Molecular Cell, Vol. 2, 605–616, November, 1998, Copyright ©1998 by Cell Press

# Crystal Structure of the Ebola Virus Membrane Fusion Subunit, GP2, from the Envelope Glycoprotein Ectodomain

**Winfried Weissenhorn,*†# Andrea Carfí,*# Kon-Ho Lee,* John J. Skehel,§ and Don C. Wiley*†‡‖**
*Laboratory of Molecular Medicine
Howard Hughes Medical Institute
The Children's Hospital
320 Longwood Avenue
Boston, Massachusetts 02115
‡Department of Molecular and Cellular Biology
Howard Hughes Medical Institute
Harvard University
Cambridge, Massachusetts 02138
§National Institute for Medical Research
The Ridgeway
Mill Hill
London, NW7 1AA
United Kingdom

## Summary

**We have determined the structure of GP2 from the Ebola virus membrane fusion glycoprotein by X-ray crystallography. The molecule contains a central triple-stranded coiled coil followed by a disulfide-bonded loop homologous to an immunosuppressive sequence in retroviral glycoproteins, which reverses the chain direction and connects to an α helix packed antiparallel to the core helices. The structure suggests that fusion peptides near the N termini form disulfide-bonded loops at one end of the molecule and that the C-terminal membrane anchors are at the same end. In this conformation, GP2 could both bridge two membranes and facilitate their apposition to initiate membrane fusion. We also find a heptad irregularity like that in low-pH-induced influenza HA2 and a solvent ion trapped in a coiled coil like that in retroviral TMs.**

## Introduction

Infections of humans with the filoviruses, Ebola virus, and Marburg virus are rare, but the resulting hemagorrhagic fevers are associated with high mortality (Siegret et al., 1967; Sanchez et al., 1995, 1998a). For example, 231 people died in an outbreak involving 300 infections of Ebola in Zaire in 1995 (Centers for Disease Control and Prevention, 1995; Sanchez et al., 1995). Four subtypes of Ebola viruses, named Zaire, Sudan, Ivory Coast, and Reston, have been identified by genome sequencing (Sanchez et al., 1996). Their single envelope glycoproteins (GP) have about 60% amino acid sequence identity; the envelope glycoprotein of Marburg virus shares only about 30% amino acid sequence identity with them (Sanchez et al., 1993). The glycoproteins mediate entry into target cells such as liver cells and endothelial cells (Schnittler et al., 1993; Becker et al., 1995; Yang et al., 1998).

Like other virus membrane glycoproteins active in viral entry by membrane fusion such as those of the myxoviruses, paramyxoviruses, and retroviruses, Ebola virus GPs are synthesized as single-chain precursors and transferred cotranslationally into the lumen of the endoplasmic reticulum where they form trimers (Wiley et al., 1977; Hunter and Swanstrom, 1990; Feldmann et al., 1991; Joshi et al., 1998; Sanchez et al., 1998a). They are then posttranslationally cleaved into two chains, GP1 and GP2, analogous to HA1 and HA2 of influenza, gp120 and gp41 of HIV-1, SU and TM of retroviruses, and F2 and F1 of paramyxoviruses (reviewed in Klenk et al., 1993; Volchkov et al., 1998). This cleavage is required to prime the glycoproteins for subsequent activation of their membrane fusion function required for infectivity (Scheid and Choppin, 1974). In influenza virus, the primed, cleaved molecules found on the viral surface are apparently in a metastable conformation, kinetically blocked behind a high activation barrier from their membrane fusion–active conformation. Activation energy to generate membrane fusion activity and to form a stable conformation is supplied by the low pH of endosomes for influenza HA. There is evidence for conformational changes triggered by receptor binding for the glycoproteins of retroviruses such as avian leukosis sarcoma virus (ALSV), and by receptor and coreceptor binding for HIV-1 envelope glycoprotein (Skehel et al., 1982; D'Souza and Harden, 1996; Herdandez et al., 1997), but how filovirus membrane fusion is activated is not known.

In the case of influenza, three conformations of the hemagglutinin have been determined by X-ray crystallography of soluble ectodomains derived by proteolysis of purified virions or infected cell membranes: the uncleaved precursor, HA0; the cleaved, metastable conformation found on infectious virions, HA1-HA2; and the stable conformation formed at fusion pH, TBHA2 (Wiley et al., 1981; Wilson et al., 1981; Bullough et al., 1994; Chen et al., 1998a). In addition, HA2 alone, expressed in bacteria in the absence of the receptor-binding domain, HA1, and lacking its N-terminal fusion peptide, was shown to fold spontaneously at neutral pH into the low-pH-induced rod-shaped conformation (Chen et al., 1995, 1998b; Carr et al., 1997). It has been reasoned that the analogous ectodomains of Moloney murine leukemia virus (Mo-55) (Fass and Kim, 1995; Fass et al., 1996) and HIV-1 (gp41) (Weissenhorn et al., 1996), and core fragments of gp41 (Lu et al., 1995; Chan et al., 1997; Tan et al., 1997; Weissenhorn et al., 1997a, 1997b) and SV5 (Joshi et al., 1998) also adopt their lowest energy states spontaneously when expressed in the absence of the receptor containing subunits and fusion peptides.

We have previously shown that the fusion subunit ectodomain of Ebola, GP2, folds spontaneously into a trimeric, highly α-helical, rod-shaped conformation when expressed in bacteria with a trimeric, isoleucine zipper derived from GCN4 (Harbury et al., 1993) in place of the N-terminal fusion peptide (Weissenhorn et al., 1998a). Here we report the determination of the three-dimensional structure of the Ebola virus GP2 to 3.0 Å resolution

‖To whom correspondence should be addressed (e-mail: dcwadmin@crystal.harvard.edu).
#These authors contributed equally to this work.

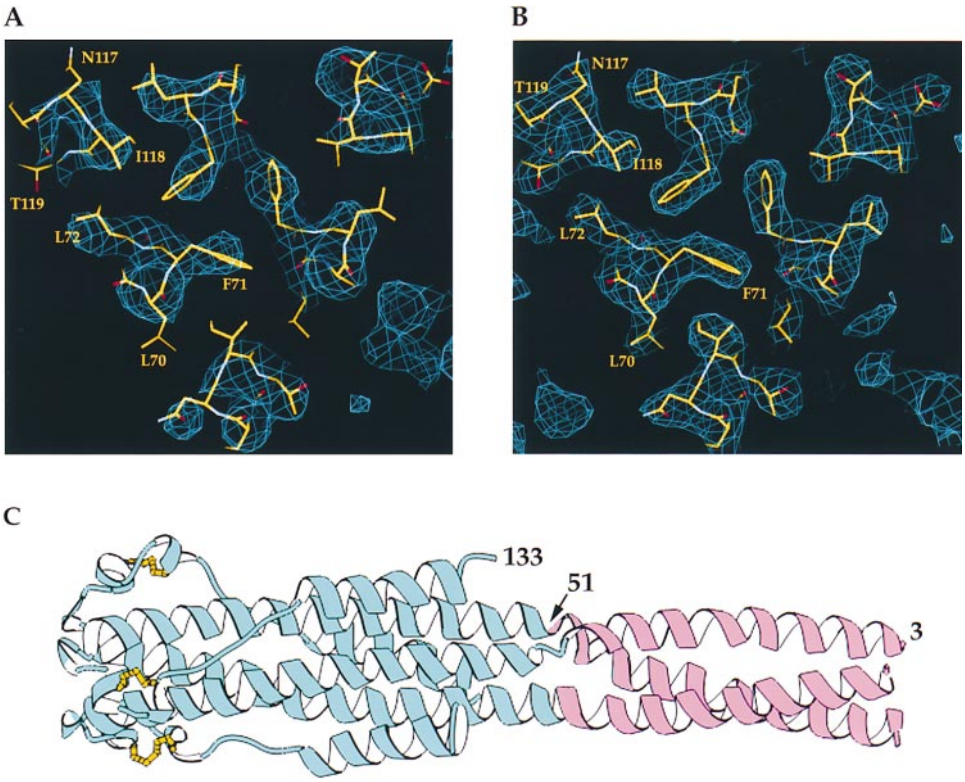


Figure 1. X-Ray Crystallographic Structure Determination

(A) Experimental electron density map calculated with single isomorphous replacement anomalous scattering (SIRAS) phases to 3.5 Å resolution (Table 1B), showing GP2 Phe-71 packed at the core of the coiled coil.

(B) Density-modified electron density map calculated after iterative real-space NCS three-fold averaging, histogram matching, and solvent flattening to improve and extend the phases to 3.0 Å resolution.

(C) Ribbon diagram of the GP2 (blue)/modified GCN4 (purple) hybrid trimer. The disulfide bond in GP2 is colored yellow. The N terminus of the GCN4 trimer is at the far right, and the C terminus of GP2 is after the outer helical layer at the center. (A) and (B) were created with O (Jones et al., 1991), (C) with Ribbons (Carson, 1991).

by X-ray crystallography. The structure shows that a disulfide-bonded loop, proposed in retroviruses to have immunosuppressive properties, connects an inner trimeric core of N-terminal $\alpha$ helices to an outer layer composed of an antiparallel, linear strand-helix-strand segment (Figure 1). We compare the structure of GP2 to those of other viral and cellular membrane fusion molecules and to a structure predicted for GP2 (Gallaher, 1996). The comparison suggests that the observed structure may be a model for both a prefusion intermediate and the postfusion conformation of GP2.

## Results and Discussion

### Structure Determination

Ebola Zaire GP2 is 179 residues long; it has a 16-residue "internal fusion peptide," residues 23–38, and a 25-residue membrane anchor, residues 150–175 (purple in Figure 2A) (Feldmann et al., 1994; Sanchez et al., 1998a). Residues 51–149 were expressed in *E. coli* with a 32-residue modified GCN4 isoleucine zipper known to trimerize (Harbury et al., 1993), replacing the N-terminal fusion peptide loop residues 1–50 as described earlier (Weissenhorn et al., 1998a). Heptads in the GCN4 zipper were placed in register with predicted heptads in the N-terminal region of GP2 (Gallaher, 1996). Free cysteines at residues 55 and 108 were substituted with glutamic acid and arginine, respectively, to minimize aggregation and to promote crystallization. Arg-108 forms a crystal lattice contact. Crystals contain two trimers of the GCN4-GP2 chimera per asymmetric unit. To generate an isomorphous heavy atom derivative, GCN4 zipper residue Leu-12c, known to be on the outside surface of the trimer, was substituted with a cysteine that was reacted with p-hydroxy-mercuribenzoate before crystallization. The structure was determined by single isomorphous replacement with anomalous scattering (SIRAS) (Figure 1A) and density modification including iterative real-space noncrystallographic averaging about the molecular three-fold symmetry axes and solvent flipping (Figure 1B). (Data and phasing statistics are presented in Tables 1A and 1B and in the Experimental Procedures.) The refined model consists of residues 3c–32c of the modified GCN4 isoleucine zipper (purple in Figure 1C) and residues 51–133 of one GP2 monomer (light blue in Figure 1C); the other monomers end at residues 131 and 132. Refinement statistics are presented in Table 1C.

### Monomer Structure and a Conserved Interhelical Linker

The GP2 ectodomain monomer is a helical hairpin about 65 Å long. A 44-residue N-terminal $\alpha$ helix (residues

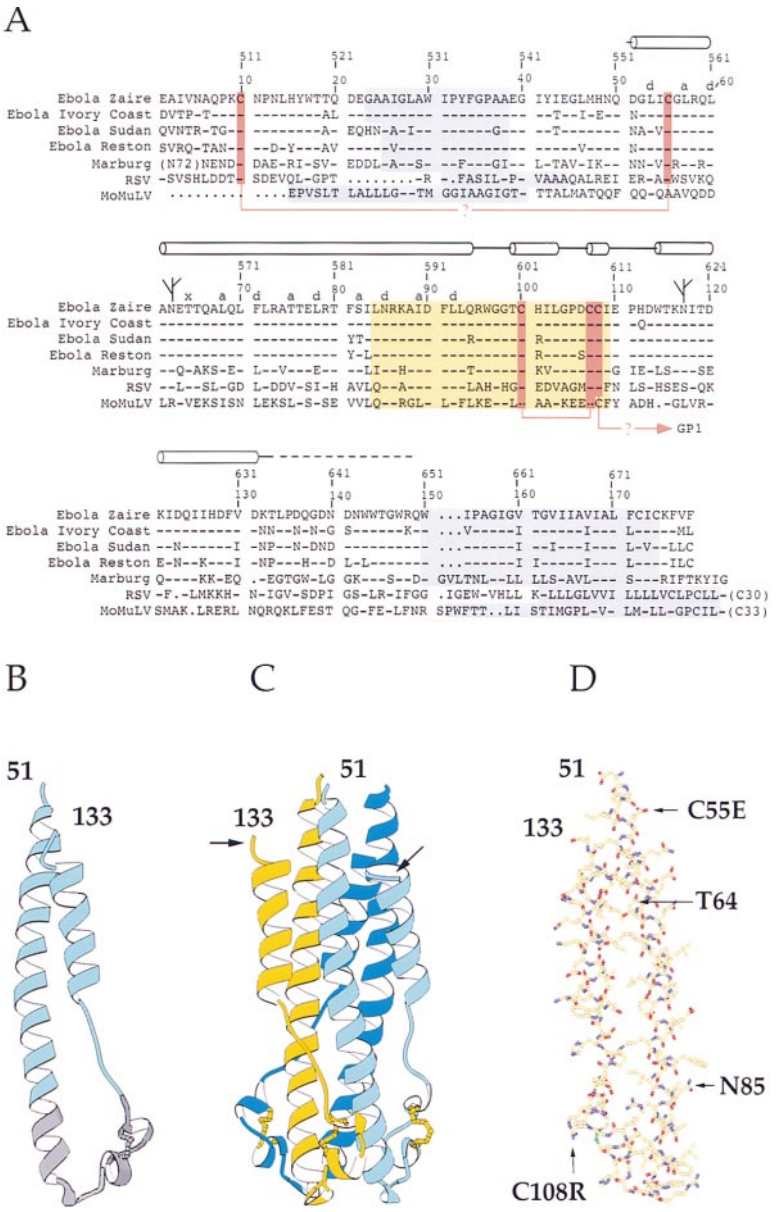


Figure 2. Sequence and Structure of Ebola Virus Zaire GP2

(A) Sequence and secondary structure (cylinders, helices; heptads, a, d, and x; dashed lines, disordered) of Ebola Zaire GP2 aligned with sequences of the other three Ebola subtypes, Marburg virus GP2, and two retrovirus Tms (RSV, Rous sarcoma virus; MoMulv, Moloney murine leukemia virus). Cysteines are marked red. The N-terminal fusion peptide and C-terminal transmembrane anchors are highlighted in purple, and the conserved "immunosuppressive motif" containing a disulfide-bonded loop and free cysteine in yellow. The disulfide bonds between Cys-10 and Cys-55 and Cys-108 and GP1 have not been demonstrated biochemically or directly in the current structure but have been suggested previously (Sanchez et al., 1998b), and their assignment is discussed in the text.

(B) Ribbon drawing of the Ebola GP2 monomer. The N-terminal residue 51 and last structured residue 155 are labeled.

(C) Ribbon drawing of the trimer of Ebola GP2. Arrows points to the residues 132–133, which are ordered but slightly different in different monomers (see text). The residues beyond 133 (to 149) are disordered. The disulfide bonds are shown as ball and stick models and colored yellow.

(D) All-atom model of the Ebola GP2 monomer. Mutated residues C55E and C108R are indicated with arrows. The side chain of Thr-64, which packs irregularly in the core coiled coil, and Asn-85, which is bound to the tentative chloride ion, are labeled.

(B), (C), and (D) were created with Ribbons (Carson, 1991).

51–94) is followed by a 19-residue link to a short, antiparallel $\alpha$ helix of 28 residues (residues 114–131) (Figure 2B). Beyond the outer layer $\alpha$ helix, two monomers extend by 1 and 2 residues, respectively, while the remaining 16 (17) residues are disordered (arrow in Figure 2C), as well as all 18 residues of the third monomer.

The interhelical linker, residues 95–113, is part of a segment, residues 84–109, containing a disulfide-bonded loop (Cys-100 to Cys-107); the whole segment shares 40%–50% sequence identity with an immunosuppressive sequence found in many retrovirus TM subunits (yellow in Figure 2A and gray in Figure 2B) (Volchkov et al., 1992). The conformation of this segment, which comprises the two C-terminal turns of the long $\alpha$ helix and a four-residue turn followed by a short $\alpha$ helix that leads into an eight-residue disulfide-bonded loop (Figure 3A), is very similar (1.5 Å rms) to that of the homologous sequence in the structure of a 55-residue segment of the Murine Moloney Leukemia virus (MuMoLv) TM subunit, Mo-55 (Figure 3A), despite the truncation of Mo-55 immediately C-terminal to the sequence for crystallization (Fass et al., 1996).

The structural conservation between filoviruses and retroviruses of this complex motif suggests that it may have been selected to serve a more complex function than merely reversing the chain direction (Figure 3A). It seems unlikely to have been conserved to have an immunosuppressive effect because that effect can be demonstrated using peptides homologous to the C terminus of the core helix, less than half the length of the immunosuppressive motif (Cianciolo et al., 1985; Becker, 1996). Similarly, such an elaborate structure also seems unnecessary simply to link two antiparallel $\alpha$ helices, a task that can be accomplished by four residues. It may be that both the linker and disulfide-bonded loop were conserved to function as a hinge that can

Table 1. Crystallographic Statistics

A. Data Collection

| | Native | | PHMB[a] | |
|---|---|---|---|---|
| | 20.0–3.0 | 3.09–3.0 | 20.0–3.5 | 3.6–3.5 |
| $R_{merge}$[b] | 0.083 | 0.29 | 0.10 | 0.23 |
| Complete (%) | 99.3 | 99.2 | 99.3 | 98.8 |
| $>3\sigma_I$ (%) | 75.0 | 48.2 | 74.4 | 56.1 |
| No. reflections | 87,380 | 4133 | 66,460 | 2745 |
| No. unique reflections | 17,123 | 1345 | 10,870 | 888 |
| Average redundancy | 5.1 | 3.07 | 6.1 | 3.0 |
| $R_{iso}$[c] | | | 0.18 | 0.20 |

B. Phasing (SHARP)

| | Phasing Power[d] | | | | Figure of Merit | |
|---|---|---|---|---|---|---|
| Resolution | Acentric | $Acentric_{anom}$ | Centric | Resolution | Acentric | Centric |
| 20–9.95 | 2.37 | 1.2 | 1.2 | 20–7.88 | 0.54 | 0.41 |
| 9.95–6.56 | 3.30 | 0.98 | 2.30 | 7.88–5.8 | 0.55 | 0.54 |
| 6.56–5.22 | 2.68 | 0.82 | 1.72 | 5.8–4.81 | 0.46 | 0.47 |
| 5.22–4.46 | 1.73 | 0.65 | 1.31 | 4.81–4.19 | 0.36 | 0.39 |
| 4.46–3.96 | 1.78 | 0.56 | 1.24 | 4.19–3.77 | 0.36 | 0.35 |
| 3.96–3.5 | 1.78 | 0.46 | 1.18 | 3.77–3.5 | 0.25 | 0.24 |
| 20–3.5 | 2.16 | 0.71 | 1.51 | 20–3.5 | 0.42 | 0.4 |

C. Refinement

| R Values and Temperature Factors | | Model Geometry | |
|---|---|---|---|
| No. reflections in working set | 16,401 | Bond length rmsd from ideal | 0.016 Å |
| $R_{cryst}$[e] | 0.239 (0.332)[f] | Bond angles rmsd from ideal | 1.54° |
| $R_{free}$[e] | 0.256 (0.337)[f] | Ramachandran plot | |
| | | % in most favored region[g] | 93.0 |
| Average B | 38.7 $Å^2$ | % in additional allowed regions[g] | 6.7 |

[a] p-hydroxy-mercuribenzoate.
[b] $R_{merge} = \Sigma_h\Sigma_i| I_i(h) - <I(h>)|/\Sigma_h\Sigma_i I_i(h)$, where $I_i(h)$ is the $i^{th}$ measurement, and $<I(h)>$ is the weighted mean of all measurements of I(h).
[c] $R_{iso} = \Sigma_h | |F_{EMP}(h)| - |F_{native}(h)||/\Sigma_h|F_{native}(h)|$.
[d] Phasing power = $<|F_H|>/E$, where $<|F_H|$ is the rms structure factor amplitude for the heavy atom, and E is the estimated lack-of-closure error.
[e] $R_{cryst}$ and $R_{free} = \Sigma_h||F(h)_{obs}| - |F(h)_{calc}||/\Sigma_h|F(h)_{obs}|$ for reflections in the working and test sets, respectively.
[f] Numbers in parentheses are for final shell 3.07–3.00 Å.
[g] As defined in PROCHECK (Laskowski et al., 1993).

change the direction of the sequences C-terminal to it. The hypothetical disulfide-bonded attachment of GP1 (and some retroviral Su subunits) to such a hinge (Gallaher, 1996; Sanchez et al., 1998b) (Figure 5B) would provide a possible path for information to transfer from receptor binding by GP1 to trigger a conformational change in GP2. There is evidence for contacts between gp120 and gp41 at the homologous loop in gp41 and for conformational changes of that region after receptor and coreceptor binding (Sattentau and Moore, 1991; Weissenhorn et al., 1996), but there are no studies to date of the conformational changes that accompany activation of filoviral membrane fusion.

### A Stutter in the Trimeric Coiled Coil Like that in Low-pH-Induced HA2

The core of the trimeric structure is a triple-stranded $\alpha$-helical coiled coil formed by the 12-turn-long, N-terminal $\alpha$ helix (residues 51–94) (Figure 2C). Near the N-terminal end, the coil is underwound, at a stutter (Brown et al., 1996), relative to classical knobs-into-holes packing predicted by Crick (Crick, 1953). Seven hydrophobic and four polar residues are packed in layers at the core of the coiled coil, generally with a heptad (3-4) periodicity, where heptad positions are denoted a–g (Figure 2A). At one position, fourth from the N terminus, an unusual 3-4-4-3 periodicity is observed (Figures 2A and 2B), caused by the packing of Thr-64 directed at the three-fold axis ("x-like" packing; Figure 3B) (see also Bullough et al., 1994; Brown et al., 1996). Low-pH-treated HA2 has a similar 3-4-4-3 periodicity near the beginning of the central helix at Thr-59, a residue that is near the beginning of an interhelical loop in metastable HA but in the coiled coil in the low-pH-treated HA2 structure (in Bullough et al., 1994; see Figures 3A and 5). It is possible by analogy that in a hypothetical metastable GP1-GP2 structure, GP2 may have a nonhelical segment between $\alpha$ helices starting near Thr-59. (A similar irregularity is found in the structure of HTLV-1 TM [B. Kobe, R. J. Center, B. E. Kemp, and P. Poumbourious, personal communication].)

A practical consequence of this unanticipated break in the 3-4 periodicity in GP2 is that we attached the GCN4 isoleucine zipper with its heptads out of phase with the actual, as opposed to predicted (Gallaher, 1996), heptads. (The remainder of the heptad prediction

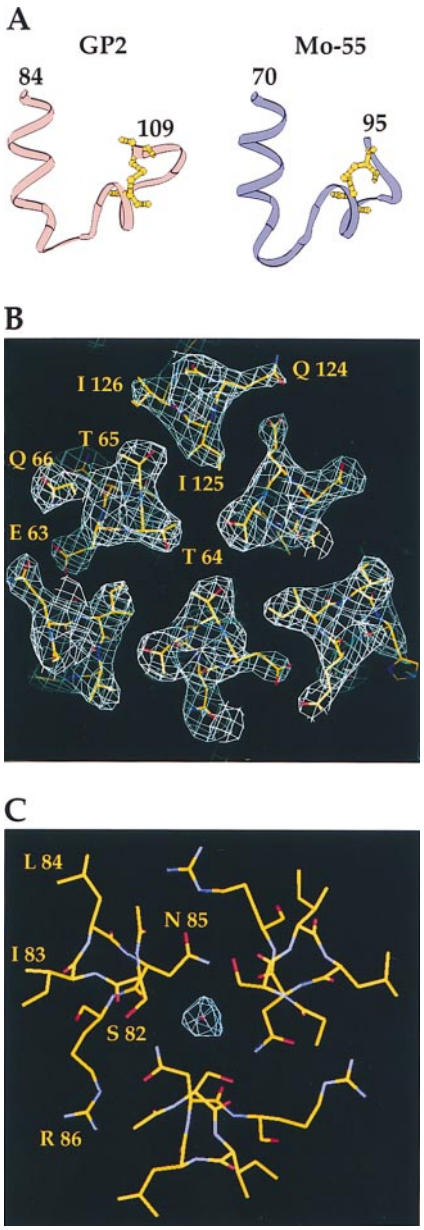


Figure 3. A Conserved Motif, a Coiled Coil Stutter, and an Ion Binding Site in GP2

(A) Comparison of the helix-turn-disulfide-bonded-loop motif of Ebola GP2 and MoMuLv Mo-55. Residues Leu-84 to Ile-109 of GP2 (yellow in Figure 2A) are shown in pink, Gln-70 to Phe-95 of Mo-55 in purple. The disulfide bond is drawn as a ball and stick model and colored yellow.

(B) The "x-like" packing of Thr-64 pointing toward the three-fold symmetry axis rather than like classical knobs-into-holes packing (e.g., Figure 1C). Electron density calculated with phases based on the refined model (Table 1C) is superimposed on the model.

(C) Electron density thought to represent a chloride ion trapped at the center of the coiled coil between Asn-85 and Ser −82. The distances between the amide nitrogen of Asn-85 and the tentative chloride ion average 3.5 Å, similar to the 3.25 Å distances observed between a chloride ion and Asn-71 (the sequence homolog of GP2 Asn-85) in Mo-55 (Fass et al., 1996).

(A) was created with Ribbons (Carson, 1991), (B) and (C) with O (Jones et al., 1991).

for the N-terminal helix is correct [Gallaher, 1996].) In the crystals, we observe the GCN4 trimer axis to make an angle of a few degrees with the GP2 trimer axis (Figure 1C) in both copies of the molecule in the asymmetric unit. The independence of the structures of the GCN4 trimer and the GP2 trimer suggests that the zipper had its intended purpose of increasing the effective molarity of the GP2 chains so that they would fold stably and of helping to solubilize the molecule by replacing the fusion peptide (Weissenhorn et al., 1998a), without imposing the regular 3-4 periodicity of the GCN4 trimer.

### A Possible Chloride Ion Trapped in the Coiled Coil as in Mo-55

Between the successive layers formed by Ser-82 and Asn-85 in the center of the coiled coil, we observe an electron dense peak from the solvent (5.9 $\sigma$, B = 39.5 Å$^2$) (Figure 3C). A similar feature was observed buried by the three Asn-71 amine groups of the Mo-55 trimer (Fass et al., 1996), which is the homolog of Asn-85 (Figure 2D) in the sequence to GP2 and is conserved in a number of sequences of retroviruses (e.g., Figure 4 in Bukreyev et al. [1993]), suggesting that the trapped chloride ion is a common feature in retroviral coiled coils. In the case of Mo-55, stability measurements of wild-type and an Asn-71 to Ile mutant in the presence of a series of anions indicated that the electron dense peak was due to a chloride ion. The crystalline GP2 has been maintained in buffers containing chloride from expression through crystallization. Based on the conservation of the Asn in retroviral and filoviral sequences (Figure 2A), and on the similarity of the electron density and coordination of the solvent peak in GP2 to that of the chloride in Mo-55, it is possible that GP2 also contains a chloride ion trapped at the homologous location near the C-terminal end of the central coiled coil. A solvent molecule thought to be an ion is also observed in the N-terminal part of the triple helical bundle at the core of the influenza HEF metastable conformation (Rosenthal et al., 1998). It has been suggested that the chloride ion in Mo-55, like other buried polar interactions in coiled coils (Harbury et al., 1993; Lumb and Kim, 1995) that promote specificity in their folding, might prove important in MoMuLv (Fass et al., 1996). Since the chloride is located at the beginning of the sequence segment (84–109; yellow in Figure 2A) that is most conserved between filoviral GP2s and retroviral TMs (see above), we speculate that it may be conserved for some function, possibly during conformational switching.

### A Hydrophobic Core at the Base of the Coiled Coil

Packed against the core coiled coil at its C terminus are residues forming a turn and the eight-residue disulfide bond–containing loop (Figures 3A and 2C), which together form a small hydrophobic core (residues F91, L93, W96, I102, and I103) like that at the base of the coiled coil in the low-pH-induced conformation of influenza HA (Figure 4C; Figure 4 in Bullough et al., [1994]) and at the homologous location in Mo-55 (Fass et al., 1996) (Figures 4A and 4B). The interhelical loop is missing from the X-ray structures of the proteolytically resistant core fragments of HIV-1 gp41 (Figure 4D). In an

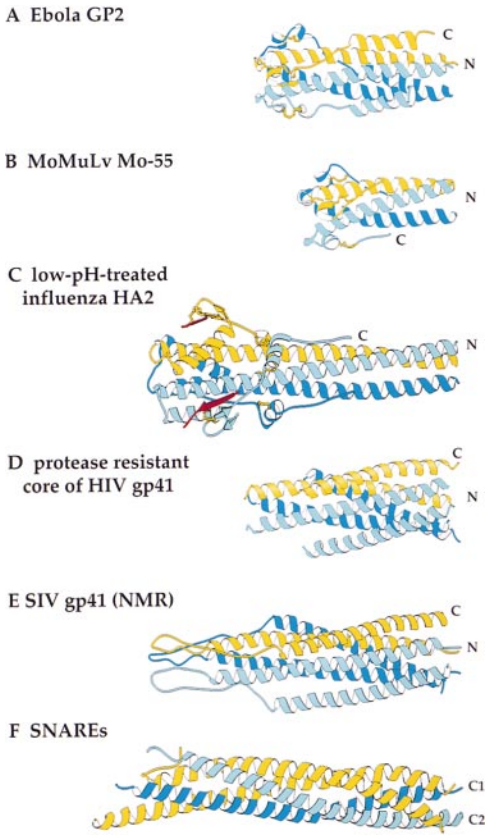


Figure 4. Comparison of GP2 with the Structures of Viral and Cellular Membrane Fusion Proteins

(A) Recombinant Ebola Zaire GP2.
(B) Recombinant Mo-55 from the TM subunit of MoMuLv (Fass et al., 1996).
(C) Low-pH-treated HA2 from influenza virus (Bullough et al., 1994).
(D) Recombinant, proteolysis-resistant core of HIV-1 gp41 (Weissenhorn et al., 1997b).
(E) Recombinant SIV gp41, NMR structure (Caffrey et al., 1998).
(F) Recombinant core coiled segments of the SNARES syntaxin 1-A (blue), synaptobrevin-II (light blue), and SNAP-25B (yellow) (Sutton et al., 1998). This figure was created with RIBBONS (Carson, 1991).

NMR structure of an ectodomain of SIV gp41, this loop extends beyond the coiled coil, turning at the apex of a disulfide-bonded loop before returning antiparallel to the outer layer $\alpha$ helix (Caffrey et al., 1998) (Figure 4E).

### An Outer Layer $\alpha$ Helix

The remaining length of the core coiled coil is stabilized by the packing of a linear strand-helix-strand segment into the groove between the core $\alpha$ helices (Figures 2C and 5A). The first extended strand of this segment, Cys-108 to Asp-113, makes nonpolar contacts into this groove with Ile-109 and Pro-111 and reaches to one core $\alpha$ helix with His-112 to make a salt bridge with Glu-77 (Figure 5A). A short, 4 1/2–turn $\alpha$ helix (residues Trp-114 to Asp-131) then packs antiparallel into the groove with nearly classical knobs-into-holes packing of the nonpolar side chains Trp-114, Ile-118, Ile-122, Ile-125, and Phe-129 (Figure 5A). This outer layer $\alpha$ helix is two turns shorter in this structure of a GCN4-GP2 hybrid than predicted in a model of GP2, although the nonpolar packing is in phase with the predicted heptads (Gallaher, 1996). (We cannot rule out the unlikely possibility that the GCN4 trimer which begins 13 Å further along the core coiled coil might have influenced the length of the outer layer $\alpha$ helix).

### A Flexible Link to the Transmembrane Anchor

Of the 19 residues connecting the outer layer $\alpha$ helix to the transmembrane anchor, only the first one or two residues per monomer (Lys-132 and Thr-133 [built as Ala]; arrow in Figure 2C) are visible in the electron density maps. The two structured residues pack along the groove between the core $\alpha$ helices of the coiled coil (Figure 5A). The C-terminal-most structured segment of the low-pH-induced conformation of influenza HA2 is an extended strand of 11 residues, 152–162, which was also observed to adopt two conformations, ordered in two monomers and disordered in one monomer, in the crystallized trimer (at C label in Figure 4C) (Bullough et al., 1994). The observation that the final structured residues of GP2 and HA2 take different paths in the different copies of the monomers in crystals suggests that these residues can adopt alternate conformations of approximately the same energetic stability.

One feature shared by the structures of GP2, low-pH-treated HA2, the proteolyticly resistant core of recombinant gp41, and the ectodomain of SIV gp41 is a long (13- to 18-residue) disordered segment between the last structured residue (labeled C in Figures 4A, 4C, 4D, and 4E) and the transmembrane anchor sequence, either not visible, although present, in the crystals as in GP2 (15 residues, 134–149) and HA2 (13 residues, 163–175 in crystal; anchor starts at 185), or inferred from its proteolytic susceptibility in gp41 (18 residues, 155–172). In HA2, these 23 disordered residues were proposed to provide a flexible link from the rod-shaped molecule to the virus membrane that would allow the opposed membranes to approach more closely than the length of the rod during fusion and allow HA2 to have both its fusion peptide and C-terminal anchor in the same membrane postfusion (Bullough et al., 1994). Immunoelectron microscopy of low-pH-treated HA in virosomes subsequently demonstrated that, in the absence of target membranes, association of HA rods with the same membrane by both the N-terminal fusion peptide and the C-terminal anchor was permitted by the flexible segment at the C terminus of the HA2 ectodomain (Wharton et al., 1995).

### A Hypothetical Model for the Full-Length GP2 Ectodomain

Based on the X-ray structure of the GP2 ectodomain without the fusion peptide, some aspects of the full-length GP2 and its attachment to GP1 can be suggested. The structure establishes that two of the three cysteines, Cys-55 and Cys-108, are too far apart to form a disulfide bond. This implies, as suggested previously but not established biochemically (Gallaher, 1996; Sanchez et al., 1998a), that Cys-10, which is N-terminal to the fusion peptide, will form a disulfide bond with Cys-55 near the top of the N-terminal $\alpha$ helix (Figure 5B). These two cysteines are conserved as a pair in filoviral and some

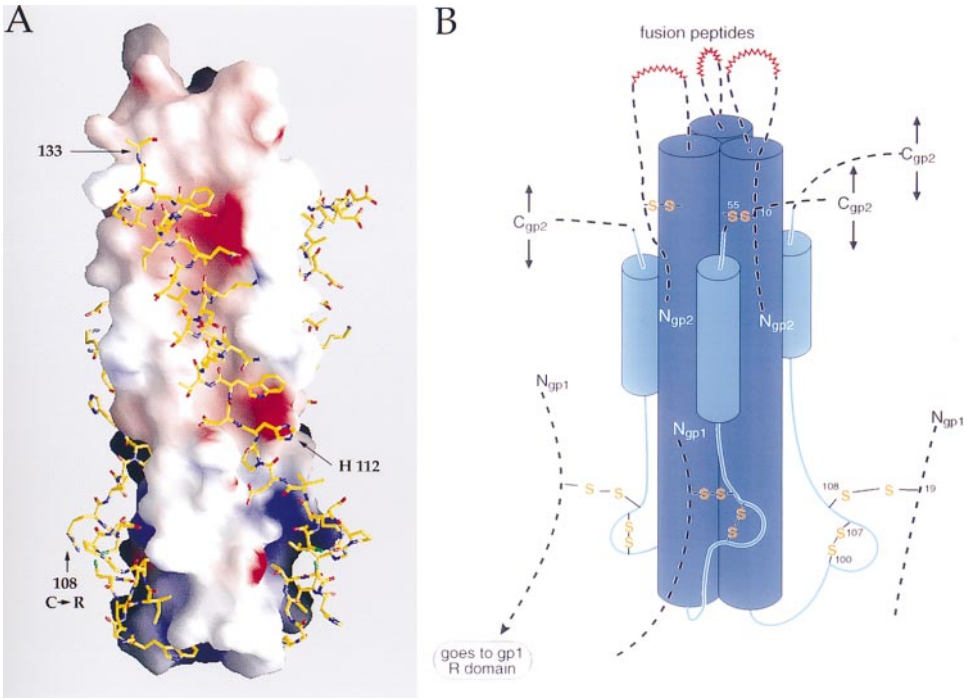


Figure 5. The Packing of the Outer Layer of the GP2 Ectodomain Trimer and a Hypothetical Model for the Whole Molecule

(A) GP2 ectodomain with an electrostatic surface potential representation of the core coiled coil and with the outer layer drawn as an atomic model. Negative electrostatic potential, red; positive, blue. The position of His-112, which makes a salt bridge to Glu-77 of the inner core, and the position of Cys-108 (mutated to Arg), which forms a disulfide bond with Gp1, are indicated with arrows. The electrostatic surface was contoured between -30 kT/e and +30 kT/e. This figure was created with GRASP (Nicholls et al., 1991).

(B) A hypothetical model of the full-length ectodomain of GP2 based on the portion of the structure determined here (X-ray structure shown in blue) showing the proposed location (dotted lines) of the fusion peptide in a hypothetical disulfide-bonded loop, the C-terminal disordered segments (Cgp2), and the proposed disulfide bond to GP1 (GP2 Cys-108 to GP1 Cys-19). Neither the Cys-10 to Cys-51 nor the Cys-108 to GP1 Cys-19 disulfide bonds have been demonstrated biochemically or structurally, but they are proposed based on sequence conservation and distances seen in the GP2 structure (see text).

retroviral sequences (e.g., Figure 2A), or both are missing, suggesting that they are linked. This arrangement would position the internal, nonpolar fusion peptide, residues 23–38 (Figure 1A), as part of a disulfide-bonded loop at one end of the rod-shaped GP2 (Figure 5B). Cysteine 55 (Figure 2D, C55E) is accessible on the outer surface of the core $\alpha$ helix in a groove between the outer layer helix-strand segments so that the returning part of the fusion peptide loop could form a second outer layer on the structure (Figure 5B).

Because they are disordered in the crystal, the location of the C-terminal 16 residues of the GP2 ectodomain is unknown beyond the outer layer $\alpha$ helix (dotted lines in Figures 2A and 5B). In full-length GP2, these residues are attached to the hydrophobic anchor in the viral membrane, so that the orientation of the rod-shaped GP2 relative to the viral membrane is determined by the structure of this apparently flexible segment. Seventeen residues cannot extend the 62 Å back to the opposite end of GP2, which suggests that in this form the GP2 rod will be oriented with the viral membrane near the same end of the rod as the fusion peptides. If, as an alternative, the last two structured residues (residues 132–133) of the ectodomain were to peel off the rod, this C-terminal region would be long enough to reach either end (hence two arrows at GP2 C termini in Figure 5B). Similar considerations apply to the C-terminal regions of HA2 and gp41.

GP1 is disulfide bonded to GP2, and the location of the bond between GP2 Cys-108 and GP1 Cys-19 has been proposed but not demonstrated biochemically (Gallaher, 1996; Sanchez et al., 1998a). The structure of GP2 reported here supports this proposal by locating Cys-108 (Figure 2D, C108R) at the opposite end of the molecule from the other free cysteines of the fusion peptide loop (Figure 5B) (although strictly an internal GP2 disulfide bond from Cys-10 to Cys-108 cannot be ruled out). Sequence comparisons suggest that in some retroviruses the cysteines homologous to GP2 Cys-108 (Figure 5B) and GP1 Cys-19 could also form the interchain disulfide bonds between the equivalent retroviral Su and Tm residues (e.g., RSV and ALSV [Herdandez et al., 1997; but see also Pinter et al., 1997]). The GP2 structure indicates that this tentative disulfide bond would attach the receptor-binding domain, GP1, at the end of the rod-shaped molecule distant from the N and C termini (Figure 5B). This is very similar to the disulfide bond attachment of influenza HA1 to HA2, which is between HA1 Cys-14, near the N-terminal of HA1, and HA2 Cys-137, located at the opposite end of the low-pH-induced HA2 rod from the fusion peptide (yellow ball and stick in Figure 4C). In Figure 5B, the N and C termini

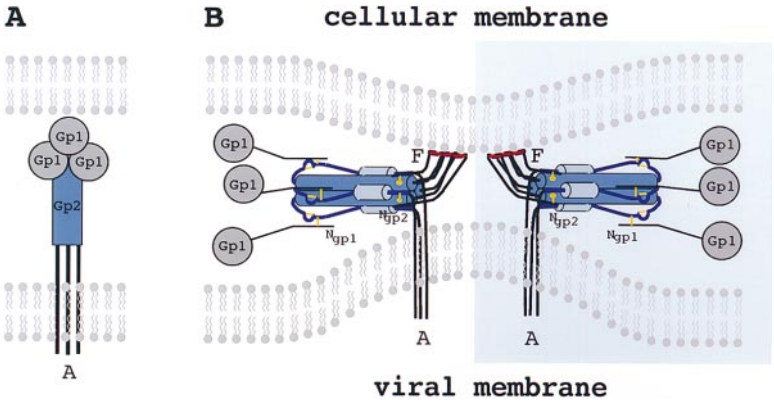


Figure 6. Hypothetical Model of Some Steps in the Membrane Fusion Mechanism

(A) and (B) represent hypothetical intermediates and are based on Figure 3 of Weissenhorn et al. (1997b) (see text). Fusion peptides labeled F are colored red. GP2 transmembrane anchors are labeled A. The GP2 outer layer α helices are colored light blue; the N-terminal, core coiled coil is dark blue; disulfide bonds are yellow. Ngp2 and Ngp1 label the proposed location of the N termini of those polypeptide chains. The gray spheres represent the receptor-binding domain of GP1. The picture also incorporates suggestions that more than one trimer might be involved in forming an initial fusion pore (Danieli et al., 1996) and that the bilayers may be distorted with molecules entering at an angle (Ruigrok et al., 1992; Tatulian et al., 1995), which in some way results in distortions to the membrane that favor membrane fusion (e.g., Chernomordik et al., 1998).

of GP1 (dotted lines) are oriented by analogy to influenza HA, in which at low pH, the β strand containing HA1 Cys-14 is inverted projecting the receptor-binding domain away from the rod (red in Figure 4C); whereas in the metastable structure, the HA receptor-binding domain points in the opposite direction (see Figure 3B in Bullough et al. [1994]).

**Membrane Fusion**

An intermediate in the membrane fusion mechanism like that in Figure 6B was proposed (Weissenhorn et al., 1997b, 1998b) to incorporate the observations that in both HA2 and gp41, the fusion peptide and membrane anchor sequences were projecting toward one end of a thin rod-shaped molecule that could facilitate close apposition of the prefusion membranes. A similar model is suggested for GP2 based on its similarity in structure to gp41. In some strains of HIV-1, gp120 dissociates from gp41, but in influenza and, as we propose here, in Ebola virus, the receptor-binding domain is attached by a disulfide bond at the opposite end of the fusion protein rod, where it could be withdrawn from the site of fusion.

**A Prefusion Intermediate and/or a Postfusion Conformation?**

In the absence of direct studies of the membrane fusion–active form of filovirus GP2, the similarity of the structure observed here to recombinant gp41 and low-pH-induced HA2 provides evidence for the functional significance of this structure. Only three conformational states of the membrane fusion subunits of viral glycoproteins have been defined that are not transient and that are sufficiently stable for biochemical and structural study (reviewed in Skehel et al., 1995; Chen et al., 1998a): the single-chain, precursor state (HA0, gp160, and GP) that is not infectious because it cannot be triggered to a membrane fusion–active conformation; the postranslationally cleaved, primed state (HA1/HA2, gp120/gp41, and GP1/GP2) that is also membrane fusion inactive but can be triggered to undergo a conformational change by different events in different viruses (receptor and coreceptor binding in HIV-1, low pH for influenza virus, and a currently unknown stimulus in filoviruses); and a stable, rod-like conformation (Figure 4) observed because they are the states formed spontaneously by the membrane fusion subunit when expressed alone (HA2, gp41, and GP2) and in the case of influenza virus, when induced at the low pH of membrane fusion.

One possibility is that the rod-like conformation observed here for GP2 ectodomain is the conformation of both a prefusion intermediate (perhaps as a component of an oligomeric pore) and the postfusion state of the ectodomain. The distinction is that in the prefusion intermediate the N and C termini are inserted in the cellular and viral membranes, respectively, whereas in the postfusion state both sets of termini are inserted in the one remaining, fused membrane. This possibility was suggested for low-pH-induced HA2 based on the apparent flexibility between the molecular rod and both its termini and the observation that all of the mutations that altered the pH of the fusion process were positioned in the structure at points of structural change leading to the rod-shaped conformation (Bullough et al., 1994). It was also suggested for gp41 because in the X-ray studies both the membrane anchor and the fusion peptides were, more clearly than in the HA2 structure, at the same end of the rod-shaped structure where they might draw two membranes together (Weissenhorn et al., 1997a). (A corollary of the possibility that there is a single rod-like prefusion intermediate and postfusion state is that whatever distortion or stress to the bilayers that might exist at the attachment positions of the prefusion intermediate to the two membranes and might favor membrane fusion [e.g., Sutton et al., 1998] would be relieved postfusion at the site of attachment to the fused membrane, extinguishing the membrane fusion activity.)

An alternative possibility, that GP2 mediates membrane fusion while its ectodomain is in a transient conformation, preceding the rod-shape structure observed here, is addressed by studies of inhibitors of membrane fusion of the similarly shaped gp41. Peptides corresponding to both the core coiled coil and the outer layer α helices of gp41 inhibit viral infection by inhibiting membrane fusion (Wild et al., 1994). Structural studies

suggested that these peptides block fusion by blocking the refolding to the rod-shaped conformation (Lu et al., 1995; Chan et al., 1997; Tan et al., 1997; Weissenhorn et al., 1997b; Malashkevich et al., 1998). Resistant mutations map to the coiled coil as expected from this proposed mechanism (Rimsky et al., 1998), and inhibition only occurs when peptide is present at the time of fusion when gp41 either becomes exposed and/or undergoes a structural change (Furuta et al., 1998). Other studies suggest that peptides may inhibit by disrupting the rod-shaped conformation after it forms (Caffrey et al., 1998) or preventing it from assembling into an oligomeric fusion pore (Munoz-Barroso et al., 1998).

### Ancestral Viral Fusion Proteins

The attachment of GP1 to GP2 by a disulfide bond near the bottom of the GP2 stem, if confirmed, suggests that the N-terminal segment of GP1 may form part of the stem of GP1-GP2 in the conformation found on virus, in a similar way that the first 50 residues of HA1 form part of the stem of influenza HA1-HA2. The recently determined structure of HEF, the hemagglutinin-esterase fusion glycoprotein of influenza C virus, suggested, by comparison with the HA structure, that the stem of HA, including N-terminal and C-terminal segments of HA1 as well as all of HA2, constitute a single domain derived from an ancestral fusion protein into which a receptor domain was inserted at a surface loop (Rosenthal et al., 1998; Zhang et al., submitted).

The attachment of GP1 to GP2 (and, e.g., RSV Su to Tm) and evidence that gp120 also contacts the disulfide-bonded loop at the base of gp41 (Sattentau and Moore, 1991; Weissenhorn et al., 1996) suggest the possibility that retrovirus and filovirus glycoproteins may derive from another ancestral stem-like fusion protein composed of the membrane fusion subunit (gp41, TM, and GP2) and the N-terminal segment of the subunit preceding the posttranslational cleavage site, and that most of the remainder of the first subunit inserted into that ancestral fusion protein to provide receptor binding activity. Whether the retroviral and filoviral fusion protein undergoes a complex refolding event to achieve the state observed for gp41 and GP2, or whether this conformation is simply exposed from a hidden location in the uninduced glycoprotein, is unknown.

### Comparison with Intracellular Vesicle Fusion

The fusion of intracellular transport vesicles involves the assembly of a four-stranded $\alpha$-helical coil from three SNARE molecules, including one anchored in the vesicle and one anchored in the target membrane (Figure 4F) (e.g., Brünger et al., 1998; Nicholson et al., 1998; Poirier et al., 1998; Sutton et al., 1998). The C-terminal membrane anchors of both membrane proteins are at the same end of the helical rod (Hanson et al., 1997; Lin and Scheller, 1997; Poirier et al., 1998; Sutton et al., 1998), suggesting that a similarly shaped molecular complex is used for bringing membranes into apposition and mediating membrane fusion by both viruses and cells. In the case of cellular vesicle fusion, protein components of both membranes contribute to the formation of the rod-shaped complex (Figure 4E); for viruses, the rod-shaped complex is formed exclusively by the virus fusion protein; its fusion peptide, by interacting with the target membrane, substitutes for the membrane anchor sequence of the cellular target membrane protein in intracellular vesicle fusion.

## Conclusions

The structure of the Ebola virus GP2 ectodomain adds to observations of similarities between virus fusion glycoproteins that have accumulated from analyses of their amino acid sequences, their activities, and their biosynthesis. For influenza, retro-, and filoviruses in particular, there is a common requirement for proteolytic cleavage of a biosynthetic precursor into distinct receptor-binding and membrane fusion domains, and common structural properties have been predicted from their sequences. These similarities are now emphasized by the observations that irrespective of either their initial structures or of differences in the processes involved in their activation for membrane fusion, the isolated fusion domains are rod-shaped coiled coils. This homology strongly suggests that the structures have a role in membrane fusion and, for HA, it reinforces previous conclusions that the changes in conformation required to form this structure are necessary for fusion activity (Bullough et al., 1994). Specific details of the Ebola GP2 structure, especially irregularities in the core coiled coil, a conserved disulfide loop segment that may be a hinge, and the likely location of the site of attachment to GP1, suggest that filovirus fusion proteins might also be required to refold in some way to mediate fusion.

## Experimental Procedures

### Expression, Purification, and Crystallization

Hybrids of a modified 30-residue GCN4 trimeric coiled coil linked to residues 51–149 of GP2 (residues 552–650 of GP) were expressed and purified in the same way as the construct pIIgp2(552–650) described earlier (Weissenhorn et al., 1998a). To promote crystallization, two residues, Cys-55 (Ser-565 in pIIgp2(552–650)) and Cys-108 (Ser-609 in pIIgp2(552–650)) were mutated to Glu and Arg, respectively, by site-directed mutagenesis, following standard PCR protocols. To make a heavy atom isomorphous derivate, Leu-12 of the GCN4 zipper was mutated to Cys, and the protein pIIGp2(51–149)cys was reacted with p-hydroxy-mercuribenzoate before crystallization.

Crystals were grown by vapor diffusion in hanging drops by combining 2 $\mu$l of protein solution (5 mg/ml in 20 mM Tris–HCl [pH 9.3]) with 1 $\mu$l of reservoir solution (100 mM Tris–HCl [pH 8.5], 30%–34% PEG 4000, 220 mM ammonium sulphate, 1.0% dioxane). Derivatized pIIGp2(51–149)cys was crystallized by the same method and from the same reservoir buffers as the wild-type pIIGp2(51–149) but with the addition of 5 mM lauryldimethylamine oxide (Hampton Research) to the crystallization drop.

### X-Ray Diffraction Data Collection

Both proteins (wild-type and derivatized form) crystallize isomorphously in space group C2 with unit cell dimensions a = 171.70 Å, b = 32.69 Å, c = 168.86 Å, $\alpha$ = 90°, $\beta$ = 119.23°, $\gamma$ = 90° with two molecules per asymmetric unit. Diffraction data from both native and derivative crystals were measured at room temperature using an 18 cm image plate scanner (MAR Research, Hamburg Germany) mounted on an Elliott GX-13 rotating anode source (Elliot, London) with mirror optics. Oscillation images were processed using DENZO (Otwinowski and Minor, 1997), and data reduction was carried out using SCALEPACK (Otwinowski, 1993) (Table 1A). The CCP4 suite

of programs was used for further data processing and analysis (Collaborative Computational Project Number 4, 1994).

### Phasing

The structure was determined by phasing with single isomorphous replacement anomalous scattering (SIRAS) and density modification using the p-hydroxy-mercuribenzoate derivative as the initial source of phase information. Six heavy atom positions (corresponding to the GCN4 Cys-12 positions) were located in an isomorphous difference Patterson map using RSPS (Collaborative Computational Project Number 4, 1994). Refinement of the heavy atom parameters was carried out with the program SHARP (Fortelle and Bricogne, 1997) (Table 1B). Three additional sites, spaced approximately 4 Å from the ones found initially, were located by a peak search using a log likelihood gradient map after SIR phasing. These sites may correspond to alternative cysteine conformations. Density modification, consisting of solvent flipping using an envelope determined by analysis of the local standard deviation of the SIRAS map, histogram matching, and phase extension from 3.5 to 3.0 Å resolution, was carried out with the program SOLOMON (Collaborative Computational Project Number 4, 1994). The resulting electron density map was of sufficient quality to fit the two GCN4 molecules (residues 3–28), a Cα trace for the inner and outer layer helices of one Ebola molecule, and three shorter fragments of the inner and outer layer helices for the second Ebola molecule (molecule 2). Although some electron density was evident for the linker loop region, no model could be built into it. From this model, it was clear that three-fold symmetry axes of the GCN4 trimers were slightly misaligned from the three-fold axes of the covalently linked Ebola GP2 trimers (for the two GCN4 trimers, $\kappa = 175.46°$, $\omega = 91.49°$, $\phi = 81.25°$; for the Ebola trimers, $\kappa = 174.93$, $\omega = 90.32$, $\phi = 90.76$).

The Cα coordinates were used to define four sets of three-fold NCS matrices (two for the GCN4 and two for the Ebola trimers) with LSQKAB (Collaborative Computational Project Number 4, 1994). The program NCSMASK (Collaborative Computational Project Number 4, 1994) was used to generate four NCS masks that did not include the GCN4–Ebola connections. Iterative solvent flattening, histogram matching, and three-fold NCS averaging with phase extension from 3.5 Å to 3.0 Å were performed with DM (Collaborative Computational Project Number 4, 1994) without phase combination to the previous cycle. The resulting experimental electron density map was of excellent quality and showed the location of most of the side chains. The unweighted phase difference at 20–3.0 Å resolution between the SIRAS phases and the phases after density modification was 74°. A partial model (residues 3–29 of GCN4 and 52–128 of Ebola GP2) of one monomer was readily built using the program O (Jones et al., 1991), and the rest of the model was generated applying the NCS operators. A short stretch of disconnected electron density (sufficient for approximately two or three residues) was observed in proximity of one monomer (F) of molecule 1 but could not be interpreted.

### Refinement

Before refinement, 5% of the reflections were set aside for calculation of the free R factor (Brünger, 1992a). An overall B factor of 30 $Å^2$ was applied to all atoms. The model was refined in X-PLOR (Brünger, 1992b) initially with tight restraints on both main and side chains. Four-group rigid body fitting, followed by 300 cycles of positional refinement and torsion angle dynamics simulating annealing with tight restraints on main and side chains for the six monomers, reduced the $R_{free}$ from 41.1% to 36.7% and the $R_{work}$ from 43.8% to 34.2% between 8.0 and 3.0 Å resolution. New phases and figures of merit were calculated from the refined model using SFALL and SIGMAA (Collaborative Computational Project Number 4, 1994) and were used in DM for a new cycle of iterative NCS averaging. Nonaveraged $\sigma_A$ weighted electron density maps were used in subsequent model-building stages during which the non-three-fold-symmetric connections between GCN4 and Ebola GP2 (residues 29 and 30 of GCN4 and 50 and 51 of GP2) were built. Three additional cycles of manual rebuilding and positional and B group refinement with tight restraints on the main chain (residues 5–28 and 55–129) and weak restraints on side chains reduced the $R_{free}$ to 27.8% and $R_{work}$ to 24.6%. At that stage, three metal ions (which may be one metal atom and two waters) involved in crystal lattice interactions and two chloride ions (inside the Ebola trimeric coiled coils) were modeled in the electron density. The three tentatively assigned metal atoms are 9.0 σ, 5.0 σ, and 4.8 σ electron density peaks (in omit difference maps calculated with phases from the refined model) and refined to B factors of 42.3 $Å^2$. The strongest peak is coordinated between trimers in the crystal by two histidine residues (His-112B and His-127D) with short (average bond distance = 2.42 Å) interatomic distances to the side-chain nitrogens, characteristic of metal ions. The other two solvent peaks, one in each trimer in the asymmetric unit, which may be metals or well-ordered waters, are at lattice contacts and are coordinated (average bond distance = 2.89 Å) between His-112 of GP2 and Gln-3 of the GCN4 sequence (from trimer chains C and E to chains F and A, respectively). The two chloride ions were modeled into 5.8 and 6.0 s electron density peaks and refined to a B factor of 39.5 $Å^2$. They are assigned as chloride ions based on the close similarity of their locations to a chloride ion in Mo-55 (Fass et al., 1996) on sequence conservation of the ligands (Asn-85) (see text). Refinement was concluded using CNS (Brünger et al., 1998) with restrained conjugate gradient minimization and restrained B group refinement. A uniform bulk solvent correction (B = 41.57 $Å^2$, k = 0.29 e/$Å^3$) and overall anisotropic thermal factor ($B_{11} = -6.46$ $Å^2$, $B_{22} = -5.11$ $Å^2$, $B_{33} = 11.546$ $Å^2$, $B_{13} = -3.41$ $Å^2$) were applied. The final model contains residues 3–131 for chains 1C and 2E; 3–132 for chains 1A, 1B, and 2D; and 3–133 for chain 2F, two chloride and three zinc ions. The model has good stereochemistry and structural quality with an average bond length and bond angle deviation of 0.016 Å and 1.54°, respectively, and no Ramachandran violations (Table 1C). Residues 32 and 51 of both monomer A and F are built in poor electron density.

The $R_{free}$ is 25.6%, and $R_{work}$ is 23.9% for the resolution range 20–3.0 Å (Table 1C). Some noninterpreted electron density, enough for 2 or 3 residues, can be observed at 0.5 σ contour level in $\sigma_A$ weighted electron density maps and may correspond to part of the disordered C terminus of the Ebola molecule.

### Acknowledgments

We thank Gary Nabel and A. Sanchez for the DNA clone of the Ebola virus glycoprotein, Dr. Bostjian Kobe for sharing unpublished data on HTLV-1 gp21, Dr. M. Clore for the SIV coordinates, M. Pietras for excellent technical assistance, and the members of the Harrison/Wiley laboratory for helpful discussions. W. W. was supported by the HHMI. A. C. is supported by a AFRT (Association Française pour la Recherche Therapeutique) postdoctoral fellowship. This work was supported by a supplement to the NIAID grant (5RO1AI13654-20) for Expanded International Research on Emerging and Re-Emerging Diseases, the Medical Research Council (UK), and the Howard Hughes Medical Institute. D. C. W. is an investigator of the HHMI.

Received October 2, 1998; revised October 21, 1998.

### References

Becker, Y. (1996). Retrovirus and filovirus "immunosuppressive motif" and the evolution of virus pathogenicity in HIV-1, HIV-2, and Ebola viruses. Virus Genes *11*, 191–195.

Becker, S., Spiess, M., and Klenk, H.D. (1995). The asialoglycoprotein receptor is a potential liver-specific receptor for Marburg virus. J. Gen. Virol. *76*, 393–399.

Brown, J.H., Cohen, C., and Parry, D.A. (1996). Heptad breaks in α-helical coiled coils: stutters and stammers. Proteins *26*, 134–145.

Brünger, A.T. (1992a). Free R value: a novel, statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472–475.

Brünger, A.T. (1992). X-PLOR: A System for X-Ray Crystallography and NMR. (New Haven, CT: Yale University).

Brünger, A.T., Adams, P.D., Clore, G.M., Gros, P., Grosse-Kuntsleve, R.W., Jiang, J.-S., Kuszerski, J., Nilges, M., Pannu, N.S., and Read, R.J. (1998). Crystallographic and NMR system (CNS): a new software system for macromolecular structure determination. ACTA Crystallogr. D *54*, 905–921.

Bukreyev, A.A., Volchkov, V.E., Blinov, V.M., and Netesov, S.V.

(1993). The GP-protein of Marburg virus contains the region similar to the 'immunosuppressive domain' of oncogenic retrovirus P15E proteins. FEBS Lett. *323*, 183–187.

Bullough, P.A., Hughson, F.M., Skehel, J.J., and Wiley, D.C. (1994). Structure of influenza haemagglutinin at the pH of membrane fusion. Nature *371*, 37–43.

Caffrey, M., Cai, M., Kaufman, J., Stahl, S.J., Wingfield, P.T., Covell, D.G., Gronenborn, A., and Clore, G.M. (1998). Three-dimensional structure of the 44 kDa ectodomain of SIV gp41. EMBO J. *17*, 4572–4584.

Carr, C.M., Chaudhry, C., and Kim, P.S. (1997). Influenza hemagglutinin is spring-loaded by a metastable native conformation. Proc. Natl. Acad. Sci. USA *94*, 14306–14313.

Carson, M. (1991). Ribbons 2.0. J. Appl. Crystallogr. *24*, 958–961.

Centers for Disease Control and Prevention (1995). Update: outbreak of Ebola viral hemorrhagic fever—Zaire, 1995. Morbid. Mortal. Week. Rep. *44*, 381–382.

Chan, D.C., Fass, D., Berger, J.M., and Kim, P.S. (1997). Core structure of gp41 from the HIV envelope glycoprotein. Cell *89*, 263–273.

Chen, J., Wharton, S.A., Weissenhorn, W., Calder, L.J., Hughson, F.M., Skehel, J.J., and Wiley, D.C. (1995). A soluble domain of the membrane-anchoring chain of influenza virus hemagglutinin (HA2) folds in *Escherichia coli* into the low-pH-induced conformation. Proc. Natl. Acad. Sci. USA *92*, 12205–12209.

Chen, J., Lee, K.H., Steinhauer, D.A., Stevens, D.J., Skehel, J.J., and Wiley, D.C. (1998a). Structure of the hemagglutinin precursor cleavage site, a determinant of influenza pathogenicity and the origin of the labile conformation. Cell *95*, 409–417.

Chen, J., Skehel, J.J., and Wiley, D.C. (1998b). A polar octapeptide fused to N-terminal fusion peptide solubilizes the influenza virus HA2 subunit ectodomain. Biochemistry *37*, 13643–13649.

Chernomordik, L.V., Frolov, V.A., Leikina, E., Bronk, P., and Zimmerberg, J. (1998). The pathway of membrane fusion catalyzed by influenza hemagglutinin: restriction of lipids, hemifusion, and lipidic fusion pore formation. J. Cell Biol. *140*, 1369–1382.

Cianciolo, G.J., Copeland, T.D., Oroszlan, S., and Snyderman, R. (1985). Inhibition of lymphocyte proliferation by a synthetic peptide homologous to retroviral envelope proteins. Science *230*, 453–455.

Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D *50*, 760–763.

Crick, F. (1953). The packing of alpha-helices: simple coiled-coils. Acta Crystallogr. *6*, 689–697.

D'Souza, M.P., and Harden, V.A. (1996). Chemokines and HIV-1 second receptors. Nat. Med. *2*, 1293.

Danieli, T., Pelletier, S.L., Henis, Y.I., and White, J.M. (1996). Membrane fusion mediated by the influenza virus hemagglutinin requires the concerted action of at least three hemagglutinin trimers. J. Cell Biol. *133*, 559–569.

Fass, D., and Kim, P.S. (1995). Dissection of a retrovirus envelope protein reveals structural similarity to influenza hemagglutinin. Curr. Biol. *5*, 1377–1383.

Fass, D., Harrison, S.C., and Kim, P.S. (1996). Retrovirus envelope domain at 1.7 Å resolution. Nat. Struct. Biol. *3*, 465–469.

Feldmann, H., Will, C., Schikore, M., Slenczka, W., and Klenk, H.-D. (1991). Glycosylation and oligomerization of the spike protein of Marburg virus. Virology *182*, 353–356.

Feldmann, H., Nichol, S.T., Klenk, H.D., Peters, C.J., and Sanchez, A. (1994). Characterization of filoviruses based on differences in structure and antigenicity of the virion glycoprotein. Virology *199*, 469–473.

Fortelle, E.d.L., and Bricogne, G. (1997). Maximum likelihood heavy atom refinement for multiple isomorphous replacement and multi anomalous diffraction methods. Methods Enzymol. *276*, 472–494.

Furuta, R.A., Wild, C.T., Weng, Y., and Weiss, C.D. (1998). Capture of an early fusion-active conformation of HIV-1 gp41. Nat. Struct. Biol. *5*, 276–279.

Gallaher, W.R. (1996). Similar structural models of the transmembrane proteins of Ebola and avian sarcoma viruses. Cell *85*, 477–478.

Hanson, P.I., Roth, R., Morisaki, H., Jahn, R., and Heuser, J.E. (1997). Structure and conformational changes in NSF and its membrane receptor complexes visualized by quick-freeze/deep-etch electron microscopy. Cell *90*, 523–535.

Harbury, P.B., Zhang, T., Kim, P.S., and Alber, T. (1993). A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. Science *262*, 1401–1407.

Herdandez, L.D., Peters, R.J., Delos, S.E., Young, J.A.T., Agard, D.A., and White, J.M. (1997). Activation of a retroviral membrane fusion protein: soluble receptor-induced liposome binding of the ALSV envelope glycoprotein. J. Cell Biol. *139*, 1455–1464.

Hunter, E., and Swanstrom, R. (1990). Retrovirus envelope glycoproteins. Curr. Top. Microbiol. Immunol. *157*, 187–253.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Joshi, S.B., Dutch, R.E., and Lamb, R.A. (1998). A core trimer of the paramyxovirus fusion protein: parallels to influenza virus haemagglutinin and HIV-1 gp41. Virology *248*, 20–34.

Klenk, H.-D., Angliker, H., Garten, W., Hallenberger, S., Ohuchi, M., Ohuchi, R., Rott, R., Shaw, E., Stieneke-Grober, A., and Vey, M. (1993). Processing of viral glycoproteins by host proteases: structural aspects and functional consequences. In Virus Strategies, pp. 195–214 (Weinheim, Germany: Verlag Chemie).

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–290.

Lin, R.C., and Scheller, R.H. (1997). Structural organization of the synaptic exocytosis core complex. Neuron *19*, 1087–1094.

Lu, M., Blacklow, S.C., and Kim, P.S. (1995). A trimeric structural domain of the HIV-1 transmembrane glycoprotein. Nat. Struct. Biol. *2*, 1075–1082.

Lumb, K.J., and Kim, P.S. (1995). A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. Biochemistry *34*, 8642–8648.

Malashkevich, V.N., Chan, D.C., Chutkowski, C.T., and Kim, P.S. (1998). Crystal structure of the simian immunodeficiency virus (SIV) gp41 core: conserved helical interactions underlie the broad inhibitory activity of gp41 peptides. Proc. Natl. Acad. Sci. USA *95*, 9134–9139.

Munoz-Barroso, I., Durell, S., Sakaguchi, K., Appella, E., and Blumenthal, R. (1998). Dilation of the human immunodeficiency virus-1 envelope glycoprotein fusion pore revealed by the inhibitory action of a synthetic peptide from gp41. J. Cell Biol. *140*, 315–323.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Prot. Struct. Funct. Genet. *11*, 281–296.

Nicholson, K.L., Munson, M., Miller, R.B., Filip, T.J., Fairman, R., and Hughson, F.M. (1998). Regulation of SNARE complex assembly by an N-terminal domain of the t-SNARE Sso1p. Nat. Struct. Biol. *5*, 793–802.

Otwinowski, Z. (1993). Oscillation data reduction program. In Data Collection and Processing, C.W. Carter, Jr., and R.M. Sweet, eds. (New York: Academic Press).

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Pinter, A., Kopelman, R., Li, Z., Kayman, S.C., and Sanders, D.A. (1997). Localization of the labile disulfide bond between SU and TM of the murine leukemia virus envelope protein complex to a highly conserved CWLC motif in SU that resembles the active-site sequence of thiol-disulfide exchange enzymes. J. Virol. *71*, 8073–8077.

Poirier, M.A., Xiao, W., Macosko, J.C., Chan, C., Shin, Y.-K., and Bennett, M.K. (1998). The synaptic SNARE complex is a parallel four-stranded helical bundle. Nat. Struct. Biol. *5*, 765–769.

Rimsky, L.T., Shugars, D.C., and Matthews, T.J. (1998). Determinants of human immunodeficiency virus type 1 resistance to gp41-derived inhibitory peptides. J. Virol. *72*, 986–993.

Rosenthal, P.B., Zhang, X., Formanowski, F., Fitz, W., Wong, C.-W.,

Meier-Ewert, H., Skehel, J.J., and Wiley, D.C. (1998). Three dimensional structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. Nature *396*, 92–96.

Ruigrok, R.W.H., Hewat, E.A., and Wade, R.H. (1992). Low pH deforms the influenza virus envelope. J. Gen. Virol. *73*, 995–998.

Sanchez, A., Kiley, M.P., Holloway, B.P., and Auperin, D.D. (1993). Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. Virus Res. *29*, 215–240.

Sanchez, A., Ksiazek, T.G., Rollin, P.E., Peters, C.J., Nicol, S.T., Khan, A.S., and Mahy, B.W.J. (1995). Reemergence of Ebola virus in Africa. Emerging Infect. Dis. *1*, 96–97.

Sanchez, A., Trappier, S.G., Mahy, B.W., Peters, C.J., and Nichol, S.T. (1996). The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. Proc. Natl. Acad. Sci. USA *93*, 3602–3607.

Sanchez, A., Trappier, S.G., Stroher, U., Nichol, S.T., Bowen, M.D., and Feldmann, H. (1998a). Variation in the glycoprotein and VP35 genes of Marburg virus strains. Virology *240*, 138–145.

Sanchez, A., Yang, Z.-Y., Xu, L., Nabel, G.J., Crews, T., and Peters, C.J. (1998b). Biochemical analysis of the secreted and virion glycoproteins of Ebola virus. J. Virol. *72*, 6442–6447.

Sattentau, Q., and Moore, J.P. (1991). Conformational changes induced in the human immunodeficiency virus envelope glycoprotein by soluble CD4 binding. J. Exp. Med. *174*, 407–415.

Scheid, A., and Choppin, P.W. (1974). The hemagglutination and neuraminidase protein of a paramyxovirus: interaction with neuraminic acid in affinity chromatography. Virology *62*, 125–133.

Schnittler, H.J., Mahner, F., Drenckhahn, D., Klenk, H.D., and Feldmann, H. (1993). Replication of Marburg virus in human endothelial cells. A possible mechanism for the development of viral hemorrhagic disease. J. Clin. Invest. *91*, 1301–1309.

Siegret, R., Shu, H.-L., Slenczka, W., Peters, D., and Muller, G. (1967). Zur Athiologie einer unbekannten von Affen ausgegangenen Infektionskrankheit. Dtsch. Med. Wochenschr. *92*, 2341–2343.

Skehel, J.J., Bayley, P.M., Brown, E.B., Martin, S.R., Waterfield, M.D., White, J.M., Wilson, I.A., and Wiley, D.C. (1982). Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. Proc. Natl. Acad. Sci. USA *79*, 968–972.

Skehel, J.J., Bizebard, T., Bullough, P.A., Hughson, F.M., Knossow, M., Steinhauer, D.A., and Wharton, S.A. (1995). Membrane fusion by influenza hemagglutinin. Cold Spring Harb. Symp. Quant. Biol. *60*, 573–580.

Sutton, R.B., Fasshauer, D., Jahn, R., and Brünger, A.T. (1998). Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4Å resolution. Nature *395*, 347–353.

Tan, K., Liu, J., Wang, J., Shen, S., and Lu, M. (1997). Atomic structure of a thermostable subdomain of HIV-1 gp41. Proc. Natl. Acad. Sci. USA *94*, 12303–12308.

Tatulian, S.A., Hinterdorfer, P., Baber, G., and Tamm, L.K. (1995). Influenza hemagglutinin assumes a tilted conformation during membrane fusion as determined by attenuated total reflection FTIR spectroscopy. EMBO J. *14*, 5514–5523.

Volchkov, V.E., Blinov, V.M., and Netesov, S.V. (1992). The envelope glycoprotein of Ebola virus contains an immunosuppressive-like domain similar to oncogenic retroviruses. FEBS Lett. *305*, 181–184.

Volchkov, V.E., Feldmann, H., Volchkova, V.A., and Klenk, H.-D. (1998). Processing of the Ebola virus glycoprotein by the proprotein convertase furin. Proc. Natl. Acad. Sci. USA *95*, 5762–5767.

Weissenhorn, W., Wharton, S.A., Calder, L.J., Earl, P.L., Moss, B., Aliprandis, E., Skehel, J.J., and Wiley, D.C. (1996). The ectodomain of HIV-1 env subunit gp41 forms a soluble, alpha-helical, rod-like oligomer in the absence of gp120 and the N-terminal fusion peptide. EMBO J. *15*, 1507–1514.

Weissenhorn, W., Calder, L.J., Dessen, A., Laue, T., Skehel, J.J., and Wiley, D.C. (1997a). Assembly of a rod-shaped chimera of a trimeric GCN4 zipper and the HIV-1 gp41 ectodomain expressed in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *94*, 6065–6069.

Weissenhorn, W., Dessen, A., Harrison, S.C., Skehel, J.J., and Wiley, D.C. (1997b). Atomic structure of an ectodomain from HIV-1 gp 41. Nature *387*, 426–430.

Weissenhorn, W., Calder, L.J., Wharton, S.A., Skehel, J.J., and Wiley, D.C. (1998a). The central structural feature of the membrane fusion protein subunit from the Ebola virus glycoprotein is a long triple stranded coiled coil. Proc. Natl. Acad. Sci. USA *95*, 6032–6036.

Weissenhorn, W., Dessen, A., Calder, L.J., Harrison, S.C., Skehel, J.J., and Wiley, D.C. (1998b). Structural basis for membrane fusion by enveloped viruses. Mol. Membr. Biol., in press.

Wharton, S.A., Calder, L.J., Ruigrok, R.W., Skehel, J.J., Steinhauer, D.A., and Wiley, D.C. (1995). Electron microscopy of antibody complexes of influenza virus haemagglutinin in the fusion pH conformation. EMBO J. *14*, 240–246.

Wild, C., Dubay, J.W., Greenwell, T., Baird, T., Jr., Oas, T.G., McDanal, C., Hunter, E., and Matthews, T. (1994). Propensity for a leucine zipper-like domain of human immunodeficiency virus type 1 gp41 to form oligomers correlates with a role in virus-induced fusion rather than assembly of the glycoprotein complex. Proc. Natl. Acad. Sci. USA *91*, 12676–12680.

Wiley, D.C., Skehel, J.J., and Waterfield, M. (1977). Evidence from studies with a cross-linking reagent that the haemagglutinin of influenza virus is a trimer. Virology *79*, 446–448.

Wiley, D.C., Wilson, I.A., and Skehel, J.J. (1981). Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature *289*, 373–378.

Wilson, I.A., Skehel, J.J., and Wiley, D.C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature *289*, 366–373.

Yang, Z.-y., Delgado, R., Xu, L., Todd, R.F., Nabel, E.G., Sanchez, A., and Nabel, G.J. (1998). Distinct cellular interactions of secreted and transmembrane Ebola virus glycoproteins. Science *279*, 1034–1037.

**Brookhaven Protein Data Bank Accession Number**

The accession number for the coordinates of the structure reported in this paper is 1ebo. The coordinates are also available from carfi@crystal.harvard.edu.