer structures was carried out using Gold suite of docking software **[19].** Gold suite is well known docking software which uses Genetic Algorithm (GA) to dock ligands flexibly to the binding site of protein **[20].** The docking settings were optimized by performing several rounds of trial and error method to obtain the most accurate docking result. The binding site was defined upto 6 Å radius, by keeping the center of c-alpha carbon atom of asparagine and the aspartic acid as the 50th residue found in the unmutated and mutated HIV-gp41, respectively. The asparagine side chain amino group (aspartic acid carboxyl oxygen in case of mutated HIV-gp41) present with a lone pair electron was chosen for linking with an atom in mannose, during the covalent docking study. The GA run was set to be very flexible to find the most accurate binding pose. The number of binding poses to generate during docking was set to 20 and top ranked solution was retained. The ChemPLP scoring function was used to rank and select the docking poses.

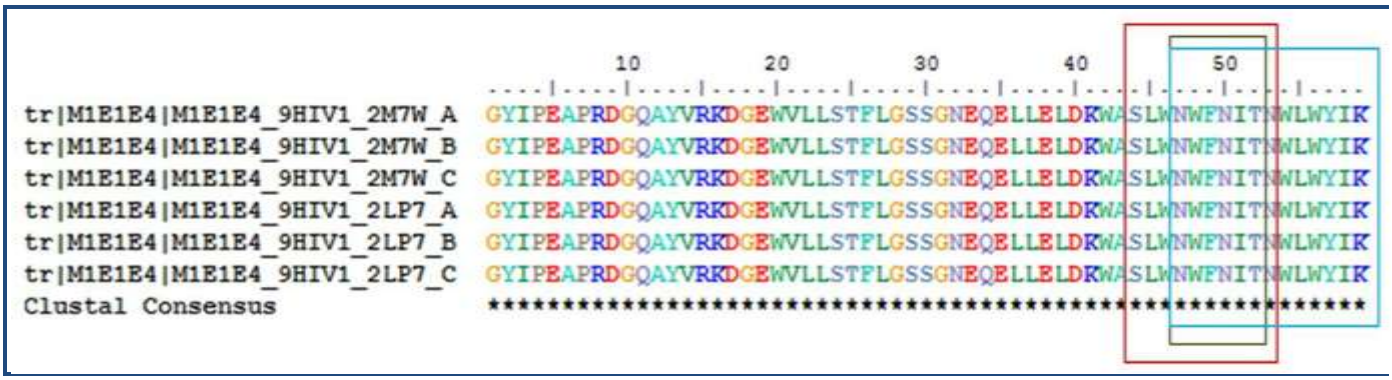

Figure 2: Multiple Sequence alignment of the monomers (Chain A, B, C) of HIV-1 gp41 homotrimer (2M7W and 2LP7) is shown here. The epitopes of the neutralizing antibodies Z13, 4E10 and 10E8 are marked in red, green and cyan colored boxes respectively.

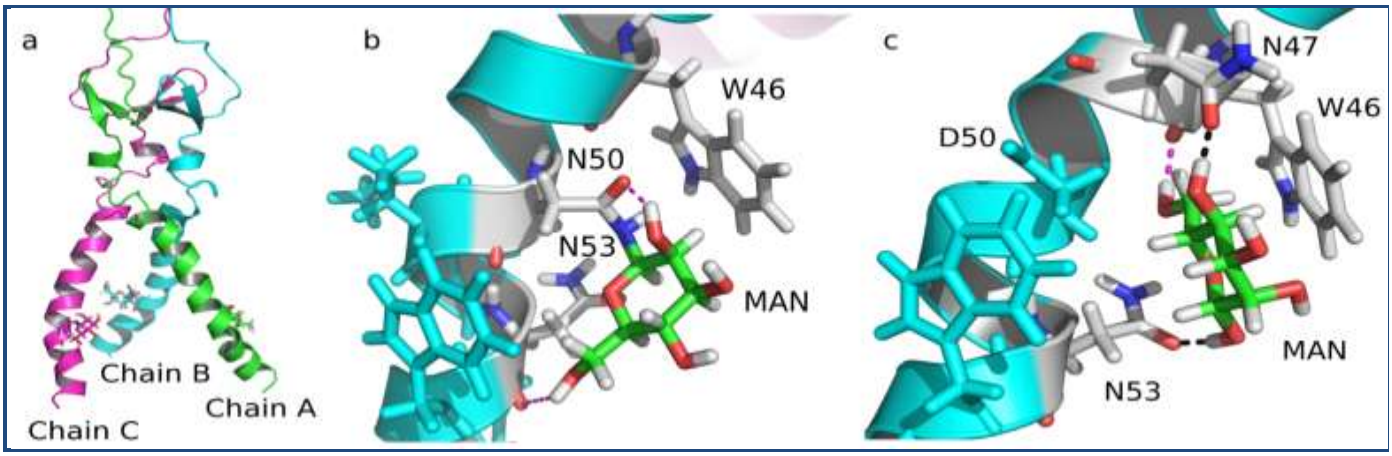

Figure 3: a) The HIV gp41 (PDB id: 2M7W) trimer structure with mannose covalently linked to N-50 is shown as cartoon model, the mannose is shown as stick model and colored according to chain information. In figure b and c the binding site residues within 6 Å radius of mannose binding is shown as stick model. The binding site residues interacting with mannose (atom color) is colored grey, other binding site residues are colored cyan. Hydrogen bonds and short contacts between protein and ligand are shown as meganta and black dash. **b)** Mannose covalently interacts with N-50 and non covalently with other binding site residues of wild gp41. c) Mannose non covalently interacts with the binding site residues of mutated gp41.

Result:

The alignment of the sequences of 3 monomers (Chains A, B and C) that are forming the trimer HIV-1 gp41 was obtained. The epitopes of Z13, 4E10 and 10E8 antibodies are enclosed in the red, green and cyan boxes in the alignment as shown in the **Figure 2.**

The results of the prediction of glycosylation on every chain of the trimer are represented in **Table 1 (see supplementary material)** that shows the estimated frequency of glycosylation at the potential sites. The N-50 located in the shared targets for Z13, 4E10 and 10E8 antibodies is the only potential site of glycosylation found on each chain forming the trimer and the frequency of glycosylation with both N and D at position 50 was calculated **(Table 1).**

The chemical interactivity of the mannose moiety with the asparagine and the aspartic acid at the 50th position in their respective forms are depicted in the **Figure 3b** & **3c** respectively.

Discussion:

In order to study the glycosylation frequency on HIV-1 gp41 trimer, each monomer sequence (Chains A, B, C) and their trimer structures were used in these analyses. The template for investigation of glycosylation frequency was chosen from PDB Id 2M7W and 2LP7. The sequence identities of the monomers were checked by the multiple sequence alignment as shown in **Figure 2.** All the 59 aminoacids in all three chains of both the trimers 2M7W and 2LP7 are identical with the N-linked glycosylation site predicted at the 50th residue found in the shared epitopes of the neutralizing antibodies Z13, 4E10 and 10E8 as shown in **Figure 2** which are colored in red, green and cyan respectively and these epitopic regions are overlapping.

The N-glycosite prediction site has located the glycosylable asparagine residue at the 50th postion in the neutralizing antibody targets. The predicted glycosylation site N-50 matches the profile in **Figure 1.** This interaction on one of the chains (Chain B of 2M7W) that forms the trimer has been represented in the figure 3b. As all the monomers are identical, the glycosylation depicted in one chain is possible on all the three monomers resulting in efficient masquerade from the immune-surveillance and thus can help the virus to escape. The recognition by both autologous and the epitope specific monoclonal antibodies **[15]** is possible due to the "glycan shield". Contrarily, the absence of such carbohydrates increases the neutralizing ability of these antibodies against SIV and

affects the infectivity also **[8, 12].** Experimentally, deglycosylation at specific residues on the envelope glycoprotein has been shown to have increased infectivity and antibody affinity **[13].**

To substantiate our sequence based prediction, structural mutation of glycosylation site in HIV-gp41 was carried out. Further, we performed covalent docking to unravel the possibility of N-linked glycosylation and its mechanism in the presence of both N-50 and D-50 which are unmutated and mutated forms respectively. It is evident that the covalent linkage is not possible in the presence of D at 50th position **(Figure 3b & 3c).** The covalent docking for unmutated HIV-gp41 resulted in covalently bound 20 different poses, among which the top ranked pose with highest CHEMPLP score was chosen for this study and shown in **Figure 3b**. The top ranked pose of mannose bound to wild and mutated gp41 showed van der Waals clashes along with hydrogen bond. The van der Waal's clashes were removed by energy minimization **(Details shown in supplementary material).**

The docking predictions shows that 1st carbon present in the mannose sugar moiety forms a glycosyl linkage (bond) with the amino group of the carboxylamine moiety on the side chain of N-50 **(Figure 3b).** On the other hand, due to the presence of carboxyl group in the side chain of D-50 of the mutated HIV-gp41, it failed to form the glycosyl linkage **(Figure 3c).** In addition to the covalent linkage, we also observed other non bonded interactions, i.e. hydrogen bonds and C-H…π -interactions between the mannose and active site residues. In the wild gp41-mannose complex **(Figure 3b)**, hydroxyl group linked to the 2nd carbon donates electron to oxygen atom of carboxamide group of N-50. The hydroxyl group of hydroxymethyl moiety attached to the 5th carbon of mannose ring donates an electron to form a hydrogen bond with the backbone carbonyl oxygen of N-53 as acceptor. The hydrogen atom linked to the 2nd carbon of the mannose ring forms a C-H…π interaction with the aromatic side chain of W-46. In contrast, mannose bound to mutated gp41 as shown in **Figure 3c**, the aromatic ring of W-46 involves in a π…C-H interaction with hydrogen attached to the 4th carbon of mannose ring. The hydroxyl group of hydroxymethyl moiety attached to the 5th carbon of mannose ring donates an electron to form a hydrogen bond with the backbone carbonyl oxygen of W-46 as acceptor. Apart from these interactions, mannose involves in two short contacts with the carbonyl oxygen of N-53 and N-47. The side chain of Aspartic acid at the 50th position in mutated gp41 is positioned away from the mannose as compared to the Aparagine of wild gp41, which shows covalent interaction with mannose. Moreover the CHEMPLP score for unmutated HIV-gp41-mannose covalent complex resulted in a better score, when compared to the mutated HIV-gp41-mannose noncovalent complex. This underscores that the glycosylation is feasible with N-50 and not with D-50.

In an *in vivo* scenario, the pressure of these surveillant antibodies can drive the viral mutations to get glycosylated with asparagines as the 50th amino acid instead of aspartic acid or any other. Such survivable variants are selectively preferred by the error prone reverse transcriptase according to the copy choice model as they can escape. The similar side chains of the two aminoacid enables all the other functions of the residue except for glycosylation. This feature comes in hand with the 'No fidelity' reverse transcriptase which can make progeny of these variants to cause the antigenic drift in HIV-1 that helps immune evasion. This unparalleled genetic plasticity of the retrovirus HIV-1 sets it apart from all other classes and families of the viruses to survive inspite of the immune pressure and therapeutic agents **[21].** So it is essential to study the glycosylation mechanisms along with viral diversity and all other factors that pose a hurdle to achieve a successful vaccine.

Conclusion:

In conclusion, the variable degree and the pattern of glycosylation plausibly affect the viral entry, tropism, immune recognition or evasion and play a key role in modulating the immune response or the viraemia and ultimately the disease progression in the host. These data entail that the glycosylation occurs only in the presence of N-50 and it is not possible with D-50 can potentially mask the targets from being recognized by the neutralizing antibodies. A major obstacle in achieving a successful vaccine to prevent AIDS is due to the enormous diversity and mutability of the virus that offers escape from the immune surveillance by the neutralizing antibodies. Glycosylation of the envelope glycoproteins in a way thus favors the virus to escape the surveillance by these circulating neutralizing antibodies. This varying glycosylation of the viral epitopes in concert with the error prone reverse transcriptase can make progeny of these mutant forms to escape the immune recognition. Thus, the glycosylation together with the genetic plasticity offers a survival advantage for the virus. These data on glycosylation frequency concord with the experimental reports, apparently shows that glycosylation interferes with the recognition and affinity of the neutralizing antibodies and their respective targets. Hence it is obvious that one should consider the glycosylation to study the immune evasion mechanisms strategies to design successful vaccines against the viral diseases.

Acknowledgement:

The authors thank KM Saravanan, Post Doctoral Fellow at the Centre of Excellence in Bioinformatics, Madurai Kamaraj University for his suggestions. This work was supported by the University Grants Commission of India and was carried out with the facilities in the Centre of Excellence in Bioinformatics, Madurai Kamaraj University sponsored by the Department of Biotechnology, Government of India. M.P thanks University Grants Commision (UGC) for providing Post Doctoral Fellowship under Dr.D.S.Kothari PDF Scheme. The authors thank S.Krishnawamy, Centre Of Excellence in Bioinformatics, for helping us with the facilities in the centre.

References:

[1] Ferris SP *et al. Dis Model Mech*. 2014 **7**: 331 [PMID: 24609034]
[2] Giltner CL *et al. Microbiol Mol Biol Rev*. 2012 **76:** 740 [PMID: 23204365]
[3] Sadat MA *et al. N Engl J Med*. 2014 **1:** 11 [PMID: 24716661]
[4] Huffman JE *et al. Mol Cell Proteomics.* 2014 **13:** 1598 [PMID 24719452]
[5] Li P *et al. Vaccine*. 2007 **25:** 1544 [PMID: 17055129]
[6] Lin G *et al.J Virol*. 2003 **77:** 1337 [PMID: 12502850]
[7] Wyatt R *et al. Science*. 1998 **280**: 1884 [PMID: 9632381]
[8] Polzer S *et al. Virology*. 2002 **304**: 70 [PMID: 12490404]

[9] Balla-Jhagjhoorsingh SS *et al. PLoS One*. 2013 **8**: e68863 [PMID: 23874792]
[10] Whiteman MC *et al. Vaccine*. 2010 **28:** 1075 [PMID: 19896447]
[11] Mascola JR & Montefiori DC, *Annu Rev Immunol*. 2010 **28**: 413 [PMID: 20192810]
[12] Martinez-Navio JM & Desrosiers RC, *Journal of Virology*. 2012 **86:** 12484 [PMID: 22993152]
[13] Ma BJ *et al. PLoS Pathog*. 2011 **7**: e1002200 [PMID: 21909262]
[14] Fauci AS *et al. Science*. 2008 **321:** 530 [PMID: 18653883]
[15] Wei X *et al. Nature*. 2003 **422:** 307 [PMID: 12646921]
[16] Huang J *et al. Nature*. 2012 **491**: 406[PMID: 23151583]
[17] Zhang M *et al. Glycobiology*. 2004 **14:** 1229 [PMID: 15175256]
[18] Guex N *et al. Electrophoresis*. 2009 **30:** S162 [PMID: 19517507]
[19] Jones G *et al. J Mol Biol*. 1995 **245**: 43 [PMID: 7823319]
[20] Verdonk ML *et al. Proteins Struct Funct Genet*. 2003 **52:** 609 [PMID: 12910460]
[21] Worobey M & Holmes EC, *J Gen Virol*. 1999 **80:** 2535 [PMID: 10573145]

Edited by P Kangueane
Citation: **Christdas *et al.*** Bioinformation 11(2): 057-062 (2015)

BIOINFORMATION open access

Supplementary material:

Energy Minimization:
It is not surprising that computational docking of small molecule to protein, will result in van der Waals clashes. Especially during covalent docking, where the ligand degree of freedom is restricted due to the linkage of ligand atom to the residue in the protein, that is treated rigidly during the docking study, because only the ligand and OH groups of residues in binding site are flexibly considered during docking in Gold. The wild and mutated gp41-mannose complex was subjected for a controlled energy minimization. The heavy atoms of binding site residues and mannose was restrained initially and gradual reduction of force constant was applied to reach a fully relaxed energy minimization of mannose and binding site. This removed van der Waals clashes between the mannose and binding site residues within 6 Å radius and at the same time preserved the other nonbonded interactions. The energy minimized wild and unmutated gp41-mannose complex was used for further interaction analysis.

Table 1: The glycosylation frequency predicted by the N-glycosite server provided at http:/ /www.hiv.lanl.gov/content/sequence /GLYCOSITE/glycosite.html with N and D at the 50th position of the monomeric chains that form the trimer. Glycosylation frequency possible at N-50 onevery chain that forms the trimer.

Monomeric chains that form the HIV- gp41 homotrimer (PDB Ids 2M7W and 2LP7)			**Amino acids at 50th position**	**Glycosylation frequency**
Biological Source	**PDB Id**	**Chain**		
HIV-1 Envelope glycoprotein41	2M7W	A	N	1
			D	0
HIV-1 Envelope glycoprotein41	2M7W	B	N	1
			D	0
HIV-1 Envelope glycoprotein41	2M7W	C	N	1
			D	0
HIV-1 Envelope glycoprotein41	2LP7	A	N	1
			D	0
HIV-1 Envelope glycoprotein41	2LP7	B	N	1
			D	0
HIV-1 Envelope glycoprotein41	2LP7	C	N	1
			D	0