Insights from the interfaces of HIV-1 envelope (ENV) trimer viral protein GP160 (GP120-GP41)

Christina Nilofer, Arumugam Mohanapriya

School of Biosciences and Technology, Vellore Institute of Technology, Vellore Campus, Tiruvalam Road, Katpadi, Vellore, Tamil Nadu - 632014, India

Article History:
Received on: 02 Dec 2020
Revised on: 31 Dec 2020
Accepted on: 02 Jan 2021

Keywords:
HIV-1, Envelope, Trimer, Glycoprotein, GP160, GP120, GP41, Protein Interface, Van Der Waals

ABSTRACT

The Human Immunodeficiency Virus (HIV-1) type 1 viral protein is a life threatening virus causing HIV/AIDS in infected humans. The HIV-1 envelope (ENV) trimer glycoprotein GP160 (GP120-GP41) is gaining attention in recent years as a potential vaccine candidate for HIV-1/AIDS. However, the sequence variation and charge polarity at the interacting sites across clades is a short-coming faced in the development of an effective HIV-1 vaccine. We analyzed the interfaces in terms of its interface area, interface size, and interface energies (van der Waals, hydrogen bonds, and electrostatics). The interfaces were divided as dominant (≥60%) and subdominant (<60%) based on van der Waals contribution to total energies. 88% of GP120 and 74% of GP41 interfaces are highly pronounced with van der Waals energy having large interfaces with interface size (98±65 (GP120) and 73±65 (GP41)) and interface area (882±1166Å² (GP120) and 921±1288Å² (GP41)). Nevertheless, 12% of GP120 and 26% of GP41 interfaces have subdominant van der Waals energies having small interfaces with interface size (58±20 (GP120) and 27±9 (GP41)) and interface area (581±1605Å² (GP120) and 483±996Å² (GP41)). It was interesting to observe GP41 small interfaces with subdominant van der Waals are stabilized by electrostatics (r²=0.63) without hydrogen bonds (r²=0). However, GP120 small interfaces were found to have two fold more hydrogen bonds (r²=0.59) than electrostatics (r²=0.20). Therefore, our previous finding stating that small protein-protein interfaces rich in electrostatics holds true in case of GP41 whereas not with GP120 protein interfaces.

INTRODUCTION

Regardless of the remarkable efforts to develop a vaccine for HIV-1/AIDS has always been a great challenge over the last decade with disappointing results in clinical trials (Shin, 2016). The unsatisfactory clinical trial results from VaxGen’s AIDSVAXgp120 vaccine and MRKAd5 HIV-1 Gag/Pol/Nef have been discussed elsewhere (Adis Editorial, 2003; Überla, 2008). This could be due to the viral human molecular mimicry, protein structural architecture, viral protein mutation and glycosylation. Despite the serious biotechnological challenges there is always an amplified energy to synthesis ENV trimer spike protein. The reasonable efficacy shown in the Thai trail vaccine (RV144 - ENV-GP120, Gag and Pro) is promising (Rerks-Ngarm et al., 2009, 2013). Post Thai trial (RV144), the focus is on envelope (ENV) as a vaccine candidate. In addition, ENV GP160 with least homology is selected by performing a sequence
comparison between HIV and human proteome (Kangueane et al., 2008). GP160 ENV trimer spike glycoprotein has gained attention as a potential vaccine candidate in the recent years. Production of native like HIV-1 envelope trimer glycoprotein is a challenge in designing, developing and validating an effective vaccine from a biochemical, structural and immunological viewpoint (Sanders and Moore, 2017; Doores, 2015).

Ringe et al. investigated on the number of factors that are importantly influencing the design, stability and purification of native like HIV-1 envelope trimer glycoprotein. Alsalmi et al. used strept tag method to purify GP160 trimer protein and was resulted with cleaved, uncleaved, fully or partially glycosylated trimers. In addition, they found cleaved gp140 were not required for trimerization, however they played a significant role in triggering conformational changes in channelizing the trimers to generate compact three blade propeller shaped trimers. Verkerke et al. used lectin affinity chromatography to purify native like trimers from diverse HIV-1 isolates. The challenges faced in the production, analysis and synthesis of GP160 ENV trimer glycoprotein are reported (Grimm et al., 2015; Guenaga, 2015). Surface mutation, charge polarity and glycosylation and sequence variation between known variants in different clades are the significant barriers causing difficulty in imitating a native-like conformation of the glycoprotein. It is evident that assembling individual GP160 into a trimer spike complex structure is a challenge from a protein-protein interaction viewpoint. A large number of GP120 and GP41 structures are available in the PDB deposited using different biophysical techniques to understand the underlying molecular mechanism of the interacting proteins.

Sowmya et al. demonstrated the correlation between sequence polarity and mean Shannon entropy by calculating sequence polarity for surface residues in GP120 and GP41 and concluded stating the use of protein modification in the enhancement of HIV-1 vaccine across different clades, blood, and brain. Nilofer et al. characterized the interfaces of GP120-GP120, GP120-GP41 and GP41-GP41 and reported that the interfaces of GP120-GP120 are largely polar. The interfaces of GP120-GP41 and GP41-GP41 are characteristics of polar and non-polar residues. We characterize a manually curated dataset of 121 GP120 and 85 GP41 (Figure 1) protein interfaces reported by Nilofer et al. using interface features including interface area, interface size (number of residues at the interface), van der Waals, hydrogen bonds and electrostatics, to verify our previous finding stating that small protein interfaces are rich in electrostatics are often linked to regulatory proteins (Nilofer et al., 2020). The residues at the interface are displayed using CPK depiction (Discovery Studio® (Systèmes, 2020)).

MATERIALS AND METHODS

Dataset

We used a dataset of 206 interfaces manually curated as reported by Nilofer et al. It consists of 121 GP120 (Table 1) and 85 GP41 (Table 2) interfaces. It should be noted that GP120 structures in the PDB are available in ligand-bound state.

Interface area

Interface area was estimated for each of 121 interfaces of GP120 and 85 interfaces of GP41 using Naccess (Hubbard and Thornton, 1993). Naccess uses Lee and Richards method (Lee and Richards, 1971), wherein a probe with radius 1.4Å (Jones and Thornton, 1996) roll over the protein complex in monomer state and dimer state to find the accessible surface area and the interface area using delta ASA. Delta ASA (change in accessible surface area) is calculated using a formula: [ASA (Monomer subunit 1) + ASA (Monomer subunit 2) - AB (Dimer complex)]/2.

Interface size & interface energies

Interface size and interface energies were estimated for each of 121 interfaces of GP120 and GP41 using PPCheck (Sukhwal and Sowdhamini, 2015). PPCheck uses distance criteria to identify the non-covalent interactions between atoms of the two interacting proteins. It should be noted that the role of water is ignored in this analysis.

Large interface area and small interface area

Interfaces with large interface size (98±65 (GP120) and 73±65 (GP41)) and interface area (882±1166Å² (GP120) and 921±1288Å² (GP41)). Having dominant van der Waals energy (≥60%) at the interface are defined as large interface whereas interfaces with small interface size (58±20 (GP120) and 27±9 (GP41)) and interface area (581±1605Å² (GP120) and 483±896Å² (GP41)) having subdominant van der Waals energy (<60%) are defined as small interface.

Dominant and subdominant van der Waals interface

Interfaces with van der Waals contribution ≥60% to total energy (sum of van der Waals, hydrogen bonds and electrostatics) is defined as dominant interfaces, while interfaces with van der Waals contribution <60% to total energy is defined as subdominant interfaces. A cutoff of 60% was used as the larger part of van der Waals contribution was at...
Table 1: Dataset of GP120 (121) having large (van der Waals Dominant) and small (van der Waals Subdominant) interfaces are listed.

| Large interfaces van der waals dominant | GP120 Small interfaces van der waals subdominant |
|-----------------------------------------|-----------------------------------------------|
| 1G9M 2NY4 4R4N 1RZ7 3IDX 4J6R 4LSS 4R4H 4YBL 5IES | 1G9N 2NY5 4RFN 1RZ8 3LQA 4JB9 4LSU 4RQS 4YC2 5IF0 |
| 1GC1 2NY6 4RFO 1RZF 3NGB 4JDT 4LSV 4RWY 4YDI 5IGX | 1NAK 3JWD 4XNZ 1RZG 3Q6G 4J03 40LU 4RX4 4YDK 5KG9 |
| 1RZJ 3JWO 4YDJ 1RZI 3SE8 4JKP 40LV 4S1Q 4YDL 5T33 | 1RZK 3MLS 4ZTO 1YYL 3SE9 4JPV 40LW 451R 4YFL 5TE4 |
| 2F58 4DVR 5F96 1YMY 3TYG 4JPW 40LX 451S 5CAY 5TE6 | 2NYX 4J01 5KJR 2B4C 3U7Y 4JZW 40LY 4XMK 5F6j - |
| 2NXZ 4J02 5KZC 2I5Y 4JAN 4JZZ 40LZ 4XML 5F90 - | 2NY0 4K0A 5TE7 2I60 4D9L 4LAI 40M0 4XM 5F9W - |
| 2NY1 4KA2 - 2NY7 4H8W 4LS steep 40M1 4XY 5FCU - | 2NY2 4LST - 2QAD 4I3R 4LSQ 4PH 4XS 5FEC - |
| 2NY3 4R4F - 3F58 4I3S 4LSR 4R2G 4XVT 5I9Q - | 2NY8 4DVR 5F9E 1YYM 3TYG 4JPW 40LX 451S 5CAY 5TE6 |

Figure 1: Examples of GP120 (PDB 3MLZ) and GP41 (PDB 2ZFC) viral protein interfaces are illustrated along with the percent contribution of each of interface energies.

this cutoff on a scale of 0-100% and hence used.

**Statistical analysis**

We calculated interface energies of GP120 and GP41 using the statistical (Microsoft® Office Excel (version 2003)) variables including mean, mode, distribution, standard deviation and frequency at definite bin and range. We also carried out multiple linear regressions analysis for each interface with interface size against van der Waals, hydrogen bonds, electrostatic, total energy and interface area using regression tool. Its co-efficient of determination ($r^2$) was predicted with an evaluation of p-value using ANOVA (statistical test) at 95% confidence limit.

**RESULTS AND DISCUSSION**

The HIV-1 envelope trimer glycoprotein GP160 is a potential vaccine candidate for HIV-1/AIDS (Burton et al., 2004). Structural data of GP160 available in
Table 2: Dataset of GP41 (85) having large (van der Waals Dominant) and small (van der Waals Subdominant) interfaces are listed.

| GP41            | Large Interfaces van der Waals Dominant | Small Interfaces van der Waals Subdominant |
|-----------------|----------------------------------------|------------------------------------------|
| 1CE0            | 1U8N                                   | 1NLD                                     |
| 1TJG            | 1U8O                                   | 1T2G                                     |
| 1TJH            | 1U8P                                   | 2CMR                                     |
| 1TJI            | 1U8Q                                   | 2FX7                                     |
| 1U8H            | 1U91                                   | 2FX9                                     |
| 1U8I            | 1U92                                   | 2Q31                                     |
| 1U8J            | 1U93                                   | 2R5B                                     |
| 1U8K            | 1U95                                   | 2R5D                                     |
| 1U8L            | 2F5B                                   | 2X7R                                     |
| 1U8M            | 2FX8                                   | 3ECB                                     |

Therefore, we characterized the interfaces of GP120 and GP41 using interface area, interface size and interface energies using PPCheck (identifies non-covalent interactions using distance criteria). To verify our previous findings, we used the manually curated dataset of 121 and 85 interfaces of GP120 and GP41 proteins. The statistical analysis show that the mean interface size (98±65 (GP120) and 73±65 (GP41)) and interface area (882±1166Å² (GP120) and 921±1288Å² (GP41)) to be in close proximity for GP120 and GP41 interfaces (Figure 2). In contrast to our previous study we observed, most of the interfaces to have an interface area <1000Å² in both GP120 (60%) and GP41 (71%) and about 25% of GP120 and 19% of GP41 to have interface area between 1500Å² to 2000Å² (Figure 3).

Subsequently, we described each interface of GP120 and GP41 using van der Waals, H-bond, electrostatics and total energies along with their varying proportion of contribution at the interface. Thus, we calculated each individual contribution in percentage towards total energy. We observed interfaces to have high percentage of van der Waals (77%) and a low percentage of hydrogen bonds (12%) and electrostatics (11%) on average for GP120 and GP41 complexes (Figure 4). In addition, we noticed the interfaces of GP120 and GP41 to be normally distributed with increasing percentage of van der Waals (Figure 5). While a proportion of the interface
Figure 4: GP120 and GP41 interfaces are shown in terms of mean percent van der Waals, hydrogen bonds and electrostatics.

Figure 5: GP120 and GP41 interfaces are shown with varying percent contribution of van der Waals, hydrogen bonds and electrostatics.

decrease with increasing percentage of hydrogen bonds and electrostatics unlike van der Waals energy. It should be noted that interfaces of GP120 and GP41 are similar with the percentage contribution of van der Waals, hydrogen bonds, electrostatics, interface area and interface size. We further grouped the interfaces of GP120 and GP41 as dominant (≥60%) and subdominant (<60%) van der Waals based on its contribution towards total energy. As a result dominant interfaces have ≥60% of van der Waals with less magnitude of hydrogen bonds and electrostatics. We observed majority of interfaces of GP120 (88%) and GP41 (74%) to be van der Waals dominant with less than 10% contribution of hydrogen bonds and electrostatics while the remaining 12% of GP120 and 26% of GP41 have subdominant van der Waals with more than 15% of hydrogen bonds and electrostatic (Figure 6). We performed statistical analysis on dominant van der Waals and subdominant van der Waals interfaces to
Figure 6: GP120 and GP41 interfaces are shown with increasing percentages of van der Waals, hydrogen bonds and electrostatics.

Figure 7: Mean percentage of hydrogen bonds and electrostatics in large and small interfaces of GP120 and GP41 are shown.
highlight the contribution of hydrogen bonds and electrostatics in the small interfaces. We observed subdominant van der Waals interfaces of GP120 and GP41 to have three fold more of both hydrogen bonds and electrostatics when compared to dominant van der Waals interfaces (Figure 7). Furthermore, we noticed subdominant van der Waals interfaces to be more pronounced with more than 20% of hydrogen bonds and electrostatics distinct compared to dominant van der Waals interfaces (Figure 8). It is evident from (Figure 9) that the interface size and interface area of small interfaces are only half when compared to the large interfaces. Most of the interfaces with subdominant van der Waals have interface area less than 500Å² (Figure 10). Therefore, it was stated that the small interfaces with subdominant van der Waals energy and small interface area are rich in electrostatics. However, in context to the small interfaces (subdominant van der Waals) of GP120 and GP41, small interfaces of GP120 are rich in hydrogen bonds and GP41 is rich in electrostatics. Interface size increases with interface area is a known fact and hence we correlated interface size and interface energies. It was reported that total...
Figure 10: Interface area in large and small interfaces among GP120 and GP41 is shown.

Figure 11: Correlation between interface energy and interface size of GP120 and GP41 interfaces is shown.
Figure 12: Correlation between interface energy and interface size for GP120 and GP41 interfaces in terms of large and small interfaces is shown.

energy, van der Waals and hydrogen bonds increase with interface size but electrostatics decrease with increasing interface size (Nilofer et al., 2020). While the results of our current study shows that van der Waals and total energies of GP120 and GP41 interfaces increase with interface size but hydrogen bonds and electrostatics decrease with increasing interface size (Figure 11). Hence we divided our interfaces as large (dominant van der Waals) and small (subdominant van der Waals) interface based on their percent contribution towards total energy to check for correlation. It has also been reported that in small interfaces, total energy, van der Waals and hydrogen bonds decreases considerably with the increasing interface size whereas electrostatics moderately increases with interface size (Nilofer et al., 2020). But, this is not the case with the small interfaces of GP120 and GP41. Surprisingly, we found electrostatics ($r^2=0.63$) (Figure 12p) to be highly pronounced in GP41 interfaces with subdominant van der Waals having van der Waals ($r^2=0.23$) (Figure 12h) and without hydrogen bonds ($r^2=0$) (Figure 12p) contribution. Contrastingly, we observed the small interfaces of GP120 to be highly stabilized by hydrogen bonds ($r^2=0.59$) (Figure 12k) followed by electrostatics ($r^2=0.20$) (Figure 12o). Hence, we report that hydrogen bonds ($r^2=0.59$) (Figure 12k) increases with the interface size in the small interfaces of GP120 and electrostatics ($r^2=0.63$) (Figure 12p) increases with the interface size in the small interfaces of GP41.

**CONCLUSIONS**

GP120 viral proteins interact with GP41 to form GP160 the HIV-1 trimer glycoprotein. Statistical analysis on the interfaces of GP120 and GP41 using interface area, interface size and interface energies including van der Waals, hydrogen bonds and electrostatics demonstrate that they are similar. 88% of GP120 and 74% of GP41 interfaces have large interface area and interface size with dominant van der Waals energy; while 12% of GP120 and 26% of GP41 interfaces have small interface area and interface size with subdominant van der Waals energy. In addition, small interfaces were observed to have three fold more of hydrogen bonds and electrostatics than large interfaces. It is shown hydrogen bonds to increase with interface size in the small interfaces of GP120; while electrostatics to increase with interface size in small interfaces of GP41 in absence of hydrogen bonds. These insights from the interfaces of GP120 and GP41 shows that our previous finding stating that small interfaces with small interface area are rich in electrostatics holds true in case of GP41 but not in the case of GP120.

**Funding Support**

The authors declare that they have no funding support for this study.

**Conflict of Interest**

The authors declare that they have no conflict of interest for this study.
REFERENCES

Abagyan, R. A., Batalov, S. 1997. Do aligned sequences share the same fold. *Journal of Molecular Biology*, 273(1):355–368.

Adis Editorial 2003. HIV gp120 Vaccine-VaxGen: AIDSVAX™, AIDSVAX™ B/B, AIDSVAX™ B/E, HIV gp120 Vaccine-Genentech, HIV gp120 Vaccine AIDSVAX-VaxGen, HIV Vaccine AIDSVAX-VaxGen. *Drugs in R & D*, 4:249–253.

Alsalmi, W., et al. 2015. A New Approach to Produce HIV-1 Envelope Trimers. *Journal of Biological Chemistry*, 290(32):19780–19795.

Burton, D. R., et al. 2004. HIV vaccine design and the neutralizing antibody problem. *Nature Immunology*, 5(3):233–236.

Chen, B., et al. 2005. Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature*, 433(7028):834–841.

Doores, K. J. 2015. The HIV glycan shield as a target for broadly neutralizing antibodies. *FEBS Journal*, 282(24):4679–4691.

Grimm, S. K., et al. 2015. Directed Evolution of a Yeast-Displayed HIV-1 SOSIP gp140 Spike Protein toward Improved Expression and Affinity for Conformational Antibodies. *PLOS ONE*, 10(2).

Guenaga, J. 2015. Well-ordered trimeric hiv-1 subtype b and c soluble spike mimetics generated by negative selection display native-like properties. *PLoS Pathogens*, 11(1).

Hubbard, S. J., Thornton, J. M. 1993. NACCESS Computer Program. *Journal of Biophysical Chemistry*, 1(3).

Jones, S., Thornton, J. M. 1996. Principles of protein-protein interactions. *Proceedings of the National Academy of Sciences*, 93(1):13–20.

Kanguane, P., et al. 2008. Designing HIV gp120 peptide vaccines: Rhetoric or reality for Neuro-AIDS. *Spectrum of Neuro-AIDS Disorders*, pages 105–119.

Kinjo, A. R., et al. 2001. Physicochemical evaluation of protein folds predicted by threading. *European Biophysics Journal*, 30(1):1–10.

Lee, B., Richards, F. M. 1971. The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology*, 55(3):379–383.

Moore, J. P., et al. 1990. Enhancement of soluble CD4-mediated HIV neutralization and gp120 binding by CD4 autoantibodies and monoclonal antibodies. *AIDS Research and Human Retroviruses*, 6(11):1273–1279.

Moore, J. P., et al. 1992. Virions of primary human immunodeficiency virus type 1 isolates resistant to soluble CD4 (sCD4) neutralization differ in sCD4 binding and glycoprotein gp120 retention from sCD4-sensitive isolates. *Journal of Virology*, 66(1):235–243.

Nilofer, C., et al. 2017. HIV-1 envelope (ENV) GP160 trimer protein complex SPIKE as a recombinant macromolecular assembly vaccine component candidate: Current opinion. *Global Virology II - HIV and NeuroAIDS*, pages 939–951.

Nilofer, C., et al. 2020. Small protein-protein interfaces rich in electrostatic are often linked to regulatory function. *Journal of Biomolecular Structure and Dynamics*, 38(11):3260–3279.

Rerks-Ngarm, S., et al. 2009. Vaccination with ALVAC and AIDSVAX to Prevent HIV-1 Infection in Thailand. *New England Journal of Medicine*, 361(23):2209–2220.

Rerks-Ngarm, S., et al. 2013. Extended Evaluation of the Virologic, Immunologic, and Clinical Course of Volunteers Who Acquired HIV-1 Infection in a Phase III Vaccine Trial of ALVAC-HIV and AIDSVAX B/E. *Journal of Infectious Diseases*, 207(8):1195–1205.

Ringe, R. P., et al. 2015. Influences on the design and purification of soluble, recombinant native-like HIV-1 envelope glycoprotein trimers. *Journal of Virology*, 89(23):12189–12210.

Sanders, R. W., Moore, J. P. 2017. Native-like Env trimers as a platform for HIV-1 vaccine design. *Immunological Reviews*, 275(1):161–182.

Shin, S. Y. 2016. Recent update in HIV vaccine development. *Clinical and Experimental Vaccine Research*, 5(1):6.

Sowmya, G., et al. 2011. HIV-1 envelope accessible surface and polarity: clade, blood, and brain. *Bioinformation*, 6(2):48–56.

Sukhwal, A., Sowdhamini, R. 2015. PPCheck: A webserver for the quantitative analysis of protein interfaces and prediction of residue hotspots. *Bioinformatics and Biology Insights*, 9:141–151.

Systèmes, D. 2020. BIOVIA, discovery studio visualizer, release 2019. *San Diego: Dassault Systèmes*.

Überla, K. 2008. HIV vaccine development in the aftermath of the STEP study: Re-focus on occult hiv infection? *PLoS Pathogens*, 4(8):e1000114.

Verkerke, H. P., et al. 2016. Epitope independent purification of native like envelope trimers from diverse HIV-1 isolates. *Journal of Virology*, 90(20):9471–9482.