SciMedCentral

Journal of Drug Design and Research

Original Research

A Comparative Analysis of HIV-1 Envelope Glycoprotein GP120 Affinity for the Binding Site of Human CD4

Yan Long*, Haifeng Liang*, Zixin Long, Piaopiao Wang, Xueli Xie#, and Jiping Shao#

Key Laboratory of Tropical Translational Medicine of Ministry of Education, School of basic medical sciences, Hainan Academy of Medical Sciences, Hainan Medical University, China

**Yan Long and Haifeng Liang contributed equally to this work*

#Jiping Shao and Xueli Xie are co-corresponding authors

***Corresponding author**

Jiping Shao, Key Laboratory of Tropical Translational Medicine of Ministry of Education, School of basic medical sciences, Hainan Academy of Medical Sciences, Hainan Medical University, China, Tel No: 66961105

Submitted: 22 January 2026

Accepted: 09 February 2026

Published: 10 February 2026

ISSN: 2379-089X

Keywords

- HIV-1
- Envelope Glycoprotein
- Protein-Protein Interaction
- Docking

Abstract

Objective: To analyze the physicochemical properties of HIV-1 gp120, and investigate the interaction with the human CD4 receptor, providing structural insights for antibody and vaccine design.

Methods: The PDB files of gp120 and CD4 were obtained from the Protein Data Bank and AlphaFold Protein Structure Database. The Expasy-ProtParam tool was utilized to analyze the physicochemical properties of HIV-1 gp120. ClusPro 2.0 was employed to perform protein-protein docking between gp120 and CD4, and PyMOL was used to visualize binding sites and analyze interacting residues.

Results: Physicochemical analysis revealed that the gp120 core (molecular weight 39,168.53 Da) has a theoretical isoelectric point of 6.84 and an instability index of 39.09, classifying it as a stable protein. The docking models identified key amino acid residues involved in the gp120-CD4 interaction, including GLN-39, ASP-35, THR-28, HIS-18, CYS-31, and GLU-44, primarily driven by hydrogen bonds and electrostatic interactions.

Conclusion: These results demonstrate that the uncleaved gp120 core maintains a native-like stable structure with specific hydrophilic and charged interaction sites for CD4. This structural stability and the identified binding interfaces support the potential of gp120 as a candidate for antibody-mediated immune response targeting and vaccine development.

INTRODUCTION

Human immunodeficiency virus 1 (HIV-1) is a retrovirus that leads to acquired immunodeficiency syndrome (AIDS). The HIV-1 envelope glycoprotein (Env) is a trimer embedded in the viral envelope, consisting of two non-covalently bound subunits: the surface subunit gp120 and the transmembrane subunit gp41 [1]. Gp120 contains a CD4 binding site, several variable regions (V1-V5), and a conserved core region that forms an exposed ring structure responsible for recognizing host cell receptors. The interaction between gp120 and the primary receptor CD4 triggers a series of conformational changes [2,3]. These changes not only facilitate the subsequent binding of co-receptors (such as CCR5 or CXCR4) but also expose the fusion peptide of gp41, driving the fusion of viral and host cell membranes [4,5]. Therefore, analyzing the molecular affinity and interaction mechanism between HIV-1 gp120 and human CD4 is critical for understanding viral entry and guiding antibody or vaccine design.

Protein-protein interactions (PPIs) are fundamental to biological functions, including intercellular communication and gene expression [6-10]. However, experimentally determining these interaction sites is often complex and time-consuming. To address this, computational docking tools have been developed to simulate protein interactions. Various algorithms exist, such as ZDOCK (Fast Fourier Transform) [11], HADDOCK (data-driven) [12], and GRAMM-X [13]. In this study, we utilized ClusPro 2.0, which applies physical principles to optimize the conformational space and identify the most likely docking configurations [14]. Furthermore, advances in bioinformatics and databases, such as the Protein Data Bank [15-19], and AlphaFold, have significantly facilitated the structural analysis of biological macromolecules [17].

Despite extensive research on HIV-1, the specific affinity dynamics and structural stability of the gp120-CD4 complex in the context of emerging structural data require further elucidation. In this paper, we employed computational approaches to analyze the physicochemical properties of the HIV-1 gp120 core and simulated its docking with CD4 using ClusPro 2.0. We describe the structural basis, affinity, and binding kinetics, exploring how CD4 targets the envelope glycoprotein surface. These findings aim to provide a theoretical basis for novel therapeutic strategies.

MATERIALS AND METHODS

Gp120 and CD4 Sequences Retrieval

The HIV-1 gp120 core structure (clade A/E 93TH057) was retrieved from the AlphaFold Protein Structure Database. The crystal structure of human CD4 (PDB ID: 1CDH) was obtained from the Protein Data Bank [19].

Prediction of Physicochemical Properties

The ProtParam tool (ExPASy proteomics server) was used to predict the physicochemical properties of gp120 and CD4. Parameters analyzed included molecular composition, molecular weight, theoretical isoelectric point (pI), instability index (II), aliphatic index, and grand average of hydropathicity (GRAVY).

Sequence Alignment and Phylogenetic Analysis

To analyze evolutionary conservation, HIV-1 env amino acid sequences (accession numbers FM865453-FM865531) from Hainan Province were retrieved from EMBL. Sequence alignment was performed, and a Maximum Likelihood tree was constructed based on entropy values and specific k-mers to depict the evolutionary history.

Computational Modeling of gp120-CD4 Complex

ClusPro 2.0 was utilized to predict the protein-protein complex structures and determine binding affinity [14-22]. PyMOL was employed to visualize the docked complexes, analyze the polarity of the interface, and identify specific interacting amino acid residues involved in hydrogen bonding.

RESULTS

Physicochemical properties of gp120 and CD4

The gp120 sequence starts with the N-terminal Val (V) and has a total of 353 amino acids. The calculated molecular weight was 39168.53. The theoretical isoelectric point (PI) of the gp120 sequence is 6.84, that is, the pH of the protein without net charge. The total number of negatively charged residues (Asp + Glu) was 34, while the total number of positively charged residues (Arg + Lys) was 33. The total number of atoms in the protein molecule is 5438, the molecular formula is C1714H2696N478O527S23, the instability index is 39.09, the aliphatic index is 76.46, and the total average hydrophobicity (GRAVY) is -0.403 (Table 1). The CD4 sequence consists of 178 amino acids with a molecular weight of 19,700.44 Da and a theoretical isoelectric point (pI) of 8.88. It contains a total of 19 negatively charged residues (Asp + Glu) and 23 positively charged residues (Arg + Lys). The molecular formula of the protein is C865H1414N242O273S4, comprising a total of 2,798 atoms. Based on the calculated instability index (II) of 44.66, this protein is classified as unstable. Additionally, the aliphatic index is 89.72, suggesting a relatively high thermostability, while the grand average of hydropathicity (GRAVY) value is -0.480, indicating that the protein is overall hydrophilic in nature.

Gp120 structure

According to structural analysis sequence alignment of gp120, the variable region and Loop domain of the sequence are indicated in Figure 1. The overall outline of HIV-1 gp120 core from AlphaFold Protein Structure Database is heart-shaped, and its skeleton structure is shown in Figure 1. This gp120 core consists of 19 beta -strands, 6 alpha -helices and 10 defined loop fragments. Among the 19 β chains, 8 pairs of anti-balanced β chains maintain the structural stability of gp120. Among the 10 defined loop segments, 5 are known loop segments (V1 \ V2, V3, V4 and V5) on the gp120 sequence. Other Loop domains belong to highly variable loops (Figure 1).

The structural characteristics of the D1D2 fragment of CD4

CD4 is of great significance because of its involvement in human immunodeficiency virus (HIV) infection. The main structure interacting with HIV-1 gp120 core is amino-terminal two-domain (D1D2) fragments of human CD4. The structure includes 1469 non-hydrogen atoms from the protein molecule (residues 1-178) and 86 water molecules. D1D2 is mainly composed of β-sheet structure, which is bound by hydrogen bonds. The D1 of CD4 has nine β-strands folded in a structure similar to immunoglobulin variable domains, and the second D2 consists of seven β-strands making a distinctive structure that has characteristics of both immunoglobulin variable domains and immunoglobulin constant domain (Figure 2a). The mutational studies are consistent in identifying the side chains of Phe43 and Arg59 (Figure 2b) as key determinants of gpl20 binding, with the phenyl group being dominant.

Table 1: Physicochemical properties of gp120 core and CD4.

	Number of amino acids	Molecular weight	PI	Formula	Instability index	Aliphatic index	GRAVY
CD4	178	19700.44	8.88	$C_{865}H_{1414}N_{242}O_{273}S_4$	44.66	89.72	-0.480
Gp120	353	39168.53	6.84	$C_{1714}H_{2696}N_{478}O_{527}S_{23}$	39.09	76.46	-0.403

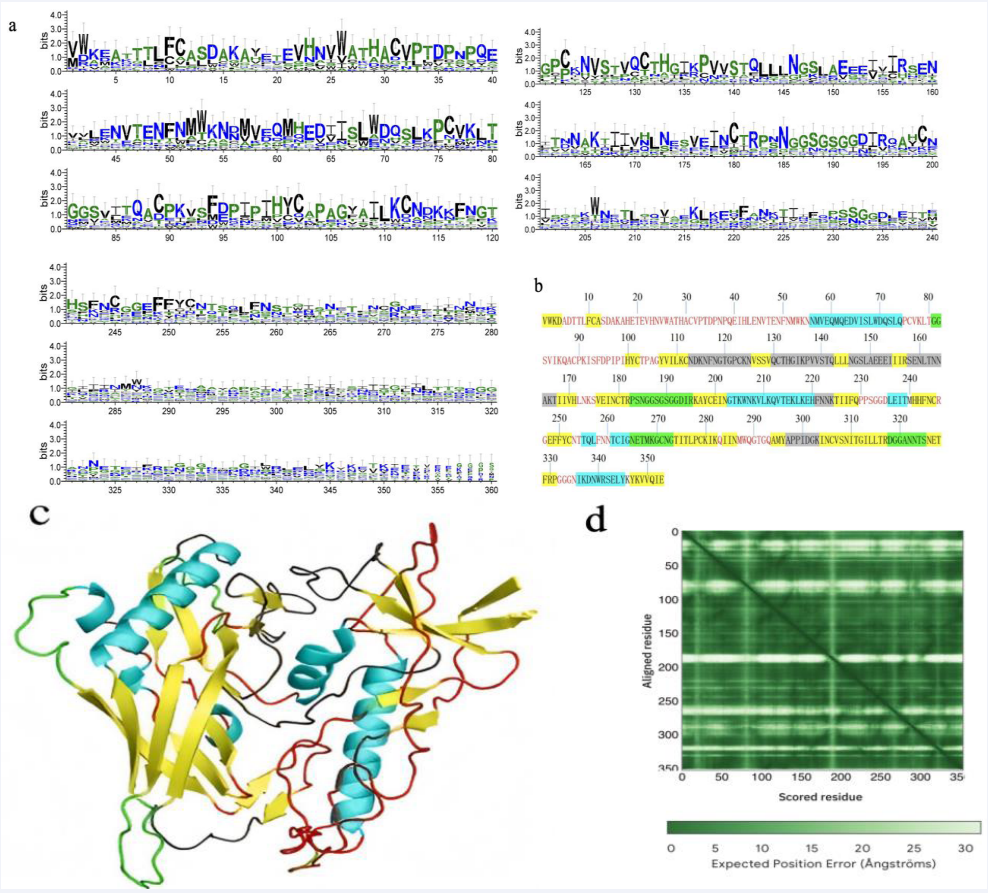

Figure 1 a. Sequence conservation of gp120 was analyzed. b. Sequence alignments. C. Cartoon of gp120. The yellow sequence represents the β chain, the blue sequence represents the α chain, the gray sequence represents the highly variable region, and the red font indicates that the sequence is conserved. d. AlphaFold confidence matrices (Predicted Aligned Error) for gp120, confirming model reliability. The asterisk (*) denotes the identical amino acid in all alignment sequences. The dash (-) denotes insertion and deletion.

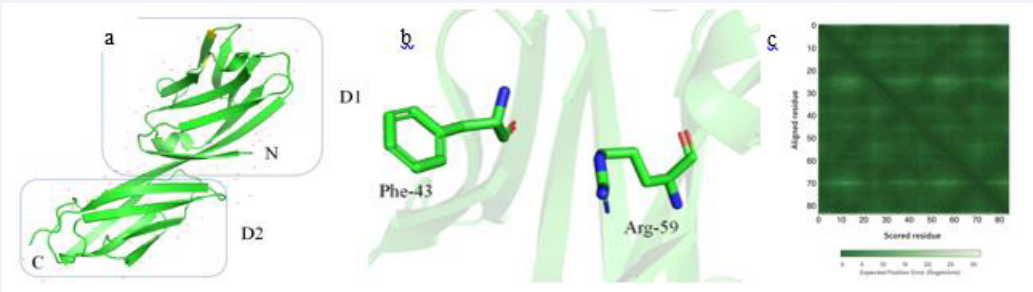

Figure 2 The structural characteristics of the D1D2 fragment of CD4. a. Cartoon D1D2 structure displayed by PyMoL. The red asterisk indicates water molecules. The yellow cartoon structure is Phe-43 and Arg-59 amino acid residues. b. Phe-43 and Arg-59 sticks. c. AlphaFold confidence matrices (Predicted Aligned Error) for CD4 confirming model reliability.

gp120-CD4 interaction

Through ClusPro 2.0: protein-protein docking, we obtained 10 display models. The polarity of 10 display models was analyzed using PyMoL. The interaction of amino acid residues between CD4 and gp120 was represented by hydrogen bonds. Eight display models are shown. The amino acid residues that interact with CD4 in these 10 display models include GLN-39, Ala-105, ASP-35, THR-28, HIS-18, CYS-31, THR-8, GLN-39, ASP-64, ALA-17, ALA-27, HIS-29, PRO-33, GLU-44 (Figure 3).

DISCUSSION

Technological advances, particularly the AlphaFold server, have revolutionized protein structure prediction, offering a viable alternative to purely experimental methods 17. In this study, we analyzed the HIV-1 gp120 core and its interaction with CD4 to understand the structural determinants of viral entry (Table 2).

Our physicochemical analysis indicates that the gp120 core is structurally stable (instability index < 40) and hydrophilic. This stability is a crucial feature for an envelope glycoprotein, as it must withstand environmental fluctuations before host engagement. The hydrophilic nature, indicated by the GRAVY score, is consistent with its role as a surface-exposed viral protein interacting with the aqueous extracellular environment and host receptors.

The docking results from ClusPro 2.0 highlight that the interaction interface is rich in both hydrophobic and charged amino acids, stabilized by hydrogen bonds. Specifically, residues such as Asp-35, Gln-39, and Arg-59 (on CD4) play pivotal roles. The engagement of CD4 induces conformational changes in gp120, which are necessary to expose the co-receptor binding site and subsequently activate gp41.

We observed that the interaction regions are concentrated near the N-terminal side of the gp120 amino acid sequence. This finding aligns with previous studies suggesting that the gp120 N-terminus is critical for the gp41-gp120 association 1. The stability of this interaction is vital; premature dissociation of gp120 would render the virus non-infectious, while delayed dissociation might hinder fusion. The specific residues identified in our models (e.g., those forming hydrogen bonds) provide precise targets for neutralizing antibodies. If an antibody can mimic the CD4 binding footprint or sterically hinder these specific residues (e.g., blocking GLN-39 or ASP-35), it could effectively neutralize the virus by preventing the initial attachment step.

In conclusion, the gp120 core maintains a stable, native-like structure capable of high-affinity binding to CD4 through a specific network of hydrogen bonds and electrostatic interactions. These structural insights reinforce the potential of the gp120 core as a stable immunogen for vaccine design and highlight specific interface residues as targets for therapeutic intervention.

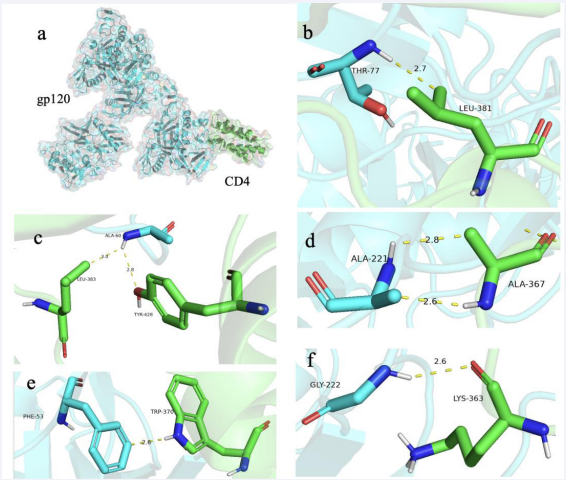

Figure 3 The docking models of gp120 and CD4. a Docking model of gp120 and CD16a. b Interaction between cyan THR-77 and green LEU-381, with a bond length of 2.7. c Interaction network involving three residues: cyan ALA-60 interacts with green LEU-383 (distance 3.0) and TYR-428 (distance 2.8). d Multiple interactions between cyan ALA-221 and green ALA-367, with distances of 2.8 and 2.6, respectively. e Interaction between cyan PHE-53 and green TRP-370 (distance 2.6), representing a typical hydrophobic/aromatic interaction. f Interaction between cyan GLY-222 and green LYS-363, with a distance of 2.6.

Table 2: Interaction between CD4 and GP120

NO.	GP120	CD4	Strength of hyrogrn bond
1	PHE-53	TRP-370	2.6
2	ALA-221	ALA-367	2.8
3	ALA-221	ALA-367	2.6
4	GLY-222	LYS-363	2.6
5	THR-77	LEU-381	2.7
6	ALA-60	LEU-383	3.0
7	ALA-60	TYR-428	2.8

AUTHOR CONTRIBUTIONS

JP conceived and designed the research; JP, LY, HF, ZX and XL wrote the manuscript; LY, HF, ZX and PP performed experiments; LY, HF, ZX and PP analyzed data. JP reviewed the manuscript.

ACKNOWLEDGEMENTS

This study was supported by research grants, Academic Enhancement Support Program of Hainan Medical University (Grant No: XSTS2025174)

REFERENCES

1. Julien JP, Cupo A, Sok D, Stanfield RL, Lyumkis D, Deller MC, et al. Crystal structure of a soluble cleaved HIV-1 envelope trimer. Science. 2013; 342: 1477-1483.
2. Lee JH, Ozorowski G, Ward AB. Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. Science. 2016; 351: 1043-1048.
3. Pancera M, Zhou T, Druz A, Georgiev IS, Soto C, Gorman J, et al. Structure and immune recognition of trimeric pre-fusion HIV-1 Env. Nature. 2014; 514: 455-461.
4. Wang H, Cohen AA, Galimidi RP, Gristick HB, Jensen GJ, Bjorkman PJ. Cryo-EM structure of a CD4-bound open HIV-1 envelope trimer reveals structural rearrangements of the gp120 V1V2 loop. Proc Natl Acad Sci U S A. 2016; 113: E7151-E7158.
5. Tamamis P, Floudas CA. Molecular recognition of CXCR4 by a dual tropic HIV-1 gp120 V3 loop. Biophys J. 2013; 105: 1502-1514.
6. Tamamis P, Floudas CA. Molecular recognition of CCR5 by an HIV-1 gp120 V3 loop. PLoS One. 2014; 9: e95767.
7. Zheng Y, Han GW, Abagyan R, Wu B, Stevens RC, Cherezov V, et al. Structure of CC Chemokine Receptor 5 with a Potent Chemokine Antagonist Reveals Mechanisms of Chemokine Recognition and Molecular Mimicry by HIV. Immunity. 2017; 46: 1005-1017.e5.
8. Qin L, Kufareva I, Holden LG, Wang C, Zheng Y, Zhao C, et al. Structural biology. Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine. Science. 2015; 347: 1117-1122.
9. Chen J, Kovacs JM, Peng H, Rits-Volloch S, Lu J, Park D, et al. HIV-1 ENVELOPE. Effect of the cytoplasmic domain on antigenic characteristics of HIV-1 envelope glycoprotein. Science. 2015; 349: 191-195.
10. Colin P, Bénureau Y, Staropoli I, Wang Y, Gonzalez N, Alcami J, et al. HIV-1 exploits CCR5 conformational heterogeneity to escape inhibition by chemokines. Proc Natl Acad Sci U S A. 2013; 110: 9475-9480.
11. Ozorowski G, Pallesen J, de Val N, Lyumkis D, Cottrell CA, Torres JL, et al. Open and closed structures reveal allostery and pliability in the HIV-1 envelope spike. Nature. 2017; 547: 360-363.
12. Manglik A, Kruse AC. Structural Basis for G Protein-Coupled Receptor Activation. Biochemistry. 2017; 56: 5628-5634.
13. Kwon YD, Pancera M, Acharya P, Georgiev IS, Crooks ET, Gorman J, et al. Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. Nat Struct Mol Biol. 2015; 22: 522-531.
14. Pan R, Qin Y, Banasik M, Lees W, Shepherd AJ, Cho MW, et al. Increased Epitope Complexity Correlated with Antibody Affinity Maturation and a Novel Binding Mode Revealed by Structures of Rabbit Antibodies against the Third Variable Loop (V3) of HIV-1 gp120. J Virol. 2018; 92: e01894-e01917.
15. Lyumkis D, Julien JP, de Val N, Cupo A, Potter CS, Klasse PJ, et al. Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. Science. 2013; 342: 1484-1490.
16. Brandenberg OF, Magnus C, Regoes RR, Trkola A. The HIV-1 Entry Process: A Stoichiometric View. Trends Microbiol. 2015; 23: 763-774.
17. Jiang X, Feyertag F, Meehan CJ, McCormack GP, Travers SA, Craig C, et al. Characterizing the Diverse Mutational Pathways Associated with R5-Tropic Maraviroc Resistance: HIV-1 That Uses the Drug-Bound CCR5 Coreceptor. J Virol. 2015; 89: 11457-11472.
18. Cai Y, Karaca-Griffin S, Chen J, Tian S, Fredette N, Linton CE, et al. Antigenicity-defined conformations of an extremely neutralization-resistant HIV-1 envelope spike. Proc Natl Acad Sci U S A. 2017; 114: 4477-4482.
19. Ru H, Chambers MG, Fu TM, Tong AB, Liao M, Wu H. Molecular Mechanism of V(D)J Recombination from Synaptic RAG1-RAG2 Complex Structures. Cell. 2015; 163: 1138-1152.
20. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods. 2017; 14: 331-332.
21. Rohou A, Grigorieff N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J Struct Biol. 2015; 192: 216-221.
22. Jin J, Momboisse F, Boncompain G, Koensgen F, Zhou Z, Cordeiro N, et al. CCR5 adopts three homodimeric conformations that control cell surface delivery. Sci Signal. 2018; 11: eaal2869.