Published in final edited form as:

Curr Opin Struct Biol. 2013 April ; 23(2): 268–276. doi:10.1016/j.sbi.2013.03.007.

HIV-1 envelope glycoprotein structure

Alan Merk and Sriram Subramaniam

Laboratory of Cell Biology, Center for Cancer Research, National Cancer Institute, NIH Bethesda, MD 20892

Abstract

The trimeric envelope glycoprotein of HIV-1, composed of gp120 and gp41 subunits, remains a major target for vaccine development. The structures of the core regions of monomeric gp120 and gp41 have been determined previously by X-ray crystallography. New insights into the structure of trimeric HIV-1 envelope glycoproteins are now coming from cryo-electron tomographic studies of the gp120/gp41 trimer as displayed on intact viruses and from cryo-electron microscopic studies of purified, soluble versions of the ectodomain of the trimer. Here, we review recent developments in these fields as they relate to our understanding of the structure and function of HIV-1 envelope glycoproteins.

Introduction

Infection of target cells by the human immunodeficiency virus (HIV), a particle with dimensions of ~ 100 nm to 150 nm, follows the interaction of its surface envelope glycoprotein (Env) with the cellular CD4 receptor and co-receptors such as CCR5 and CXCR4 [1–8]. As displayed on the surface of the viral membrane, Env is a trimer, composed of three copies of non-covalently associated heterodimers of gp120, the component that interacts with cellular receptors, and gp41, the transmembrane component necessary for mediating fusion between viral and target membranes. Trimeric Env, like HIV itself, is heterogeneous in almost every possible respect: in addition to constant mutations that alter the genetic composition of the virus in infected hosts and the variable number of Env displayed on the membrane surface, each circulating virus can be studded with assorted host membrane proteins, differently sized, with Env that is differentially glycosylated and conformationally flexible. This extensive repertoire of variability, maintained in conjunction with selected conserved structural features that enable targeting and infection of host cells, is central to how HIV successfully circumvents capture by the immune system. In order to define Env function in detail, we therefore need three-dimensional structures of trimeric Env at the highest possible resolution in a spectrum of different conformations and a complete understanding of the molecular basis of its extraordinary heterogeneity. In this review, we survey selected recent developments in the structural biology of Env, and highlight how X-ray crystallography, cryo-electron microscopy and cryo-electron tomography are all coming together to assemble a composite view of Env structure as it relates to neutralization mechanisms and viral entry.

Address for correspondence: ssl@nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Structural studies of gp120 by X-ray crystallography

The first insights into the structure of HIV-1 gp120 were reported in a seminal paper by Kwong et al. [9], which presented the crystal structure of a monomeric gp120 core bound to a 2-domain soluble CD4 construct and a Fab fragment of the monoclonal antibody 17b. Although the gp120 used for crystallization was that of the deglycosylated truncated core, representing ~ 60% of the polypeptide, this structure unveiled its overall architecture, and the molecular interfaces involved in binding to CD4 and the co-receptor mimic, 17b. The knowledge gained from this structure was utilized to improve the potency and breadth of small-molecule mimics of CD4 [10,11] and to design a probe for the isolation of several broadly neutralizing antibodies, such as VRC01[12]. Since this original report, over two dozen crystal structures of monomeric HIV-1 gp120 cores in the unliganded state, or in complex with various antibodies and/or reagents that target the CD4 binding site have been published [11,13–24]. These have included structures of gp120 complexes with neutralizing and non-neutralizing antibodies, as well as ligands that act as either agonists or antagonists for potentiating HIV entry. The overall conformation of gp120 in all of these structures is virtually identical (Figure 1), with a clear separation between “inner” and “outer” domains that refer to portions of gp120 that are respectively closer to the inner and outer regions of the trimeric spike. The structural origins of how various antibodies and ligands affect HIV function must therefore lie either in minute differences at the respective molecular interfaces involved in binding, in quaternary conformational changes in the Env trimer, and/or in the more mobile regions of the protein not easily accessible to crystallographic methods [25].

The variations that do exist between the reported gp120 structures are restricted largely to the variable loop regions V1-V5 (Figures 1B, 1D), which are also the most disordered portions of the overall structure. Even different copies of gp120 present within a single three-dimensional crystal are not necessarily identical in loop conformation (Figure 1F). The V1V2 loop, expressed in the context of a non-HIV scaffold or as a 19-mer peptide, has been crystallized in complex with three Fabs, with three disparate conformations for V1V2 in each of these complexes [26,27]. The structure of the V3 loop is well-defined in some of the ternary complexes formed with soluble CD4 and selected Fab fragments [14], showing that it forms an extended structure whose tip protrudes ~ 30 Å from the main portion of gp120. However, multiple conformations are observed in crystal structures of V3 loop peptides, depending on which antibody is bound [28], suggesting that the structure of this loop may vary depending on the nature of the interactions at the gp120 surface. The V4-V5 regions are highly disordered in most crystal structures, as expected both from their location at the molecular periphery and the variability of the sequence in this region. In summary, these recent studies inform us that the V1-V5 loop residues don't always adopt the same conformation when analyzed in contexts that are removed from that of native Env, but knowledge of the actual conformation of these loops in native Env trimers remains elusive. We also do not know yet how closely the structures of the truncated gp120 cores themselves correspond to the conformation of native gp120 in trimeric Env.

The only report of a crystallographic structure of gp120 that is significantly different from all of the essentially identical HIV-1 gp120 structures shown in Figure 1 is from the simian immunodeficiency virus (SIV) envelope glycoprotein [29], crystallized without any bound ligand. The inner domain in this SIV gp120 construct adopts a different conformation from that seen for HIV-1 gp120, and led to derivation of a molecular model for trimeric Env in which the V1V2 loops were suggested to be located at the base of the spike. However, subsequent tomographic studies of intact Env trimers (reviewed in the next section) established that the V1V2 loops are, in fact, located at the apex of the spike, and also suggested that this structure of the unliganded SIV gp120 protomer does not match the

density maps obtained for native HIV-1 and SIV Env trimers using cryo-electron tomographic analyses.

Cryo-electron tomography of native trimeric Env

Cryo-electron tomography has emerged as a powerful tool to bridge the gap between structural and cell biology by providing intermediate resolution maps of complex and heterogeneous molecular assemblies that cannot be analyzed by crystallography [30–32]. Tomographic methods allow three-dimensional reconstruction of the structures of entire HIV-1 virions trapped in a near-native state. By employing computational methods that combine the information from sub-volumes corresponding to individual Env spikes, density maps of the Env spike can be obtained at resolutions of ~ 20 Å. Further, tomographic analyses of viruses that have been incubated with various antibodies and/or ligands can provide structures of native, fully glycosylated, functional Env trimers *in situ*, captured in an array of conformational states.

By combining cryo-electron tomography with image averaging and fitting of X-ray structures of selected gp120 conjugates, the quaternary conformations of trimeric Env complexes displayed on intact viruses or as soluble ectodomains have been determined, and early controversies about whether trimeric Env form a single stalk or a separated tripod have been resolved [33–39]. These combined approaches have led to definitive identification of the locations of the variable loop regions on intact Env trimers [38,40]. When trimeric Env is in the unliganded state, or when it is bound to antibodies such as VRC01, it is in a “closed” conformation, with the V1V2 loops located close to the apex of the spike (Figure 2). When trimeric gp120 is bound to CD4, or “CD4i” proteins such as the monoclonal antibody 17b or small domain antibody m36, it transitions from the “closed” state to an “open” state, in which the three gp120 monomers undergo a large rearrangement involving rotations of each gp120 monomer, relocating the V1V2 loops to the periphery of the trimer. The fact that co-receptor mimics such as 17b and m36 can bind and generate the same types of conformational changes observed with CD4 binding is an important mechanistic discovery about the mode of binding of these molecules, providing a structural context to understand why some HIV-1 isolates can enter cells that lack cell-surface CD4 [41,42]. When trimeric Env is bound to the monoclonal antibody b12 or small antibody derivative A12, there is a partial opening of the trimer, with a slight rearrangement of each gp120 monomer. It is conceivable that these different conformational states of Env are in dynamic equilibrium; alternatively, it is possible that the transitions to the partially and fully open states are irreversible, triggered changes, driven by the binding of the respective ligands, representing conformational intermediates generated in the sequence of events that lead ultimately to fusion of viral and plasma cell membranes. Our understanding of these changes can be expected to improve as more structural information is obtained on intact Env trimers using crystallographic, cryo-electron tomographic and other biophysical/biochemical approaches.

X-ray crystallography and cryo-electron tomography each provide unique information necessary to understand molecular mechanisms of Env function. For example, X-ray crystallographic studies of Env bound to either CD4 (which initiates infection) or to the broadly neutralizing antibody VRC01 (which blocks infection) are remarkably similar, with almost identical molecular interfaces involved in the contact of gp120 with these two ligands (Figures 3A, 3B). However, when the structures of VRC01 or CD4 bound to Env on intact HIV-1 virions are determined using cryo-electron tomography, there are dramatic differences in quaternary structure which help explain an important mechanistic aspect of the broad neutralization by VRC01 [36]. Thus, VRC01 locks Env in a closed conformation in which gp41 and its fusogenic components are buried at the core (Figure 3C), while CD4

binding induces an open conformation (Figure 3D), which enables exposure of gp41 and the initiation of subsequent steps in the entry process. The profound differences observed between the outcomes of CD4 or VRC01 binding, despite the fact that they bind similar sites on monomeric gp120, underscore the importance of methods such as cryo-electron tomography for determining quaternary structures of native Env. These differences are also mirrored by biochemical studies that show that Env trimers expressed on cell surfaces bound to VRC01 do not display the extensive conformational changes that are observed when CD4 binds [43], consistent with the findings from cryo-electron tomography.

Structural studies of gp41 by X-ray crystallography and cryo-electron microscopy

While the gp120 subunit of HIV Env is responsible for virion attachment to target cells, the gp41 region mediates fusion between viral and target cell membranes. The first insights into the structure of gp41 were obtained in 1997, when multiple groups reported the crystal structure of the post-fusion conformation of the gp41 core. In this conformation, the C-terminal helices of gp41 pack around the N-terminal helices in an antiparallel fashion, forming a structure called the six-helix bundle (Figure 4A, 4B) [44–46]. Although the fusion inhibitor enfuvirtide (T-20) was discovered several years before the first publication of the six-helix bundle structure [47,48], these structures helped explain the mechanism of action by T-20 and have led to the design of more potent variants of T-20 [49–51]. All reported gp41 core structures determined by X-ray crystallography or by NMR spectroscopy [52] (of the related simian immunodeficiency virus protein) have exhibited conformations that are nearly identical to those of the original gp41 structures (see superposition of 25 HIV-1 gp41 structures in Figures 4A, 4B), despite the fact that the gp41 core has been crystallized using diverse media [53], bound to different ligands (both neutralizing and non-neutralizing) [54–57], and with various mutations [58–64].

Knowledge of the structure of gp41 at high resolution in any pre-fusion step remains a critical gap in the structural biology of Env, but some progress is being made. A 9 Å resolution structure of a pre-fusion intermediate in the HIV entry process was recently determined by single particle cryo-electron microscopy [36]. This structure provides unexpected insights into the structural rearrangements that occur before formation of the six-helix bundle. When complexed to the co-receptor mimic 17b, trimeric Env, in both soluble and native membrane-bound forms, undergoes structural changes that lead to formation of an open, activated conformation. In this intermediate, each gp120 protomer is rotated outwards, and gp41 is exposed, as visualized by three clearly resolved densities at the center of the spike. Tran et al [36] assigned these densities to the three N-terminal rather than C-terminal gp41 helices, and noted that the helices in this novel, activated Env conformation are held apart by their interactions with the rest of Env, and are less compactly packed than in the post-fusion, six-helix bundle state. This likely represents a conformational intermediate that captures gp41 at a key vulnerable step in the entry process. It will be interesting to explore how changes in gp41 conformation after binding of a co-receptor mimic compare with changes that occur with CD4 binding or with true co-receptor binding. Further, the gp41 sequence in HIV has many organizational similarities to similar components in the fusion machinery employed by viruses such as Ebola and influenza [45,65]. In this regard, we note that broadly neutralizing antibodies against the membrane proximal external region (MPER) of HIV Env, similar mechanistically to the anti-stem antibodies seen in influenza, have been isolated and structurally characterized using various methods [66–69]. However, while there is considerable structural information from X-ray crystallography and cryo-electron tomography on the complexes formed between trimeric hemagglutinin and various anti-stem antibodies [70–74], there is no structural information available yet for similar complexes of MPER antibodies with trimeric Env. Structures are

available, however, for isolated gp41 peptides in complex with several MPER antibodies such as 2F5, 4E10, Z13e1 and 10E8. Interestingly, the conformations of the peptides are different depending on the bound antibody, likely reflecting the structural plasticity of the MPER region when it is removed from the native context of intact Env. Additionally, there is no consensus on precisely how the MPER antibodies bind their cognate epitopes, as some have found that the antibodies extract their epitopes from the membrane [68] while others have found that these antibodies bind Env only after CD4 binding [75]. High-resolution structural information of the MPER region in the context of native trimeric Env could allow interpretation of these results and make it easier to assess the similarities and differences in the entry mechanisms employed by HIV and related viruses (reviewed in [76–78]).

Conclusions and future perspective

Understanding both the structure and the structural variation of native trimeric Env, on the surface of infectious virions, and at the highest possible resolutions, remains one of the most important challenges for Env structural biology. This is also of central interest for the rational design of a vaccine to protect against HIV/AIDS. Recent advances in Env structural biology illustrate that the integration of complementary information from X-ray crystallography, NMR spectroscopy, cryo-electron microscopy and tomography can be a powerful tool to understand Env structure and viral entry mechanisms. However, the limitations inherent to each of these technologies are also evident. While X-ray crystallographic studies have provided important information on the core regions of monomeric gp120 and on the stable, post-fusion, six-helix bundle formed by the N and C-terminal helical regions of the gp41 ectodomain, it appears unlikely that they will provide reliable atomic structures of the flexible or conformationally variable components of Env. Because X-ray crystallography requires ordered 3-dimensional crystals to obtain structural information, its use is restricted to a small number of simplified model proteins, precluding atomic level description of the enormous variation in structures of Env in different HIV-1 strains. The size of trimeric Env (~400 kD) also places it essentially outside the range that is currently accessible to NMR spectroscopy. In principle, cryo-electron microscopy and cryo-electron tomography have the potential to describe the structural heterogeneity and conformational states of native Env on intact virions as demonstrated in recent studies with HIV and influenza [74,79,80]. However, these methods have not yet proven to be useful to obtain structural information at resolutions in the 3 – 4 Å range that will be required for a proper understanding of the chemistry involved, although the demonstration that structures of small protein complexes can be reconstructed at resolutions of ~ 8 Å starting from cryo-electron tomographic data is a promising start [81].

Given the limitations of each of these approaches, the challenge for the future is for structural biologists in the HIV field to aim higher. Thus, we need crystallographic structures of a variety of physiologically relevant, full-length, native Env trimers. We need strategies to use NMR spectroscopy to capture dynamics of the entire Env trimer. We need cryo-electron microscopy and tomography to break out of present-day resolution barriers so that the quaternary structures of native Env are not restricted to low resolution envelopes, but are good enough for *de novo*, and unambiguous determination of the secondary structure of native Env on intact, infectious viruses. Computational modeling analyses need to find ways to integrate all of this experimental information into meaningful structural models. Achieving these goals will likely require significant technical advances in each of these disciplines, but they are not beyond reach, and work along these lines will continue to be inspired by the quest of applying tools in modern structural biology to design a safe and effective vaccine against HIV/AIDS.

Acknowledgments

This work was supported by funds from the intramural program of the National Cancer Institute, NIH, Bethesda, MD. We thank Tom Goddard and Elaine Meng for advice with the use of UCSF Chimera for figure preparation, and Jacqueline Milne, Erin Tran, Lesley Earl, Joel Meyerson and Alberto Bartesaghi for helpful comments on the manuscript.

References

1. Choe H, Farzan M, Sun Y, Sullivan N, Rollins B, Ponath PD, Wu L, Mackay CR, LaRosa G, Newman W, et al. The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell*. 1996; 85:1135–1148. [PubMed: 8674119]
2. McDougal JS, Kennedy MS, Sligh JM, Cort SP, Mawle A, Nicholson JK. Binding of HTLV-III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule. *Science*. 1986; 231:382–385. [PubMed: 3001934]
3. Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*. 1996; 272:872–877. [PubMed: 8629022]
4. Maddon PJ, Dalgleish AG, McDougal JS, Clapham PR, Weiss RA, Axel R. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell*. 1986; 47:333–348. [PubMed: 3094962]
5. Alkhatib G, Combadiere C, Broder CC, Feng Y, Kennedy PE, Murphy PM, Berger EA. CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science*. 1996; 272:1955–1958. [PubMed: 8658171]
6. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhart M, Di Marzio P, Marmon S, Sutton RE, Hill CM, et al. Identification of a major co-receptor for primary isolates of HIV-1. *Nature*. 1996; 381:661–666. [PubMed: 8649511]
7. Doranz BJ, Rucker J, Yi Y, Smyth RJ, Samson M, Peiper SC, Parmentier M, Collman RG, Doms RW. A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors. *Cell*. 1996; 85:1149–1158. [PubMed: 8674120]
8. Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, Cayan C, Maddon PJ, Koup RA, Moore JP, et al. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature*. 1996; 381:667–673. [PubMed: 8649512]
9. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*. 1998; 393:648–659. [PubMed: 9641677]
10. Madani N, Schon A, Princiotta AM, Lalonde JM, Courter JR, Soeta T, Ng D, Wang L, Brower ET, Xiang SH, et al. Small-molecule CD4 mimics interact with a highly conserved pocket on HIV-1 gp120. *Structure*. 2008; 16:1689–1701. [PubMed: 19000821]
11. LaLonde JM, Kwon YD, Jones DM, Sun AW, Courter JR, Soeta T, Kobayashi T, Princiotta AM, Wu X, Schon A, et al. Structure-based design, synthesis, and characterization of dual hotspot small-molecule HIV-1 entry inhibitors. *J Med Chem*. 2012; 55:4382–4396. [PubMed: 22497421]
12. Wu X, Yang ZY, Li Y, Hogerkorp CM, Schief WR, Seaman MS, Zhou T, Schmidt SD, Wu L, Xu L, et al. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science*. 2010; 329:856–861. [PubMed: 20616233]
13. Kwong PD, Wyatt R, Majeed S, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates. *Structure*. 2000; 8:1329–1339. [PubMed: 11188697]
14. Huang CC, Tang M, Zhang MY, Majeed S, Montabana E, Stanfield RL, Dimitrov DS, Korber B, Sodroski J, Wilson IA, et al. Structure of a V3-containing HIV-1 gp120 core. *Science*. 2005; 310:1025–1028. [PubMed: 16284180]
15. Huang CC, Lam SN, Acharya P, Tang M, Xiang SH, Hussan SS, Stanfield RL, Robinson J, Sodroski J, Wilson IA, et al. Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. *Science*. 2007; 317:1930–1934. [PubMed: 17901336]

16. Zhou T, Xu L, Dey B, Hessel AJ, Van Ryk D, Xiang SH, Yang X, Zhang MY, Zwick MB, Arthos J, et al. Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature*. 2007; 445:732–737. [PubMed: 17301785]
17. Pancera M, Majeed S, Ban YE, Chen L, Huang CC, Kong L, Kwon YD, Stuckey J, Zhou T, Robinson JE, et al. Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. *Proc Natl Acad Sci U S A*. 2010; 107:1166–1171. [PubMed: 20080564]
18. Chen L, Kwon YD, Zhou T, Wu X, O'Dell S, Cavacini L, Hessel AJ, Pancera M, Tang M, Xu L, et al. Structural basis of immune evasion at the site of CD4 attachment on HIV-1 gp120. *Science*. 2009; 326:1123–1127. [PubMed: 19965434]
19. Diskin R, Marcovecchio PM, Bjorkman PJ. Structure of a clade C HIV-1 gp120 bound to CD4 and CD4-induced antibody reveals anti-CD4 polyreactivity. *Nat Struct Mol Biol*. 2010; 17:608–613. [PubMed: 20357769]
20. Zhou T, Georgiev I, Wu X, Yang ZY, Dai K, Finzi A, Kwon YD, Scheid JF, Shi W, Xu L, et al. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science*. 2010; 329:811–817. [PubMed: 20616231]
21. Wu X, Zhou T, Zhu J, Zhang B, Georgiev I, Wang C, Chen X, Longo NS, Louder M, McKee K, et al. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science*. 2011; 333:1593–1602. [PubMed: 21835983]
22. Pejchal R, Doores KJ, Walker LM, Khayat R, Huang PS, Wang SK, Stanfield RL, Julien JP, Ramos A, Crispin M, et al. A Potent and Broad Neutralizing Antibody Recognizes and Penetrates the HIV Glycan Shield. *Science*. 2011
23. Diskin R, Scheid JF, Marcovecchio PM, West AP Jr, Klein F, Gao H, Gnanapragasam PN, Abadir A, Seaman MS, Nussenzweig MC, et al. Increasing the potency and breadth of an HIV antibody by using structure-based rational design. *Science*. 2011; 334:1289–1293. [PubMed: 22033520]
24. Kwon YD, Finzi A, Wu X, Dogo-Isonagie C, Lee LK, Moore LR, Schmidt SD, Stuckey J, Yang Y, Zhou T, et al. Unliganded HIV-1 gp120 core structures assume the CD4-bound conformation with regulation by quaternary interactions and variable loops. *Proc Natl Acad Sci U S A*. 2012; 109:5663–5668. [PubMed: 22451932]
25. Guttman M, Kahn M, Garcia NK, Hu SL, Lee KK. Solution structure, conformational dynamics, and CD4-induced activation in full-length, glycosylated, monomeric HIV gp120. *J Virol*. 2012; 86:8750–8764. [PubMed: 22674993]
26. McLellan JS, Pancera M, Carrico C, Gorman J, Julien JP, Khayat R, Louder R, Pejchal R, Sastry M, Dai K, et al. Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature*. 2011; 480:336–343. [PubMed: 22113616]
27. Liao HX, Bonsignori M, Alam SM, McLellan JS, Tomaras GD, Moody MA, Kozink DM, Hwang KK, Chen X, Tsao CY, et al. Vaccine Induction of Antibodies against a Structurally Heterogeneous Site of Immune Pressure within HIV-1 Envelope Protein Variable Regions 1 and 2. *Immunity*. 2013
28. Stanfield R, Cabezas E, Satterthwait A, Stura E, Profy A, Wilson I. Dual conformations for the HIV-1 gp120 V3 loop in complexes with different neutralizing fabs. *Structure*. 1999; 7:131–142. [PubMed: 10368281]
29. Chen B, Vogan EM, Gong H, Skehel JJ, Wiley DC, Harrison SC. Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature*. 2005; 433:834–841. [PubMed: 15729334]
30. Milne JL, Subramaniam S. Cryo-electron tomography of bacteria: progress, challenges and future prospects. *Nat Rev Microbiol*. 2009; 7:666–675. [PubMed: 19668224]
31. Meyerson JR, White TA, Bliss D, Moran A, Bartesaghi A, Borgnia MJ, de la Cruz MJ, Schauder D, Hartnell LM, Nandwani R, et al. Determination of molecular structures of HIV envelope glycoproteins using cryo-electron tomography and automated sub-tomogram averaging. *J Vis Exp*. 2011
32. Milne JL, Borgnia MJ, Bartesaghi A, Tran EE, Earl LA, Schauder DM, Lengyel J, Pierson J, Patwardhan A, Subramaniam S. Cryo-electron microscopy - a primer for the non-microscopist. *FEBS J*. 2013; 280:28–45. [PubMed: 23181775]

33. Meyerson JR, Tran EE, Kuybeda O, Chen W, Dimitrov DS, Gorlani A, Verrips T, Lifson JD, Subramaniam S. Molecular structures of trimeric HIV-1 Env in complex with small antibody derivatives. *Proc Natl Acad Sci U S A*. 2013; 110:513–518. [PubMed: 23267106]
34. Harris A, Borgnia MJ, Shi D, Bartesaghi A, He H, Pejchal R, Kang YK, Depetris R, Marozsan AJ, Sanders RW, et al. Trimeric HIV-1 glycoprotein gp140 immunogens and native HIV-1 envelope glycoproteins display the same closed and open quaternary molecular architectures. *Proc Natl Acad Sci U S A*. 2011; 108:11440–11445. [PubMed: 21709254]
35. Bartesaghi A, Sprechmann P, Liu J, Randall G, Sapiro G, Subramaniam S. Classification and 3D averaging with missing wedge correction in biological electron tomography. *J Struct Biol*. 2008; 162:436–450. [PubMed: 18440828]
36. Tran EE, Borgnia MJ, Kuybeda O, Schauder DM, Bartesaghi A, Frank GA, Sapiro G, Milne JL, Subramaniam S. Structural mechanism of trimeric HIV-1 envelope glycoprotein activation. *PLoS Pathog*. 2012; 8:e1002797. [PubMed: 22807678]
37. White TA, Bartesaghi A, Borgnia MJ, de la Cruz MJ, Nandwani R, Hoxie JA, Bess JW, Lifson JD, Milne JL, Subramaniam S. 3D structures of sCD4-bound states of trimeric SIV envelope glycoproteins determined using cryo-electron tomography. *J Virol*. 2011
38. White TA, Bartesaghi A, Borgnia MJ, Meyerson JR, de la Cruz MJ, Bess JW, Nandwani R, Hoxie JA, Lifson JD, Milne JL, et al. Molecular architectures of trimeric SIV and HIV-1 envelope glycoproteins on intact viruses: strain-dependent variation in quaternary structure. *PLoS Pathog*. 2010; 6:e1001249. [PubMed: 21203482]
39. Moscoso CG, Sun Y, Poon S, Xing L, Kan E, Martin L, Green D, Lin F, Vahlne AG, Barnett S, et al. Quaternary structures of HIV Env immunogen exhibit conformational vicissitudes and interface diminution elicited by ligand binding. *Proc Natl Acad Sci U S A*. 2011; 108:6091–6096. [PubMed: 21444771]
40. Hu G, Liu J, Taylor KA, Roux KH. Structural comparison of HIV-1 envelope spikes with and without the V1/V2 loop. *J Virol*. 2011; 85:2741–2750. [PubMed: 21191026]
41. LaBranche CC, Hoffman TL, Romano J, Haggarty BS, Edwards TG, Matthews TJ, Doms RW, Hoxie JA. Determinants of CD4 independence for a human immunodeficiency virus type 1 variant map outside regions required for coreceptor specificity. *J Virol*. 1999; 73:10310–10319. [PubMed: 10559349]
42. Dumonceaux J, Nisole S, Chaneil C, Quivet L, Amara A, Baleux F, Briand P, Hazan U. Spontaneous mutations in the env gene of the human immunodeficiency virus type 1 NDK isolate are associated with a CD4-independent entry phenotype. *J Virol*. 1998; 72:512–519. [PubMed: 9420253]
43. Li Y, O'Dell S, Walker LM, Wu X, Guenaga J, Feng Y, Schmidt SD, McKee K, Louder MK, Ledgerwood JE, et al. Mechanism of neutralization by the broadly neutralizing HIV-1 monoclonal antibody VRC01. *J Virol*. 2011; 85:8954–8967. [PubMed: 21715490]
44. Chan DC, Fass D, Berger JM, Kim PS. Core structure of gp41 from the HIV envelope glycoprotein. *Cell*. 1997; 89:263–273. [PubMed: 9108481]
45. Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, Wiley DC. Atomic structure of the ectodomain from HIV-1 gp41. *Nature*. 1997; 387:426–430. [PubMed: 9163431]
46. Tan K, Liu J, Wang J, Shen S, Lu M. Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc Natl Acad Sci U S A*. 1997; 94:12303–12308. [PubMed: 9356444]
47. Wild C, Greenwell T, Matthews T. A synthetic peptide from HIV-1 gp41 is a potent inhibitor of virus-mediated cell-cell fusion. *AIDS Res Hum Retroviruses*. 1993; 9:1051–1053. [PubMed: 8312047]
48. Wild CT, Shugars DC, Greenwell TK, McDaniel CB, Matthews TJ. Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. *Proc Natl Acad Sci U S A*. 1994; 91:9770–9774. [PubMed: 7937889]
49. He Y, Xiao Y, Song H, Liang Q, Ju D, Chen X, Lu H, Jing W, Jiang S, Zhang L. Design and evaluation of sifuvirtide, a novel HIV-1 fusion inhibitor. *J Biol Chem*. 2008; 283:1112–1134.
50. Chong H, Yao X, Sun J, Qiu Z, Zhang M, Waltersperger S, Wang M, Cui S, He Y. The M-T hook structure is critical for design of HIV-1 fusion inhibitors. *J Biol Chem*. 2012; 287:34558–34568. [PubMed: 22879603]

51. Ouyang W, An T, Guo D, Wu S, Tien P. The Potent Human Immunodeficiency Virus Type 1 (HIV-1) Entry Inhibitor HR212 Blocks Formation of the Envelope Glycoprotein gp41 Six-Helix Bundle. *AIDS Res Hum Retroviruses*. 2013
52. Caffrey M, Cai M, Kaufman J, Stahl SJ, Wingfield PT, Covell DG, Gronenborn AM, Clore GM. Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *EMBO J*. 1998; 17:4572–4584. [PubMed: 9707417]
53. Shu W, Ji H, Lu M. Interactions between HIV-1 gp41 core and detergents and their implications for membrane fusion. *J Biol Chem*. 2000; 275:1839–1845. [PubMed: 10636883]
54. Frey G, Chen J, Rits-Volloch S, Freeman MM, Zolla-Pazner S, Chen B. Distinct conformational states of HIV-1 gp41 are recognized by neutralizing and non-neutralizing antibodies. *Nat Struct Mol Biol*. 2010; 17:1486–1491. [PubMed: 21076402]
55. Gustchina E, Li M, Louis JM, Anderson DE, Lloyd J, Frisch C, Bewley CA, Gustchina A, Wlodawer A, Clore GM. Structural basis of HIV-1 neutralization by affinity matured Fabs directed against the internal trimeric coiled-coil of gp41. *PLoS Pathog*. 2010; 6:e1001182. [PubMed: 21085615]
56. Luftig MA, Mattu M, Di Giovine P, Geleziunas R, Hrin R, Barbato G, Bianchi E, Miller MD, Pessi A, Carfi A. Structural basis for HIV-1 neutralization by a gp41 fusion intermediate-directed antibody. *Nat Struct Mol Biol*. 2006; 13:740–747. [PubMed: 16862157]
57. Sabin C, Corti D, Buzon V, Seaman MS, Lutje Hulsik D, Hinz A, Vanzetta F, Agatic G, Silacci C, Mainetti L, et al. Crystal structure and size-dependent neutralization properties of HK20, a human monoclonal antibody binding to the highly conserved heptad repeat 1 of gp41. *PLoS Pathog*. 2010; 6:e1001195. [PubMed: 21124990]
58. Bai X, Wilson KL, Seedorff JE, Ahrens D, Green J, Davison DK, Jin L, Stanfield-Oakley SA, Mosier SM, Melby TE, et al. Impact of the enfuvirtide resistance mutation N43D and the associated baseline polymorphism E137K on peptide sensitivity and six-helix bundle structure. *Biochemistry*. 2008; 47:6662–6670. [PubMed: 18507398]
59. Izumi K, Nakamura S, Nakano H, Shimura K, Sakagami Y, Oishi S, Uchiyama S, Ohkubo T, Kobayashi Y, Fujii N, et al. Characterization of HIV-1 resistance to a fusion inhibitor, N36, derived from the gp41 amino-terminal heptad repeat. *Antiviral Res*. 2010; 87:179–186. [PubMed: 20438763]
60. Ji H, Shu W, Burling FT, Jiang S, Lu M. Inhibition of human immunodeficiency virus type 1 infectivity by the gp41 core: role of a conserved hydrophobic cavity in membrane fusion. *J Virol*. 1999; 73:8578–8586. [PubMed: 10482611]
61. Liu J, Shu W, Fagan MB, Nunberg JH, Lu M. Structural and functional analysis of the HIV gp41 core containing an Ile573 to Thr substitution: implications for membrane fusion. *Biochemistry*. 2001; 40:2797–2807. [PubMed: 11258890]
62. Lu M, Stoller MO, Wang S, Liu J, Fagan MB, Nunberg JH. Structural and functional analysis of interhelical interactions in the human immunodeficiency virus type 1 gp41 envelope glycoprotein by alanine-scanning mutagenesis. *J Virol*. 2001; 75:11146–11156. [PubMed: 11602754]
63. Shu W, Liu J, Ji H, Radigen L, Jiang S, Lu M. Helical interactions in the HIV-1 gp41 core reveal structural basis for the inhibitory activity of gp41 peptides. *Biochemistry*. 2000; 39:1634–1642. [PubMed: 10677212]
64. Wang S, York J, Shu W, Stoller MO, Nunberg JH, Lu M. Interhelical interactions in the gp41 core: implications for activation of HIV-1 membrane fusion. *Biochemistry*. 2002; 41:7283–7292. [PubMed: 12044159]
65. Lee JE, Fusco ML, Hessel AJ, Oswald WB, Burton DR, Saphire EO. Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature*. 2008; 454:177–182. [PubMed: 18615077]
66. Ofek G, Tang M, Sambor A, Katinger H, Mascola JR, Wyatt R, Kwong PD. Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. *J Virol*. 2004; 78:10724–10737. [PubMed: 15367639]
67. Cardoso RM, Zwick MB, Stanfield RL, Kunert R, Binley JM, Katinger H, Burton DR, Wilson IA. Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41. *Immunity*. 2005; 22:163–173. [PubMed: 15723805]

68. Pejchal R, Gach JS, Brunel FM, Cardoso RM, Stanfield RL, Dawson PE, Burton DR, Zwick MB, Wilson IA. A conformational switch in human immunodeficiency virus gp41 revealed by the structures of overlapping epitopes recognized by neutralizing antibodies. *J Virol*. 2009; 83:8451–8462. [PubMed: 19515770]
69. Huang J, Ofek G, Laub L, Louder MK, Doria-Rose NA, Longo NS, Imamichi H, Bailer RT, Chakrabarti B, Sharma SK, et al. Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature*. 2012; 491:406–412. [PubMed: 23151583]
70. Corti D, Voss J, Gamblin SJ, Codoni G, Macagno A, Jarrossay D, Vachieri SG, Pinna D, Minola A, Vanzetta F, et al. A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. *Science*. 2011; 333:850–856. [PubMed: 21798894]
71. Dreyfus C, Laursen NS, Kwaks T, Zuijdgeest D, Khayat R, Ekiert DC, Lee JH, Metlagel Z, Bujny MV, Jongeneelen M, et al. Highly conserved protective epitopes on influenza B viruses. *Science*. 2012; 337:1343–1348. [PubMed: 22878502]
72. Ekiert DC, Bhabha G, Elsliger MA, Friesen RH, Jongeneelen M, Throsby M, Goudsmit J, Wilson IA. Antibody recognition of a highly conserved influenza virus epitope. *Science*. 2009; 324:246–251. [PubMed: 19251591]
73. Ekiert DC, Friesen RH, Bhabha G, Kwaks T, Jongeneelen M, Yu W, Ophorst C, Cox F, Korse HJ, Brandenburg B, et al. A highly conserved neutralizing epitope on group 2 influenza A viruses. *Science*. 2011; 333:843–850. [PubMed: 21737702]
74. Harris AK, Meyerson JR, Matsuoka Y, Kuybeda O, Moran A, Bliss D, Das SR, Yewdell JW, Sapiro G, Subbarao K, et al. Structure and accessibility of HA trimers on intact 2009 H1N1 pandemic influenza virus to stem region-specific neutralizing antibodies. *Proc Natl Acad Sci U S A*. 2013
75. Alam SM, Morelli M, Dennison SM, Liao HX, Zhang R, Xia SM, Rits-Volloch S, Sun L, Harrison SC, Haynes BF, et al. Role of HIV membrane in neutralization by two broadly neutralizing antibodies. *Proc Natl Acad Sci U S A*. 2009; 106:20234–20239. [PubMed: 19906992]
76. Blumenthal R, Durell S, Viard M. HIV Entry and Envelope Glycoprotein-mediated Fusion. *J Biol Chem*. 2012; 287:40841–40849. [PubMed: 23043104]
77. Klasse PJ. The molecular basis of HIV entry. *Cell Microbiol*. 2012; 14:1183–1192. [PubMed: 22583677]
78. Wilen CB, Tilton JC, Doms RW. HIV: cell binding and entry. *Cold Spring Harb Perspect Med*. 2012; 2.
79. Frank GA, Bartesaghi A, Kuybeda O, Borgnia MJ, White TA, Sapiro G, Subramaniam S. Computational separation of conformational heterogeneity using cryo-electron tomography and 3D sub-volume averaging. *J Struct Biol*. 2012; 178:165–176. [PubMed: 22248450]
80. Kuybeda O, Frank GA, Bartesaghi A, Borgnia M, Subramaniam S, Sapiro G. A collaborative framework for 3D alignment and classification of heterogeneous subvolumes in cryo-electron tomography. *J Struct Biol*. 2013; 181:116–127. [PubMed: 23110852]
81. Bartesaghi A, Lecumberry F, Sapiro G, Subramaniam S. Protein secondary structure determination by constrained single-particle cryo-electron tomography. *Structure*. 2012; 20:2003–2013. [PubMed: 23217682]

Highlights

- HIV-1 envelope glycoproteins (Env) are trimers of gp120/gp41 complexes
- HIV-1 Env is target for vaccine design against HIV/AIDS
- X-ray structure of a core fragment of monomeric gp120 has been determined
- Cryo-electron tomography is emerging technology to study Env trimers on intact virus
- Determining high resolution structures of Env trimers is an important goal

Figure 1. gp120 structures obtained by X-ray crystallography

To provide a comparison of the similarities and differences between the various gp120 structures determined by X-ray crystallography, three different sets of superpositions are presented. (A) Superposition of all 24 reported structures of gp120 variants. The PDB IDs of entries included in the superposition are 3NGB, 3TGT, 3SE9, 3SE8, 4DKR, 4DKQ, 4DKP, 4DKO, 3U7Y, 3RJQ, 3TGS, 3TGR, 3TIH, 3TGQ, 3JWD, 2B4C, 2QAD, 3HI1, 1GC1, 2I5Y, 1YYL, 3LQA, 3IDX, and 2NY7. The 3TYG structure was excluded because it does not contain the inner domain of gp120, although the rest of the polypeptide assumes the same conformation as the structures shown here [22]. (B) Color-coded representation of the superposition shown in (A) to display the extent of variation observed in different regions of gp120. 3NGB coordinates are used as the reference structure. The root-mean-square-deviation of the C α backbone of gp120 between all 24 sets of coordinates is < 1 Å for blue regions, 2 Å for white regions, and 4 Å for red regions. N- and C-termini of the 3NGB gp120 core are marked. The dashed line illustrates the overall organization of gp120 into an inner domain that faces the interior and an outer domain that faces the exterior. (C, D) Superpositions of the 14 most recent sets of gp120 coordinates, reported between 2010 and 2012, are displayed as in panels (A) and (B). (E, F) As in panels (A) and (B), superpositions

of the four variants of gp120 structure reported to be present within the same three-dimensional crystal of the gp120 core bound to Fab fragments of the VRC01 antibody are included. The core regions of the gp120 structures are remarkably similar to each other, while the stumps of the variable loop regions included in the crystallized polypeptides are most prone to variation (red).

Figure 2. Changes in molecular architecture of trimeric gp120 complexed to different CD4-binding site and co-receptor binding site ligands

(A–F) Top views of density maps of native trimeric Env in unliganded (A), VRC01-bound (B), b12-bound (C), A12-bound (D), soluble CD4-bound (E) and 17b-bound (F) states. In each case, density maps at resolutions of ~ 20 Å are shown fitted with three copies either of gp120 coordinates or of gp120 bound to the respective ligands: PDB IDs are 3DNN, 3NGB, 2NY7, 3RJQ, 1GC1 and 1GC1, respectively. Chains are colored red for gp120 core, blue for VRC01, cyan for b12, light sea green for A12, yellow for soluble CD4, and forest green for 17b. (G) Schematic representation of trimeric Env in various states, presented with the same coloring arrangement and in the same order in which they appear in panels (A–F). Trimeric gp120 is in the “closed” state in unliganded and VRC01-bound states, “partially open” in b12- and A12-bound states and “fully open” in soluble CD4- and 17b-bound states.

Figure 3. Comparison of effects of VRC01 and sCD4 binding on gp120 monomers vs. native Env trimers

(A, B) Views from two different directions of the superposed structures of the gp120-sCD4-17b complex (PDB ID: 1GC1) and the gp120-VRC01 complex (PDB ID: 3NGB). For clarity, only those regions of VRC01 (blue) and sCD4 (yellow) that are in close proximity to gp120 (light grey for the VRC01 bound conformation, and dark grey for sCD4 bound conformation) are shown. The orientation of the stumps of the V1V2 (red) and V3 (green) loops on the gp120 surface provides a visual marker for gp120 conformation. (C, D) Top views of the surface representations of trimeric gp120 derived by fitting three copies of the gp120-VRC01 structure (C) or gp120-sCD4 structure (D) into density maps determined by cryo-electron tomography from the respective complexes of native trimeric Env. The quaternary conformation of trimeric gp120 in the VRC01 complex is closed, and similar to unliganded trimeric Env, while the quaternary conformation of the sCD4 complex is open, with large rearrangements of gp120 as indicated by changes in position of the V1V2 and V3 loop regions. The color scheme is the same as is used in panels (A) and (B).

Figure 4. Structures of gp41 trimers visualized by X-ray crystallography and of gp41 helices within soluble gp140 trimers visualized by cryo-electron microscopy

(A, B) Superposition of 24 structures reported for trimeric variants of gp41 N-terminal and C-terminal helices in the canonical post-fusion “six-helix bundle” conformation, shown as top (A) and front (B) views. All gp41 coordinates were aligned to the 1AIK structure [44]. The PDB IDs of entries included in the superposition are 1AIK, 2X7R, 2XRA, 3MAC, 3MA9, 2CMR, 3VIE, 1F23, 3AHA, 2Z2T, 3P30, 1ENV, 1K34, 1DLB, 1SZT, 1DF4, 3CP1, 3CYO, 2OT5, 1QR9, 1I5X, 1I5Y, 3VTP, and 3K9A. (C, D) Top and front views of the structure, determined by cryo-electron microscopy, at ~ 9 Å resolution, of trimeric gp140 in an activated, but pre-fusion conformation [36]. The structure represents a complex between the entire ectodomain of trimeric Env and Fab fragments of the monoclonal antibody 17b. The three central densities are assigned to the three copies of the N-heptad repeat helix in gp41, which form a three-helix motif that is more open than that observed in the post-fusion structures shown in panels (A) and (B). The density map is shown fitted with coordinates for the gp120 core (red), heavy (green) and light (yellow) chains of the Fv fragment of the 17b antibody, and the gp41 N-terminal helices (cyan). The gp120 and 17b coordinates are from PDB structure 1GC1 [9], while the gp41 coordinates are from PDB structure 1AIK [44].