

Recent advance in the structural analysis of HIV-1 envelope protein

MATSUDA Zene^{1,2}

¹Research Center for Asian Infectious Diseases, Institute of Medical Science, the University of Tokyo, Tokyo 108-8639, Japan;

²Laboratory of Structural Virology and Immunology, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

Received January 23, 2015; accepted March 3, 2015; published online April 15, 2015

Human immunodeficiency virus type I (HIV-1), a causative agent of AIDS, is affecting today more than 35 millions of people worldwide. The advance of anti-HIV chemotherapy has made AIDS a chronic non-fatal disease in resourceful countries. Long-awaited anti-HIV-1 vaccine is still not with us yet; however, great progress in structural analyses of the envelope protein of HIV-1 in recent years starts to shed light on rational intervention targeted at the envelope protein, as will be reviewed in this article.

HIV-1, envelope glycoprotein, CryoEM, structure, membrane fusion, vaccine

Citation: Matsuda Z. Recent advance in the structural analysis of HIV-1 envelope protein. *Sci China Life Sci*, 2015, 58: 420–424, doi: 10.1007/s11427-015-4852-1

It has been more than 30 years now since the discovery of human immunodeficiency virus type I (HIV-1) as a causative agent of AIDS. During the past three decades, researchers have elucidated the complex nature of this unique retrovirus. Significant advance in chemotherapy has made HIV-1 infection as a chronic manageable disease in developed countries. Another arm of intervention, a vaccine against HIV-1, is unfortunately not with us yet [1,2]. However, our understanding of the structure of the envelope protein (Env) of HIV-1, a major target of vaccine, is significantly advanced in recent years. In addition to the X-ray crystallographic study, the cryo-electron microscopy (CryoEM) technology is helping us grasp the native structure of HIV-1 Env. Now we are able to describe the mechanism of membrane fusion and antibody-mediated neutralization at the molecular level. Here I will review some of recent advance of the structural analysis of HIV-1 Env.

1 Virology of HIV-1

Before getting into the discussion of the structure of Env, I

will briefly describe the virological characteristics of HIV-1. HIV-1 is a complex retrovirus belonging to the lentivirus family. HIV-1 is not an oncogenic retrovirus, but induces immunodeficiency. Common to other replication competent retroviruses, HIV-1 has three structural genes encoding Gag (group specific antigen), Pol (polymerase), and Env. Besides these structural genes, HIV-1 has six additional genes. Two genes, *tat* and *rev*, are classified as essential genes and play a critical role in transcription and transport of mRNAs, respectively [3]. Other four genes are called accessory genes because they are dispensable for viral replication in T-cell lines *in vitro*. However, accessory gene products are needed to modify the host environment to allow efficient virus replication *in vivo*. Many accessory gene products work as adaptors to destroy host restriction factors. Due to the limitation of its small genome size, the virus seems to target host factors whose presence is very detrimental to its replication. The initial study of Vif led to the identification of an APOBEC family as a major restriction factor [4]. Subsequent studies of Vpu, Vpr, Nef, and finally Vpx of the related virus HIV-2/SIV identified respective restrictive host factors [5]. These studies are tightly linked with the cell biology, and the researchers in the field are now more

email: zmtsuda@ibp.ac.cn

cautious about the choice of cell lines they use, since some of the cellular restriction factors are abundant in a particular cell line and not in others.

2 Host immune responses and viral mutations

The above-mentioned accessory gene products deal with the host innate immunity. HIV-1 copes with the dynamic acquired immunity, both humoral and cellular immunities, by mutations. The robust HIV-1 replication mediated by the error- and recombination-prone reverse transcriptase contributes to them. Shortly after the infection, HIV-1 establishes quasispecies that adapt to the pressure of the immune responses [6]. Being on the surfaces of virions and infected cells, Env is a major target of the host immune responses. The diversity of Env sequence is enormous and antibody responses usually cannot catch up with the mutations of Env. This is further worsened because HIV-1 targets the key player of the host immune system, CD4 positive T cells, and induces their death or dysfunction [7]. Most of the HIV-1 infected individuals would succumb to death without an appropriate intervening therapy. However, there are some people who can manage HIV-1 infection. Such people called elite controllers can control HIV-1 replication and stay disease-free without the anti HIV-1 chemotherapy. Analyses of these elite controllers, especially elite neutralizers, would lead to the identification of broadly neutralizing antibodies.

There are only handful of classes of broadly neutralizing antibodies (bnAbs) [8,9]. The breadth of the neutralizing activity of a particular bnAb is evaluated from the efficiency of its neutralization of different subtypes of HIV-1. The subtypes of HIV-1 are called clades. Clades are based on the nucleotide and amino acid sequences, and the sequences can differ from clade to clade by about 20%–30%. The bnAbs usually manifest cross-clade neutralization. One should be aware that a clade is not necessarily directly linked with any biological phenotypes and should not be confused with serotypes in other viruses.

3 Envelope protein of HIV-1

The target of the bnAbs is the Env protein of HIV-1. HIV-1 Env is a heavily glycosylated membrane protein composed of the heterodimer of gp120 and gp41. These two subunits are generated from a precursor protein called gp160 by proteolytic processing by cellular proteases. The mature HIV-1 Env is a trimer of these heterodimers. The gp120 subunit is responsible for receptor/coreceptor recognition. The receptor of HIV-1 Env is CD4 and there are two major coreceptors for HIV-1, namely CXCR4 and CCR5 [10]. Upon the interaction between the gp120 and receptor/coreceptor, gp120 will undergo the conformational changes and this, in

turn, is believed to trigger the cascade of conformational changes of the associated gp41 subunit, which plays a major role in induction of membrane fusion. The gp41 molecule belongs to a class I fusion protein like HA2 of influenza virus [11]. The class I fusion protein is composed of trimer of mainly helix proteins and forms a quaternary structure called six-helix bundle (6HB) during the membrane fusion. The atomic structures of various 6HB have been determined [12]. For the long time, the only available structures for gp41 were those of 6HB.

4 Structure of HIV-1 Env

The gp120 subunit, being heavily glycosylated (nearly the half of the mass is from sugar moieties) and with five highly variable flexible loops, has been a difficult target for the structural analysis. The first seminal structure was determined for the engineered monomeric gp120, in which the most of the variable loops were replaced with short linker sequences. Furthermore, the engineered gp120 was stabilized by mixing it with the fragment of CD4 and a monoclonal antibody, 17 β [13]. Since then, several structures of gp120, both in the free form or complex with other ligands, have been determined. However, the quaternary structure containing both gp120 and gp41 in the trimeric form has been resistant to crystallization. Meanwhile, CryoEM started to provide the quaternary structure of Env [14–17] in low resolution. Applying both crystallography and CryoEM approaches, the group at Scripps Institute revealed the overall structure of gp120/gp41 complex in a relatively high resolution in 2013 [18,19]. This is another landmark achievement in the field.

The HIV-1 Env protein used for both CryoEM and crystallographic analyses was derived from the BG505 strain belonging to clade A. The protein was truncated at residue 664 in the beginning of membrane proximal external region (MPER), and remains as a soluble non-membrane anchored form. Several modifications were introduced to facilitate structural analyses: (i) gp120 and gp41 are connected by an artificial disulfide bond (between residues 501 of gp120 and 605 of gp41, termed SOS) to avoid shedding of gp120, (ii) the isoleucine residue at 559 of gp41 was mutated to proline (termed IP) to stabilize gp41-gp41 association. These two modifications have been shown not to affect the overall structures of gp120. Thus modified HIV-1 Env called BG505 SOSIP.664 was more thermostable than the native Env. In addition to these modifications, the engineered Env was stabilized by complexing with neutralizing antibodies.

The overall structure of gp120 subunit of BG505 SOSIP.664 is similar to the previously determined structure. The structural information on variable loops often missed in the previous crystal structure was provided except for V4. V1/V2 and V3 loops are located in the central apical region of the gp120 trimer. V5 and V4 loop are located to the pe-

ripheral region of the gp120 trimer. The part of the V3 loop was located underneath the V1/V2 loop (Figure 1).

Unfortunately the resolution of gp41 portion was not high enough to allow the assignment of individual amino acid residues in gp41. However, both the N-terminal heptad repeat (NHR) and C-terminal heptad repeat (CHR) were resolved as separated helices. NHR stood perpendicular to the viral membrane and CHR lying horizontally under the NHR trimers. The C-terminal half of NHR forms a trimeric helix bundle-like structure. The similar trimeric helix-bundle of NHR is observed in the post fusion six-helix bundle (6HB) structure, but the revealed three-helix bundle most likely represents the prefusion state rather than activated intermediate. The N-terminal region of gp41 containing fusion peptide (FP) and the connection to the NHR region was not resolved. The CHRs located near the bottom of the NHR bundle are wrapping around the base of NHR trimer to form a triangular shape. It was shown that the middle portion of CHR has a strong interaction with the bottom of NHR. This structure denies the presence of pre-

formed 6HB in the native HIV-1 Env as proposed before [20], and suggests the complex conformational rearrangements during the formation of 6HB.

5 Further update of HIV-1 Env structural and its implication for membrane fusion

The molecular details of HIV-1 Env-mediated membrane fusion have been elusive. The above mentioned architectures of gp120 and gp41 clarify some aspects of structural rearrangement of the gp120/gp41 during membrane fusion. In the prefusion state, the exterior hydrophobic surface of the NHR is covered by the surrounding inner domain of gp120. This surface of the NHR bundle will eventually interact with the CHR to form the six-helix bundle. The receptor binding will induce positional shift of gp120 from the closed to open state; this allows the exterior surface of NHR to interact with the corresponding region of CHR. How this association of NHR and CHR is exactly achieved

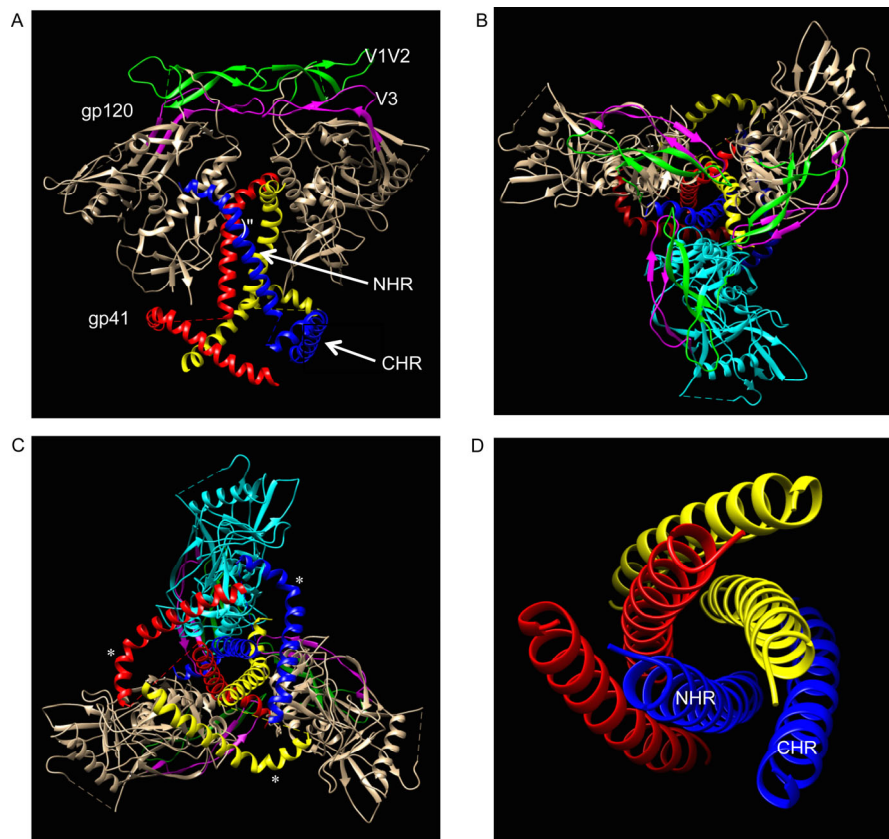


Figure 1 Structure of BG505 SOSIP.664 determined by CryoEM (PDB 4NCO). Note that bound PGV04 bnAb is not shown to show HIV-1 Env region clearly. A, Side view. Main chains of two gp120 subunits are shown in gold. V1V2 and V3 loops are shown in green and magenta, respectively. They are in the apical center of trimer complex. Three gp41 subunits are shown in blue, red, and yellow. For the clear view of the gp41 bundle, the gp120 subunit corresponding to blue gp41 is removed. NHR and CHR helices are indicated. The kink in NHR is shown with an asterisk. B, Top view corresponding to the panel A is shown. The gp120 subunit removed in the panel A is shown in cyan. The V1V2, V3 are shown in the same color as in A. Dashed lines indicate the position of the missing V4 loops. C, The bottom view (seen from the viral membrane). Three CHRs of gp41 form a triangle at the base of the NHR bundle. The positions of the kink in CHR are shown with the asterisks. D, The structure of six-helix bundle (6HB) of gp41 (PDB1AIK) is shown for a reference. In the 6HB, CHR is tightly docked into the grooves generated by the central bundle of NHR. NHR and CHR of the blue subunit are indicated. The images were generated by the UCSF Chimera program.

is the interest of the researchers like us who are interested in the mechanism of membrane fusion. Molecular insight of these conformational changes are finally revealed by the Group of P. Kwong from National Institute of Allergy and Infectious Diseases [21] with the revelation of gp120/gp41 structure with a higher resolution in 2014. They worked on the same BG505 SOSIP 664 envelope protein, but they added the second bnAb (35O22) recognizing the gp120/gp41 boundary region. This improved the resolution of the crystal structure and assignment of the amino acid residues of gp41 in prefusion status was provided for the first time. This landmark work provided the model for the possible structural rearrangements of gp41 during membrane fusion. By the comparison of the new structure with the previously determined 6HB structure, a region of gp41 was shown to undergo a structural change from the non-helical state to the helical structure during membrane fusion. This structural rearrangement is similar to the previously shown spring-loading model in influenza HA2 during membrane fusion. Here a common structural rearrangement for both class I fusion proteins, gp41 and HA2, was established. This model will guide us to further understand the molecular basis for the detailed membrane fusion mechanism.

We have been interested in the structure and function of gp41. Different from another report of nearly full length HIV-1 Env [22], the current structure of BG505 SOSIP.644 misses the structural information on the MPER, membrane-spanning domain (MSD) and cytoplasmic region of gp41. MPER is one of the major targets of the bnAbs and MSD clearly plays a critical role in biological function of HIV-1 Env [23]. Although it would be a challenging task, further elucidation of the structure of full length Env including MSD and cytoplasmic tail together with lipid bilayers is highly anticipated.

6 Broadly neutralizing antibodies and vaccine

Determination of the structure of the complex between SOSIP and bnAbs provides us the molecular basis for the mechanism of neutralization by bnAbs. Since not so much molecular information is available on the vast majority of other viruses, this understanding of neutralization at the molecular level is somehow unique for HIV-1. Several different categories of bnAbs against HIV-1 have been identified [8]. The bnAb directed to the receptor/coreceptor binding sites of gp120 represents one category and its mechanism of neutralization is relatively straightforward: it interferes with the interaction between the receptor/coreceptor and Env, and blocks the subsequent conformational changes of Env. Other categories of bnAbs bind to the different regions of Env and they are now shown to interfere with the structural changes of Env necessary for membrane fusion. These may include bnAbs binding to (i) V1/V2 loops of

Env with associated glycans, (ii) the apex of gp120 with association with attached glycans, (iii) gp41 MPER. The most recently identified bnAb targets the gp120/gp41 cleavage site [24,25].

Glycans are once regarded as an obstacle for the development of effective antibody responses, since it is rather immunologically inert to the host. When we learn the fact that the interactions of many bnAbs and HIV-1 Env depend on the presence of glycans, we may have to modify our view. Probably some of the glycan is essential for the structural integrity of HIV-1 Env while others are placed as interfering barriers against antibodies. Recent progress in the technologies now allows us to examine types of different sugar moieties attached to Env. One interesting insight obtained through these detailed structural information on Env-bnAb complexes is that not only the target region but the way the particular antibody binds determine the outcome. Two antibodies bound to the very close regions may become neutralizing or non-neutralizing depending on the direction of the antibody binding.

As discussed above, we now have much better understanding of the structures of native Env itself and those of Env-bnAb complexes. The next remaining issue is how to induce such bnAbs responses in humans. A recent report by Doria-Rose et al. [26] showed the development pathway for potent V1/V2- directed neutralizing antibodies in an HIV-1 infected individual. This study involves a sophisticated *in vitro* clonal analysis of antibody producing B cells derived from elite neutralizers in combination with structural analyses of antibodies. In the body of this person, there is clear evolution of potent bnAbs directed against HIV-1. To our surprise, this evolution took place in relatively short time. The question is how general this kind of observation is and how to elicit this kind of Ab by man-made immunogens. It is interesting to see how the research pathway in this field will evolve in future.

- 1 Ensoli B, Cafaro A, Monini P, Marcotullio S, Ensoli F. Challenges in HIV vaccine research for treatment and prevention. *Front Immunol*, 2014, 5: 417
- 2 Kim JH, Excler JL, Michael NL. Lessons from the RV144 Thai Phase III HIV-1 vaccine trial and the search for correlates of protection. *Ann Rev Med*, 2015, 66: 423–437
- 3 Nekhai S, Jeang KT. Transcriptional and post-transcriptional regulation of HIV-1 gene expression: role of cellular factors for Tat and Rev. *Future Microbiol*, 2006, 1: 417–426
- 4 Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature*, 2002, 418: 646–650
- 5 Strebel K. HIV accessory proteins versus host restriction factors. *Curr Opin Virol*, 2013, 3: 692–699
- 6 Smyth RP, Davenport MP, Mak J. The origin of genetic diversity in HIV-1. *Virus Res*, 2012, 169: 415–429
- 7 Doitsh G, Galloway NL, Geng X, Yang Z, Monroe KM, Zepeda O, Hunt PW, Hatano H, Sowinski S, Munoz-Arias I, Greene WC. Cell death by pyroptosis drives CD4 T-cell depletion in HIV-1 infection. *Nature*, 2014, 505: 509–514
- 8 Kwong PD, Mascola JR. Human antibodies that neutralize HIV-1:

- identification, structures, and B cell ontogenies. *Immunity*, 2012, 37: 412–425
- 9 Mouquet H. Antibody B cell responses in HIV-1 infection. *Trends Immunol*, 2014, 35: 549–561
 - 10 Loftin LM, Kienzle M, Yi Y, Collman RG. R5X4 HIV-1 coreceptor use in primary target cells: implications for coreceptor entry blocking strategies. *J Trans Med*, 2011, 9: S3
 - 11 White JM, Delos SE, Brecher M, Schornberg K. Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit Rev Biochem Mol Biol*, 2008, 43: 189–219
 - 12 Chan DC, Fass D, Berger JM, Kim PS. Core structure of gp41 from the HIV envelope glycoprotein. *Cell*, 1997, 89: 263–273
 - 13 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
 - 14 Zhu P, Liu J, Bess J Jr., Chertova E, Lifson JD, Grise H, Ofek GA, Taylor KA, Roux KH. Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature*, 2006, 441: 847–852
 - 15 Roux KH, Taylor KA. AIDS virus envelope spike structure. *Curr Opin Struct Biol*, 2007, 17: 244–252
 - 16 Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S. Molecular architecture of native HIV-1 gp120 trimers. *Nature*, 2008, 455: 109–113
 - 17 Merk A, Subramaniam S. HIV-1 envelope glycoprotein structure. *Curr Opin Struct Biol*, 2013, 23: 268–276
 - 18 Lyumkis D, Julien JP, de Val N, Cupo A, Potter CS, Klasse PJ, Burton DR, Sanders RW, Moore JP, Carragher B, Wilson IA, Ward AB. Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science*, 2013, 342: 1484–1490
 - 19 Julien JP, Cupo A, Sok D, Stanfield RL, Lyumkis D, Deller MC, Klasse PJ, Burton DR, Sanders RW, Moore JP, Ward AB, Wilson IA. Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science*, 2013, 342: 1477–1483
 - 20 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
 - 21 Pancera M, Zhou T, Druz A, Georgiev IS, Soto C, Gorman J, Huang J, Acharya P, Chuang GY, Ofek G, Stewart-Jones GB, Stuckey J, Bailer RT, Joyce MG, Louder MK, Tumba N, Yang Y, Zhang B, Cohen MS, Haynes BF, Mascola JR, Morris L, Munro JB, Blanchard SC, Mothes W, Connors M, Kwong PD. Structure and immune recognition of trimeric pre-fusion HIV-1. *Nature*, 2014, 514: 455–461
 - 22 Mao Y, Wang L, Gu C, Herschhorn A, Xiang SH, Haim H, Yang X, Sodroski J. Subunit organization of the membrane-bound HIV-1 envelope glycoprotein trimer. *Nat Struct Mol Biol*, 2012, 19: 893–899
 - 23 Kondo N, Miyauchi K, Meng F, Iwamoto A, Matsuda Z. Conformational changes of the HIV-1 envelope protein during membrane fusion are inhibited by the replacement of its membrane-spanning domain. *J Biol Chem*, 2010, 285: 14681–14688
 - 24 Falkowska E, Le KM, Ramos A, Doores KJ, Lee JH, Blattner C, Ramirez A, Derking R, van Gils MJ, Liang CH, McBride R, von Bredow B, Shivatare SS, Wu CY, Chan-Hui PY, Liu Y, Feizi T, Zwick MB, Koff WC, Seaman MS, Swiderek K, Moore JP, Evans D, Paulson JC, Wong CH, Ward AB, Wilson IA, Sanders RW, Poignard P, Burton DR. Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved envelope trimers. *Immunity*, 2014, 40: 657–668
 - 25 Blattner C, Lee JH, Sliepen K, Derking R, Falkowska E, de la Pena AT, Cupo A, Julien JP, van Gils M, Lee PS, Peng W, Paulson JC, Poignard P, Burton DR, Moore JP, Sanders RW, Wilson IA, Ward AB. Structural delineation of a quaternary, cleavage-dependent epitope at the gp41-gp120 interface on intact HIV-1 Env trimers. *Immunity*, 2014, 40: 669–680
 - 26 Doria-Rose NA, Schramm CA, Gorman J, Moore PL, Bhiman JN, DeKosky BJ, Ernandes MJ, Georgiev IS, Kim HJ, Pancera M, Staube RP, Altae-Tran HR, Bailer RT, Crooks ET, Cupo A, Druz A, Garrett NJ, Hoi KH, Kong R, Louder MK, Longo NS, McKee K, Nonyane M, O'Dell S, Roark RS, Rudicell RS, Schmidt SD, Sheward DJ, Soto C, Wibmer CK, Yang Y, Zhang Z, Program NCS, Mullikin JC, Binley JM, Sanders RW, Wilson IA, Moore JP, Ward AB, Georgiou G, Williamson C, Abdool Karim SS, Morris L, Kwong PD, Shapiro L, Mascola JR. Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature*, 2014, 509: 55–62

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.