Ebola virus mechanism of infection

The Ebola virus (EBOV) is an enveloped, non-segmented, negative-strand RNA virus, which together with Marburg virus, makes up the *filoviridae* family. The virus causes severe hemorrhagic fever associated with 50-90% human mortality[1]. Four species of the virus (Zaire, Sudan, Côte d'Ivoire, and Reston ebolavirus) have thus far been identified, with Zaire typically associated with the highest human lethality[2]. A fifth EBOV species is confirmed in a 2007 outbreak in Bundibugyo, Uganda[3, 4]. Infection with EBOV results in uncontrolled viral replication and multiple organ failure with death occurring 6-9 days after onset of symptoms[5]. Fatal cases are associated with high viremia and defective immune responses, while survival is associated with early and vigorous humoral and cellular immune responses[6-9]. Although preliminary vaccine trials in primates have been highly successful[10-13], no vaccines, specific immunotherapeutics, or post-exposure treatments are currently approved for human use. Since 1994, EBOV outbreaks have increased more than four-fold, thus necessitating the urgent development of vaccines and therapeutics for use in the event of an intentional, accidental or natural EBOV release.

The EBOV genome contains seven genes, which direct the synthesis of eight proteins.

Transcriptional editing of the fourth gene (GP) results in expression of a 676-residue transmembrane-linked glycoprotein termed GP, as well as a 364-residue secreted glycoprotein termed sGP[14, 15]. EBOV GP is the main target for the design of vaccines and entry inhibitors.

GP is post-translationally cleaved by furin16 to yield disulfide-linked GP1 and GP2 subunits17. GP1 effects attachment to host cells, while GP2 mediates fusion of viral and host membranes16, 18-20. EBOV is thought to enter host cells through receptor-mediated endocytosis via clathrin-coated pits and caveolae21, followed by actin and microtubuledependent transport to the endosome21, where GP is further processed by endosomal cathepsins22-24. Essential cellular receptor(s) have not yet been identified, but DC-SIGN/LSIGN25, hMGL26, β-integrins27, folate receptor-α28 and Tyro3 family receptors29 have all been implicated as cellular factors in entry. Here, we report the crystal structure of EBOV GP, at 3. 4 Å resolution, in its trimeric, pre-fusion conformation in complex with neutralizing antibody Fab KZ52. GP1 is responsible for cell surface attachment, which is probably mediated by a region including residues 54-20132. GP1 is composed of a single domain (~65 Å × 30 Å × 30 Å), arranged in the topology shown in Fig. 2a, and can be further subdivided into the (I) base, (II) head and (III) glycan cap regions (Fig. 2a and Supplemental Fig. S3). The base (I) subdomain is composed of two sets of β sheets, forming a semi-circular surface which clamps the internal fusion loop and a helix of GP2 through hydrophobic interactions (Fig. 2b). Moreover, this subdomain contains Cys53, which is proposed to form an intermolecular disulfide bridge to Cys609 of the GP2 subunit17. Cys53 resides near GP2 in the β2-β3 loop at the viral membraneproximal end of the base subdomain (Fig. 2a-b). Our EBOV GP contains an intact GP1-GP2 disulfide bridge, based on reducing and non-reducing SDS-PAGE analysis. However, the region containing the counterpart GP2 cysteine is disordered, which may reflect functionally important mobility in the region. The head (II) is located between the base and glycan

cap regions towards the host membrane surface. Two intramolecular disulfide bonds stabilize the head subdomain and confirm the biochemically determined disulfide bridge assignments17.

Cys108-Cys135 connects a surface-exposed loop (β8-β9 loop) to strand β7, while Cys121-Cys147 bridges the β8-β9 and β9-β10 loops (Fig. 2a). The glycan cap (III) contains four predicted N-linked glycans (at N228, N238, N257 and N268) in an α/β dome over the GP1 head subdomain (Fig. 1b and 2a). This subdomain does not form any monomer-monomer contacts and is fully exposed on the upper and outer surface of the chalice. The central β sheets from the head and glycan cap together form a fairly flat surface and, in the context of the GP trimer, form the three inner sides of the chalice bowl.

Ebola virus GP2

GP2 is responsible for fusion of viral and host cell membranes and contains the internal fusion loop and the heptad repeat regions, HR1 and HR2. Many viral glycoproteins have fusion peptides, located at the N terminus of their fusion subunit, which are released upon cleavage of the precursor glycoprotein. By contrast, class II and class III fusion proteins, as well as class I glycoproteins from Ebola, Marburg, Lassa and avian sarcoma leukosis viruses, contain internal fusion loops lacking a free N terminus. The crystal structure reveals that the EBOV GP internal fusion loop, which encompasses residues 511-556, utilizes an antiparallel β stranded scaffold to display a partially helical hydrophobic fusion peptide (L529, W531, I532, P533, Y534 and F535) (Fig. 2c). The side chains of these hydrophobic residues pack into

a region on the GP1 head of a neighboring subunit in the trimer, reminiscent of the fusion peptide packing in the pre-fusion parainfluenza virus 5 F structure33. A disulfide bond between Cys511 at the base of β19 and Cys556 in the HR1 helix covalently links the antiparallel β sheet. This disulfide bond between the internal fusion loop and HR1 is conserved among all filoviruses, and is analogous to a pair of critical cysteines flanking the internal fusion loop in avian sarcoma leukosis virus34, 35. Interestingly, the EBOV internal fusion loop has features more similar to those observed in class II and III viral glycoproteins (in particular to flaviviruses) than those previously observed for class I glycoproteins (Supplemental Fig. S4). It thus appears that regardless of viral protein class, internal fusion loops share a common architecture for their fusion function.

EBOV GP2 contains two heptad repeat regions (HR1 and HR2), connected by a 25-residue linker containing a CX6CC motif and the internal fusion loop. The crystal structures of postfusion GP2 fragments30, 31 have revealed that the two heptad repeat regions form antiparallel α helices and that a CX6CC motif forms an intrasubunit disulfide bond between Cys601 and Cys608 (Supplemental Fig. S5). In the pre-fusion EBOV GP, HR2 and the CX6CC motif are disordered. By contrast, the HR1 region is well ordered and can be divided into four segments: HR1A, HR1B, HR1C and HR1D (Fig. 2c), which together assemble the cradle encircling GP1.

Similarly, heptad repeat regions in influenza and parainfluenza viruses also contain multiple segments in their pre-fusion helices that substantially rearrange in their post-fusion conformations33, 36, 37.

The first two segments, HR1A and HR1B (residues 554-575), together form an α helix with an ~40° kink at T565, which delineates HR1A from HR1B. Interestingly, the bend between HR1A and HR1B contains an unusual 3-4-4-3 stutter, which may act as a conformational switch31, rather than the typical 3-4 periodicity of heptad repeats (Supplemental Fig. S6). A similar stutter has also been noted in parainfluenza virus 5 F33. The Ebola virus HR1C (residues 576-582) forms an extended coil linking HR1B to the 14-residue α helix of HR1D (residues 583-598). HR1D forms an amphipathic helix and the hydrophobic faces of each HR1D join to form a three-helix bundle at the trimer interface. Although the breakpoint maps directly to a Lee et al. Page 3

Nature . Author manuscript; available in PMC 2009 June 22.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript chloride ion binding site in the post-fusion conformation of GP230, 31 and at least two other viruses38, 39, no chloride ion is observed here as HR1 and HR2 do not come together to form the six-helix bundle. Instead, the pre-fusion GP2 adopts a novel conformation, intimately curled around GP1 (Fig. 1c).

Ebola virus GP-KZ52 interface

KZ52 is an antibody isolated from a human survivor of a 1995 outbreak in Kikwit, Democratic Republic of the Congo (formerly Zaire)40. This antibody neutralizes *Zaire ebolavirus in vitro* 40 and offers protection from lethal EBOV challenge in rodent models41, but has minimal effects on viral pathogenicity in non-human primates42. KZ52 is directed towards a vulnerable, non-glycosylated epitope at the base of the GP chalice, where it

engages three discontinuous segments of EBOV GP: residues 42-43 at the N terminus of GP1, and 505-514 and 549-556 at the N terminus of GP2 (Fig. 3 and Supplemental Fig. S7). Although the majority of the GP surface buried by KZ52 belongs to GP2, the presence of both GP1 and GP2 are critical for KZ52 recognition43. It is likely that GP1 is required to maintain the proper pre-fusion conformation of GP2 for KZ52 binding. Indeed, KZ52 is the only antibody known to bridge both attachment (GP1) and fusion (GP2) subunits of any viral glycoprotein. Given that KZ52 requires a conformational epitope seen only in the GP2 pre-fusion conformation and that the KZ52 epitope is distant from the putative receptor-binding site (RBS), KZ52 likely neutralizes by preventing rearrangement of the GP2 HR1A/HR1B segments and blocking host membrane insertion of the internal fusion loop. Alternatively, IgG KZ52 may sterically hinder access to the RBS or to a separate binding site of another cellular factor, especially if multiple attachment events are required for entry.

The KZ52 epitope of GP is convex and does not have a high shape complementarity to the antibody (Sc index of 0. 63), although ~1600 Å2 of each GP monomer are occluded upon KZ52 binding. The antibody contacts a total of 15 GP residues by van der Waals interactions and 8 direct hydrogen bonds (Supplemental Fig. S7). Ten out of 15 residues in the structurally defined KZ52 epitope are unique to *Zaire ebolavirus* (Supplemental Fig. S6), thus explaining the Zaire specificity of KZ52.

Ebola virus GP glycosylation

We generated a fully glycosylated molecular model of EBOV GP to illustrate the native GP trimer as it exists on the viral surface (Fig. 4). The majority of N-linked glycosylation sites are concentrated in the mucin-like domain and glycan cap of GP1. Given that the mucin-like domain is ~75 kDa in mass (protein and oligosaccharide), the volume of this domain is predicted to be similar to each GP monomer observed here. The crystal structure suggests that the mucin-like domain is linked to the side of each monomer and may further build up the walls of the chalice, forming a deeper bowl (Fig. 4). Although a mixture of complex, oligomannose and hybrid-type glycans are found on intact, mucin-containing GP144, those glycans outside the mucin-like domain are likely to be complex in nature: the mucin-deleted GP used for crystallization is sensitive to PNGaseF, but not to EndoH treatment (Supplemental Fig. S8).

Modeling of complex-type oligosaccharides on the EBOV GP indicates that the majority of the GP trimer is cloaked by a thick layer of oligosaccharide, even without the mucin-like domain (Fig. 4). The ~19 additional oligosaccharides on the full-length GP (17 on the mucinlike domain and 2 more on GP1, disordered here) further conceal the sides and top of the chalice. The KZ52 binding site and, presumably, the flexible regions of HR2 and the membrane-proximal external region (MPER) remain exposed and perhaps vulnerable to binding of antibodies and inhibitors.

Lee The development of neutralizing antibodies is limited in natural Ebola virus infection. Many survivors have low or insignificant titres1, 7, and those antibodies that are elicited preferentially recognize a secreted version of the viral glycoprotein that features an alternate quaternary structure and lacks

the mucin-like domain43. The glycocalyx surrounding EBOV GP likely forms a shield that protects it from humoral immune responses and/or confers stability insideor outside a host. The mucin-like domain and glycan cap sit together as an external domain to the viral attachment and fusion subunits, reminiscent of the glycan shields of HIV-1 gp12045, 46 and Epstein-Barr virus gp35047, perhaps pointing to a common theme for immune evasion. Alignment of filoviral sequences indicate that regions involved in immune evasion have a low degree of sequence conservation [i. e. GP1 glycan cap (~5%) and mucin-like domain (0%)], but the N-glycosylation sites in the glycan cap are mostly conserved among all EBOV subtypes (Supplemental Fig. S6), indicating the functional importance of these posttranslational modifications.

Sites of receptor binding and cathepsin cleavage

Although a definitive receptor for EBOV remains to be identified, previous studies32, 48, 49 have determined that residues 54-201, which map to the base and head subdomains of GP1, form a putative receptor-binding site (RBS) for attachment to host cells. Additional experimental studies have identified at least 19 GP1 residues, assigned into four groups based on the location in the structure, that are critical for viral entry48-50 (Fig. 5). Many of these residues are apolar or aromatic and are involved in maintaining the structural integrity of GP1 for receptor binding or fusion. However, six residues (K114, K115, K140, G143, P146 and C147) cluster within a ~20 × 15 Å surface in the inner bowl of the chalice and may thus represent important receptor contact sites. All residues in the putative RBS are highly conserved among Ebola virus species (Supplemental Fig. S6).

Importantly, this putative RBS is recessed beneath the glycan cap and perhaps further masked by the mucin-like domain (Fig. 4), suggesting that additional conformational change or removalof the mucin-like domain could reveal additional surfaces required for receptor or cofactor binding. It has been demonstrated that endosomal proteolysis of EBOV GP by cathepsin L and/or B removes the mucin-like domain to produce a stable ~18 kDa GP1 intermediate which has enhanced viral binding and infectivity[22-24]. The precise site of cathepsin cleavage is unknown and the role of cathepsins in natural infection is as yet unclear. However, formation of an ~18 kDa GP1 fragment implies that cathepsin may cleave near the GP1 β13-β14 loop (residues 190-213). Indeed, this loop is unresolved in the pre-fusion structure, suggesting enhanced mobility and accessibility to enzymatic cleavage. Cleavage within this loop would remove the entire mucin-like domain and glycan cap region (Fig. 5). As a result, β7 to β9 strands and their associated loops would become exposed. These regions of GP are in proximity to the previously identified residues critical for viral entry. The fold, location and physicochemical properties of this site should now provide new leads in the search for the elusive filoviral receptor(s).

A summary of the Ebola virus mechanism of infection, including the events of cathepsin cleavage and conformational changes to GP2 during fusion, is presented.