AIDS virus envelope spike structure
Kenneth H Roux and Kenneth A Taylor

The envelope (Env) spikes on HIV-1 and closely related SIV define the viral tropism, mediate the fusion process and are the prime target of the humoral response. Despite intensive efforts, Env has been slow to reveal its structural and functional secrets. Three gp120 subunits comprise the ‘head’ of Env and three gp41 subunits comprise the ‘stalk’ and other membrane-associated elements. The recent description of the core structure of unliganded (untriggered) gp120, compared to earlier CD4-liganded atomic structures, reveals dramatic conformational reorganization of the components and suggests a mechanism for the initiation of fusion. The structure of the key V3 loop, both in isolation and in association with the liganded core, helps define its role in fusion and as a prime target of neutralizing antibodies. Additional details are emerging regarding the structure of gp41 as it transitions from the preliganded configuration to the fusion intermediate (fusion-active or prehairpin intermediate) configuration, although much remains speculative. Recent advances in cryoelectron tomography are giving us the first glimpses of the overall three-dimensional structure of Env, which, when fitted with the available component atomic structures, provides new insights into the organization of the structural elements within the trimeric spike.

Addresses
Department of Biological Science and Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306–4370, USA

Corresponding author: Roux, Kenneth H (roux@bio.fsu.edu)

Current Opinion in Structural Biology 2007, 17:244–252
This review comes from a themed issue on Macromolecular assemblages
Edited by Eva Nogales and Wesley I Sundquist

Available online 28th March 2007
0959-440X/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
DOI 10.1016/j.sbi.2007.03.008

Introduction
Recent insights into the structure of the HIV and SIV envelope (Env; see Box 1 for glossary of terms) spike, primarily from X-ray crystallography of core component peptides and cryoelectron microscopy (cryo-EM) of virion-associated Env spikes, have provided considerable insight into the basic structure and probable function of these molecules, although many important features and mechanisms are still poorly defined. Elucidation of the structure of the Env spike will probably facilitate the development of long-anticipated humoral immunity vaccines.

Env spikes are derived from gp160, a precursor protein enzymatically cleaved into gp120 and gp41 subunits by cellular furin protease before cell-surface expression and incorporation into budding virus particles. Heterotrimeric Env spikes consist of three surface (SU, gp120) domains atop three transmembrane (TM, gp41) domains (Figure 1a). Heavily glycosylated SU is divided into five variable (V) and five constant (C) regions, which, when folded, comprise an inner domain, an outer domain and a bridging sheet (Figure 1b). Functional sites include the CD4 binding site and the chemokine receptor (CR) binding site. CD4 and the CR represent the receptor and coreceptor molecules, respectively, on the target cell used by the virus to gain entry. TM is subdivided into an ectodomain (ECTO), a membrane-spanning domain (MSD) and a long cytoplasmic tail (GT) (Figure 2a). The N- to C-terminal segments of TM\textsubscript{ECTO} are the fusion peptide (FP), N-terminal heptad repeat (helical region, HR), loop region, C-terminal HR and the membrane-proximal external region (MPER).

A variety of Env spike structural elements (i.e. core fragments and peptides) have been described at the atomic level, often in complex with ligands or monoclonal antibodies (mAbs) [1,2,4–7,8,9]. The core fragments, although only a fraction of the mass of the monomeric subunits, have been sufficiently complete to enable the proposal of molecular models of trimeric arrays of SU or TM, if not intact Env spikes [1,2,3,8,10,11]. As enticing as these models are, they nevertheless represent educated guesses that need independent verification. Recent advances in cryoelectron tomography (cryo-ET) have facilitated the direct observation and structural analysis of native Env spikes \textit{in situ}, albeit at less than atomic resolution [12,13]. Despite this limitation, the current resolution permits initial efforts at fitting the known atomic structures into the cryo-ET reconstructions.

Here, we describe the current state of atomic structures of the HIV and SIV Env spike components, as well as initial efforts to integrate these structures into averaged 3D cryo-ET images of spikes on mature virions.

Atomic structure of SU and TM
Structure of SU
There are now several crystal structures of core fragments (i.e. fragments missing the more flexible elements, such as V1/V2 loops, V3 loops, N- and C-terminal peptides, and carbohydrates) representing two of the functional states...
of AIDS virus gp120: SIV gp120 in a preliganded (pre-fusion) presumptive native state conformation (Figure 1b) [2**]; and HIV-1 gp120 liganded with CD4 and a mAb Fab (Figure 1c, ligands not shown) [1,3**]. Comparison of the two forms indicates profound conformational changes, particularly of the bridging sheet and inner domain, with segmental movements of up to 4 nm probably resulting from CD4 binding (compare Figure 1b,c) [2**]. In agreement with these data, small-angle X-ray scattering studies on fully glycosylated recombinant soluble gp120–gp41ECTO show a structure with several prominent protrusions, some of which (including the presumptive V3 loop) undergo significant conformational rearrangement upon CD4 binding [14].

The composition of the ~35 amino acid (aa) V3 loop influences CD4 binding, coreceptor specificity and the degree to which neutralizing epitopes are expressed [15]. V3 displays a highly variable sequence, but maintains a rather conserved β-turn structure, with a disulfide-stabilized base and minimal sequence variation at the tip of the loop [16–18]. The V3 loop is believed to be partially occluded by the V1/V2 loop in the unliganded state [19*]. However, as revealed by the comparison of atomic models, it undergoes pronounced rearrangement to most likely protrude from the apex of the trimer spike upon CD4 binding (Figure 1d), whereupon it probably engages the CR [2**,3**]. The V3 loop is an important target of neutralizing Abs [19*].

Another prominent loop, V1/V2, is thought to partially shield the V3 loop, the coreceptor-binding site and the CD4 binding site, and might do so, in part, by overlapping a neighboring protomer [1,2**,19*,20,21]. The loop is otherwise not well defined structurally and is not typically the target of neutralizing Abs [19*,22–24].

Structure of TM

TM (gp41) is a 344 aa glycoprotein composed of a 172 aa ECTO, a 21 aa MSD, and an ~150 aa (for HIV-1) or ~164 aa (for SIV) CT (Figure 2a). Very little is known about the gp41 prefusion structure [25], with the exception of two adjacent segments of the MPER defined by, and co-crystallized with, their respective neutralizing Fab mAbs, 2F5 and 4E10 [4**,5*]. The region encompassed by the helical heptad repeats, N-HR and C-HR, is apparently folded in such a way as to bring FP into close proximity with MPER (Figure 2a) [26]. Ab binding studies demonstrate that the bulk of TM, that is, N-HR, C-HR and the

---

Box 1  Glossary of terms

- **C**: constant region of SU
- **CR**: chemokine receptor
- **Cryo-ET**: cryoelectron microscopy tomography
- **CT**: cytoplasmic tail of TM
- **ECTO**: ecto(external) domain of TM
- **Env**: envelope glycoprotein
- **FP**: fusion peptide
- **MPER**: membrane-proximal external region of TM
- **MSD**: membrane-spanning domain of TM
- **N-HR, C-HR**: N- and C-heptad repeat (helical region) of TM
- **SU**: surface (gp120) portion of Env
- **TM**: transmembrane (gp41, unless truncated) portion of Env
- **V**: variable loop of SU

---

Figure 1

Env spike and atomic structures of SU. (a) Schematic representation of an Env spike, showing the trimer of heterodimers configuration. (b) Atomic structure of the unliganded SIV gp120 core (PDB code 2BF1; [2**]). Arrows indicate structural rearrangement following ligand binding. (c) Atomic structure of CD4-liganded HIV-1 gp120 core (CD4 ligand not shown) (PDB code 1GC1; [1]). (d) Atomic structure of liganded HIV-1 gp120 core containing the V3 loop (PDB code 2B4C; [50]). In (b–d), blue, outer domain; green, inner domain; yellow, bridging sheet; red, CD4 binding site residues; mauve, V3 base (a,b) or loop (c); white, V1/V2 stem. The Env trimer axis would be to the right of each monomer, with the viral membrane below.
connecting loop, is essentially obscured in the native Env spike [6\textsuperscript{*}] and perhaps might not be in helical form at this stage [25]. A highly engineered version of gp140, in which the SU has been largely truncated, shows TM ECTO (minus the MPER) to be in a rod-like configuration, with both N-HR and C-HR exposed [6\textsuperscript{*}], a configuration that could represent a fusion (prehairpin) intermediate configuration (Figure 2b). The fusion intermediate state is believed, by extrapolation from the homologous structure of hemagglutinin (HA) from influenza virus, to be in the form of an extended three-helix rod, perhaps stabilized by association with SU, that inserts the three N-terminal hydrophobic FPs into the outer leaflet of the plasma membrane of the target cell [10,11]. Further evidence of the existence of the three-helix rod derives from a crystal structure showing the neutralizing human mAb D5 engaged in primarily hydrophobic interactions with an artificially stabilized N-HR trimer [7\textsuperscript{*}]. Similar neutralization potencies of single-chain Fv, Fab and IgG D5 [27] and other neutralizing Abs [28] for the three-helix rod argue for a significant gap between the viral and target membranes during the transient (~15 min) fusion intermediate phase [29]. However, data demonstrate that a small C-HR peptide (C37) is sterically restricted from binding to the N-HR helix when tethered to molecules considerably smaller than IgG Fab [30\textsuperscript{*}]. Subtle structural differences in the mode of access to the N-HR core, shown in the same orientations, may represent the conformation of this region in the prehairpin intermediate state depicted in (b). N-HR, gold; C-HR, red.
support the contribution of multiple activated Env spikes to the fusion event [32].

In contrast to the uncertainty described above, what is widely believed to be the post-fusion configuration of the gp41 trimer is well-defined structurally as a highly stable six-helix bundle in which the three C-HRs have folded back in a hairpin configuration and are nestled in an antiparallel orientation within the largely hydrophobic grooves formed by the compact N-HR triad core (Figure 2d,g–j) [8–11]. The transition from the energetically unstable extended three-helix rod to the energetically stable compact six-helix bundle would serve to pull the viral and target cell membranes toward one another. The appeal of this model is complicated by data indicating that the formation of the six-helix bundle might not be complete until after the fusion pore complex has formed [33]. Accumulating evidence indicates that the unique properties of the FP allow it to first insert into the target membrane outer leaflet and adopt a β structure, perhaps in association with FPs from adjacent protomers, leading to membrane destabilization [34,35]. It is tempting to speculate that ligand-induced conformational changes in the MPER–MSD–CT region perform a similar function in the viral membrane, leading to the coalescence of the outer leaflets of the two membranes followed by the formation of a fusion pore (Figure 2b). The transition from the three-helix rod to the six-helix bundle represents a topological conundrum [8,10] that probably involves complex molecular maneuvers, as illustrated by the hypothetical models of the fusion process depicted in Figure 2c,d and e,f.

Structure of the envelope spike

The Env spikes on the surface of HIV-1 and SIV have long been thought to be trimeric [10], a common motif among Env spikes on other viruses [36]. Scanning transmission electron microscopy (STEM) of fixed detergent-solubilized HIV-1 and SIV Env glycoproteins showed a predominance of trimeric structures [37,38]. Clear visual evidence of the trimeric nature of the Env spikes on the surface of SIV and HIV-1 was obtained by EM tomography of virions preserved in negative stain [39]. Averaged images showed a three-bladed propeller-like structure.

Several SU trimer models have been published based on extrapolation from atomic structures of monomeric gp120 cores of liganded HIV-1 and unliganded SIV [1,2,3,20,21]. SIV and HIV-1 Env are considered similar enough in sequence to allow valid structural comparisons [2,40]. The proposed liganded trimer models differ considerably due largely to differences in the underlying assumptions upon which they were built. One of the consequences of these assumptions is that the CD4 binding sites are oriented distally and tilted slightly toward the threefold axis in a liganded spike model proposed by Chen et al. [2] and away from the threefold axis in the models of Kwong et al. [21] and Huang et al. [3]. The functional implications of these differences, with regard to fusion and Ab-mediated neutralization mechanisms, are significant.

It is widely assumed that the failure of Env-based vaccine candidates to date relates, in part, to the difficulty in generating soluble versions of Env proteins that faithfully mimic key structural features of native in situ Env. Immunization with monomers and smaller fragments proved ineffective early on, and was followed by various strategies aimed at generating soluble trimers. Recombinant soluble gp160ECTO in which furin protease cleavage sites were mutated to prevent processing of gp160 to gp120–gp41ECTO yielded soluble oligomers, but didn’t allow proper folding or subunit association. The introduction of trimerization motifs, such as GCN4 and FT, in association with the C termini of gp120 has yielded stable trimers with improved qualities, but a loose trimerization morphology and other physical and immunological abnormalities persist. The addition of disulfide-bond-stabilized gp41ECTO did not help appreciably (trimerization strategies are reviewed in [19*]). A promising advance is the introduction of a disulfide bond to stabilize the otherwise somewhat labile association between gp120 and gp41ECTO coupled with a point mutation in gp41 that stabilizes trimer subunit association. These constructs form tight stable trimers that resemble those imaged on virions, at least at the low resolution afforded by negative-stain EM; however, these are still not faithful mimics of Env spikes in some respects [41]. Another approach is to immunize using non-infectious virus particles or virus-like particles (VLPs), which presumably present Env spikes in their native form. As with the other approaches, these are yet to prove themselves [19*]. Ideally, a crystallized intact or substantially intact trimer would be structurally definitive, but the likelihood of success is less than assured in light of the known flexibility of various protomer components, possible flexibility between protomers, probable lipid bilayer association with MPER and the extensive carbohydrate shield on gp120. In fact, large regions appear to be in constant motion as part of the conformational masking defense of potential epitopes from Abs [42,43].

As a further step in determining the 3D structure of the trimeric Env spike and the location of key features, we have applied cryo-ET to the analysis of Env spikes on a mutant version of SIV previously shown, by biochemical analyses and negative-stain EM, to express high levels of Env [12*]. Subsequently, Zanetti et al. [13] produced a second cryo-ET spike model. Both laboratories took advantage of a very similar mutant SIV provided by the same source, but used different data collection and analysis approaches. The resultant tomographic density map images display some surprisingly different features
Our reconstruction shows a complex surface for the SU trimer consisting of four features, including the main, lateral and proximal lobes, which constitute the bulk of the spike volume (Figure 3a,b). Three distal projections (arches) form a peak at the apex of the SU trimer that obscures a hollow cavity, a feature more apparent in surface-rendered models using a higher density threshold (Figure 3c). This feature accounts for most of the difference in height between the two reconstructions. Emanating from each proximal lobe is the fourth feature, a short leg that angles outward, whereupon it meets the viral membrane and continues laterally just above the plane of the membrane to form a foot. The three splayed legs of the model create a distinct cavity at the threefold axis just above the membrane. The legs and feet are believed to comprise the solvent-exposed portion of the external surface of the SU trimer.
domain of TM and indicate that the MPERs of the subunits are not, as widely assumed, self-associated.

By contrast, the Zanetti et al. [13*] reconstruction consists of three distinct SU subunits perched upon a narrow TM stalk, more in line with the traditional depiction of the Env spike (Figure 3d,e). No evidence of a tripod configuration for gp41 is apparent, although the base of the TM stalk does appear to be somewhat flared.

It may be instructive to compare the two cryo-ET SIV spike averages to that of Moloney murine leukemia virus (MMLV), the only other retroviral Env spike similarly analyzed to date [44*]. The MMLV spike also shows threelfold symmetry, but is more box-like than the SIV spike models (Figure 3f,g). As in our model, it displays multiple, although more pronounced, distal and lateral projections. Both the MMLV and our SIV models share the novel open tripod leg feature.

**Fitting atomic structures into cryoelectron tomography density maps**

To explore how the cryo-ET models reconcile with the available crystallographic data, both groups attempted to place the atomic model of the unliganded SIV gp120 core monomer of Chen et al. [2**] into their respective Env spike density maps, despite the fact that the atomic model is missing substantial components, including the N- and C-terminal peptides and most of the V loops (Figure 3h–j).

The missing parts of the atomic structures precluded attempts at automated fitting. By employing manual fitting and considering the probable placement of the N and C termini, and the solvent exposure of the glycans, we found that the optimum fit fortuitously oriented the truncated stems of the V1/V2 and V3 loops toward the otherwise unoccupied proximal and lateral lobes, respectively (Figure 3h). We speculate that the mass projecting from the main lobe toward the peak apex (arch in Figure 3c) might represent an alternative position of the flexible V3 loop (i.e. projecting toward the target cell surface) on a fraction of the Env subunits. A recently proposed trimer model based on the atomic structure of the CD4-bound fusion-active HIV-1 protomer shows the V3 loops in a similar configuration [3**]. Our model places the CD4 binding site (red peptides) on the outside of the spike, projecting disto-laterally, similar to the position originally proposed by Kwong et al. [21] for the fusion-active conformation, rather than projecting somewhat axially, as proposed by Chen et al. [2**,45].

Zanetti et al. proposed two potential placements of the Chen et al. [2**] unliganded SIV gp120 core into their cryo-ET density map. The rotational orientation of one of the fitted monomers is roughly similar to ours, positioning the CD4 binding site in a disto-lateral position (compare Figure 3i,j), but the distal portion is canted medially compared to our placement. In the second model of Zanetti et al., the orientation about the long axis of the protomer is reversed, resulting in a more axially (inward) oriented CD4 binding site (compare Figure 3i,j). A significant difference between the Env reconstructions is the volume available for the gp120 atomic structure. Whereas our model can accommodate at least some of the projecting mass of the V3 and V1/V2 loops (mesh-enclosed volumes, Figure 3h), the head of the Zanetti et al. spike is completely filled by the atomic core structure. They propose that excessive flexibility in the loops interfered with effective averaging of much of the spike density. For comparison, the MMLV core SU atomic structure fitted in the MMLV density volume is also shown (Figure 3k) [44*].

The other significant difference between the two SIV cryo-ET models is the configuration of the presumptive MPER of gp41ECTO (i.e. the leg/stalk region). The atomic structure of unliganded TM available for placement is considerably less complete than that of unliganded SU and is limited to two adjacent short peptides of the MPER, which were independently co-crystallized with the broadly neutralizing mAbs 2F5 and 4E10 [4**,5**]. The amphipathic nature of the MPER has been alternatively interpreted as contributing to trimerization (via a rod-like stalk configuration) or fostering an extensive interaction with the viral membrane.

We docked the 2F5 and 4E10 peptides into the tripod-like legs of our SIV spike model under the assumption that this highly conserved region would be similarly configured in HIV-1 and SIV (Figure 3h). This assumption is strengthened by the recent observation that both epitopes, when molecularly grafted into homologous positions in SIV Env, are immunologically indistinguishable from those in HIV-1, rendering the resultant chimera neutralizable by 2F5 and 4E10 mAbs [40]. This leg/foot docking configuration precludes trimerization contacts between the MSDs, displacing them ~4–5 nm from the threefold axis. In support of this model are data showing that both epitopes are optimally expressed when associated with lipid [4**,5**]. The splayed legs of our model readily accommodated not only the peptides, but also their associated Fab mAbs (not shown) [12**]. Placement of the 2F5 and 4E10 peptides into a compact vertical stalk (as required by the Zanetti et al. cryo-ET density model) largely removes them from the lipid environment and is inconsistent with the proposed contact between the membrane and the cognate mAbs [4**,5**,46]. These two cryo-ET studies used somewhat different approaches for data collection and analysis. Subramaniam [47] has made an independent comparative assessment of these approaches.

**Conclusions**

Despite considerable recent advances in our understanding of the Env spike structure, still uncertain are the...
distribution and mobility of Env spikes on various forms of the viruses, the orientation of the monomeric core subunits within the trimeric spike, the orientation of the variable loops of SU and the CT of TM, some of the structural changes that occur during the binding and fusion processes, and the features that foster or inhibit Ab-mediated neutralization. Also unknown are the structural changes that result from the intracellular conversion of gp160 to gp41–gp120 and changes in structure or mobility that might accompany protease-mediated virus maturation following budding.

Recently described atomic structures of the SU core in the unliganded configuration and of the peptides comprising the MPER of TM have served to emphasize the conformational dynamics of the Env spike components following ligand-induced triggering. Interpretation of the functional significance of an atomic structure, such as that of the core peptides of the Env spike, is, however, fraught with uncertainty. Any sort of fragmented expression or even the production of otherwise intact soluble Env trimers dissociated from their membrane anchors and CTs will undoubtedly yet unpredictably relieve some of the internal tensions in this metastable assemblage, probably leading to deviation from the native structure [43]. Estimating or back-calculating to the original in situ spike structure might not be possible.

Cryo-ET is a viable approach for examining the structure of Env complexes in situ and in a near-native environment, and offers an opportunity to fit the available atomic structures into physiologically relevant orientations [48**]. The technology is rapidly evolving, as evident from a comparison of the latest (2005) comprehensive review [49***] to the most recently published applications. The visualization of individual molecular structures within a complex environment remains one of the most exciting and challenging aspects of cryo-ET. Application of the method to enveloped viruses and to less symmetric elements of icosahedral viruses is still at a very early stage of development. Methods of specimen preparation, data acquisition, and tilt series and subvolume alignment and averaging are rapidly evolving and should facilitate more detailed analyses in the future.

Acknowledgements
The authors wish to thank Ping Zhu for assistance with the figures, and the National Institute of Allergy and Infectious Disease for funding (grant number AI0155461).

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


The atomic structure of the core of unliganded gp120 is described and compared to earlier liganded structures, revealing large conformational changes upon CD4 binding. A tentative arrangement of the core within the Env trimer is proposed.


The gp120 core structure in complex with CD4 and MAb 17b Fab is presented. The V3 loop forms a β-hairpin structure and protrudes from the core peptide. When modeled as a trimer, the loops protrude toward the target cell, where they may interact with the CR.


The atomic structure of the 2F5 peptide of the MPER in complex with its cognate mAb reveals that both the peptide and the Ab probably make contact with, or are partially imbedded in, the viral membrane.


Like the 2F5 peptide, the atomic structure of the adjacent 4F10 peptide of the MPER in complex with its cognate mAb suggests that both the peptide and Ab probably make contact with, or partially protrude into, the viral membrane.


The authors designed a minimal SU fragment in association with gp41ECTO that stabilizes gp41ECTO in what might be the prefusion configuration and performed structural studies.


The crystal structure of the D5 mAb in complex with the therapeutic agent 5-helix is determined. Modeling demonstrates how this Ab can neutralize during the fusion intermediate stage through binding to the N-HR of the three-helix rod of gp41.


Cryo-ET was used to analyze the number and distribution of Env on HIV-1 and SIV virions, and to generate a 3D tomographic average of the SIV Env trimeric spike volume. The MPER appears to form an open trypod configuration. Relevant gp120 and gp41 atomic structures were provisionally fitted into the averaged spike volume.


This second cryo-ET report on SIV Env differs from that of Zhu et al. [12**] in that the gp120 heads are smaller and less complex, and a compact gp41 stalk, rather than tripod legs, is observed. Both studies used similar viruses, but somewhat different analytical techniques.
The authors investigated the steric accessibility of the N-HR during the fusion intermediate stage in a clever set of experiments in which molecules of varying sizes coupled to a C-peptide are used as probes.


The biophysical action of the FPs of the influenza and HIV-1 viruses on lipid membranes in a model system is compared.


In this comprehensive review, the relationships between gp120 structure, immunogenicity and Ab-mediated neutralization are discussed.


immunodeficiency virus envelope glycoproteins.


47. Subramaniam S: The SIV surface spike imaged by electron tomography: one leg or three? PLOS Pathogens 2006, 2:e91.

Recent advances in the technology of cryo-ET and its application to the analysis of viral structure are reviewed.


The best and most thorough recent review published on the technique of electron tomography of biological specimens. It is a great starting place for readers unfamiliar with the method.

