

Protein Dissection of the Ebola Virus Glycoprotein

by

Brian J. Schneider

B.S., Biological Sciences: Biochemistry
University of Rochester, 1995

Submitted to the Department of Biology
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Biology

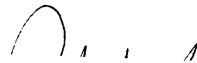
at the

Massachusetts Institute of Technology

June 1999

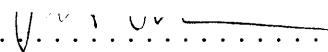
© 1999 Massachusetts Institute of Technology
All Rights Reserved

Signature of Author



Department of Biology
May 7, 1999

Certified by



Peter S. Kim
Professor of Biology
Thesis Supervisor



Accepted by



Alan D. Grossman
Co-Chair, Graduate Committee
Department of Biology

PROTEIN DISSECTION OF THE
EBOLA VIRUS GLYCOPROTEIN

by

BRIAN J. SCHNEIDER

Submitted to the Department of Biology
on May 7, 1999 in partial fulfillment of the
requirements for the Degree of
Master of Science in Biology

ABSTRACT

The Ebola virus causes a hemorrhagic fever in humans and primates with high mortality. The envelope protein necessary for viral fusion with cellular membranes is the glycosylated protein (GP) which is post-translationally cleaved into GP1 and GP2. A 95 amino acid portion of the GP2 between the transmembrane and fusion domains (Ebo-95) was recombinantly expressed. Thermal denaturation and proteolytic stability experiments on different disulfide permutations show that a region having similarity to the immunosuppressive region of retroviruses has the same disulfide connectivity as Moloney murine leukemia virus (Mo-MLV). Biochemical analyses indicate that the correctly disulfide bonded form (Ebo-95.3) has a trimeric helical-hairpin structure. A proteolytically stable fragment of Ebo-95.3 containing core 74 residues (Ebo-74) was then crystallized, and its x-ray structure determined at 1.9-Å resolution. Ebo-74 forms a trimer in which a long, central three-stranded coiled coil is surrounded by shorter C-terminal helices packed in an antiparallel orientation into hydrophobic grooves on the surface of the coiled coil. Our results confirm the previously anticipated structural similarity between the Ebola GP2 ectodomain and the core of the transmembrane (TM) subunit from oncogenic retroviruses. The Ebo-74 structure likely represents the fusion-active conformation of the protein, and its overall architecture resembles several other viral membrane-fusion proteins, including those from HIV and influenza.

Thesis Supervisor: Peter S. Kim

Title: Professor of Biology

Acknowledgments

I owe all that I have accomplished to my parents, Barbara T. Schneider and Paul F. Schneider, Sr. for creating a nurturing home environment and instilling a strong sense of integrity. I am forever grateful for my future wife Katherine T. Bacon for her love, support, and understanding through the peaks and valleys of this research project. I am indebted to Dr. Peter S. Kim for providing the resources and having the patience to allow me to find my own path and Dr. Vladimir N. Malashkevich for solving the X-ray crystal structure of Ebo-74. This project would not have been completed without the help of Michael A. Milhollen who made essential constructs and taught me everything I know about purifying pMmHb-fusion proteins, Peg L. McNally who purified the Ebo-74 protein for the crystallographic studies, and James X. Pang who had the patience to operate the often finicky LCQ. Dr. Deborah Fass, Dr. Lawren C. Wu, Dr. David C. Chan, and Dr. Michael J. Root were members of the Kim lab whose critical evaluations, support, and mentorship helped me through my years at MIT. I would also like to thank my fellow classmates in the Kim Lab, Debra Eckert, John Newman, and David Akey, for important motivation, comic relief, and assistance throughout the course of our years at MIT. Finally, I am thankful for my roommate Jeffrey M. Trimarchi whose late night conversations, teaching assistant antics, and stress release sessions helped me immensely.

Introduction

Ebola virus, a filovirus, is the causative agent responsible for lethal outbreaks of hemorrhagic fever (1, 2). The non-segmented, negative-stranded RNA virus encodes seven regulatory and structural proteins (3, 4) including a single surface transmembrane glycoprotein (GP). The Ebola GP is encoded as two open reading frames. One produces the secreted form (sGP), 50 to 70 kDa, that is synthesized in large amounts early in infection, while an alternative transmembrane form (GP), 120-150 kDa, arises from transcriptional editing and is incorporated into the virion (5, 6). The sGP likely affects the host response to infection by interacting with the Fc γ receptor III of neutrophils, whereas the transmembrane GP is believed to mediate fusion by binding to endothelial cells and may contribute to the hemorrhagic symptoms of the disease (7). Although synthesized as a single precursor molecule, the 676 amino acid transmembrane GP is cleaved into two subunits (GP1 and GP2) by a cellular convertase furin (8, 9).

Despite belonging to different viral classifications, there are shared features among the GP1/GP2 complex in Ebola, the gp120/gp41 complex in HIV-1 and SIV and the SU/TM complex in retroviruses. The external component is responsible for cell binding while the transmembrane portion is necessary for fusion of the viral and cellular membranes (10). Coiled coil regions adjacent to fusion peptides have been identified in viral transmembrane proteins and similar α -helix models have been proposed for the HIV-1 gp41, Avian sarcoma virus TM, and Ebola GP2 transmembrane subunits (11, 12, 13, 14). The three dimensional structures of the low-pH induced conformation of the Influenza hemagglutinin (HA) 2 (15), a protease resistant fragment of the HIV-1 (16, 17, 18) and SIV (19, 20), a portion of the HTLV-1 transmembrane region (21), and a fragment of the Moloney murine leukemia virus (Mo-MLV) TM protein (22) revealed that a trimeric coiled coil lie at the center of these molecules. In all but the Mo-MLV structure, α -helical regions bound in an antiparallel orientation into the hydrophobic grooves formed by the central

trimeric helices. Furthermore, the SIV and HTLV-1 structures demonstrate the helical hairpin structure in which the inner and outer helices are linked by a loop region. The Ebola virus GP2 transmembrane protein has many similarities to these structures (Figure 1). There is a fusion peptide (23) N terminal to a region predicted to have coiled coil properties (14, 24) followed by a loop region similar to the immunosuppressive regions with a CX₆CC consensus sequence identified in retroviruses including Mo-MLV and Rous sarcoma virus (13, 22, 25). There is also a region between the immunosuppressive-like domain and the hydrophobic membrane anchoring region with α -helical character (13).

Protein dissection has been used to identify key portions of viral envelope glycoproteins for further characterization by biochemical and biophysical means. The Kim lab has successfully identified and characterized stable proteolytic fragments of the ectodomains of Mo-MLV (22, 26), HIV-1 (16, 27, 28), and SIV(19). First, we develop a peptide model expressed recombinantly in *E. coli* which contains a portion of the viral glycoprotein ectodomain between the hydrophobic transmembrane and fusion peptides. These hydrophobic regions are typically left out to avoid aggregation of the model peptide. Next, the peptide is subjected to limited proteolysis to identify stable fragments. Biochemical studies including analytical centrifugation and circular dichroism are used to characterize these fragments which are then used for X-ray crystal structure determination.

This study represents the protein dissection of a 95 amino acid fragment of the Ebola virus GP2 (Figure 1). After expressing and purifying the fragment from *E. coli*, we identified the construct whose disulfide connectivity is the most stable based on thermal denaturation and proteolysis studies. This construct, Ebo-95.3, is trimeric with significant α -helical content according to analytical ultracentrifugation and circular dichroism studies. In addition, proteolysis experiments indicate that the Ebo-95.3 folds upon itself to form a helical hairpin structure. Proteolysis also revealed a stable fragment, Ebo-74, whose anion binding properties were similar to the ectodomain of Mo-MLV. This GP2 core forms a

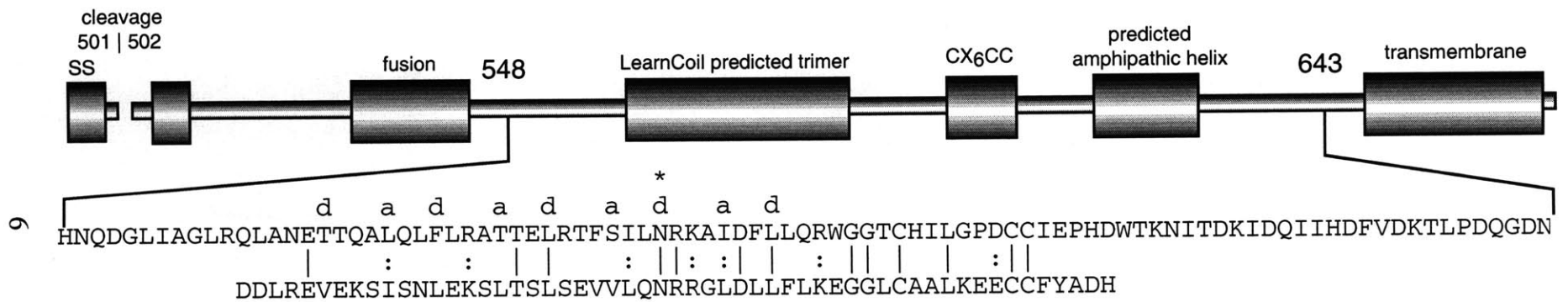


Figure 1. Representation of the Ebola virus glycoprotein. Sections of interest are noted by thicker cylinders. Ebo-95 sequence is shown aligned with Mo-55 sequence. Vertical lines and colons indicate identical and similar residues, respectively. The a and d notations represent positions of the helical wheel for Mo-55. The * denoted a conserved asparagine. CX₆CC - conserved cysteine motif evidenced in retroviruses (13) SS - signal sequence.

trimeric, helical hairpin structure with the same disulfide connectivity and similar anion binding properties as seen in the Moloney murine leukemia virus.

Results

To begin the biochemical studies on the Ebola virus glycoprotein, a 95 amino acid fragment of the GP2 protein was expressed in *E. coli* with a plasmid containing a synthetic gene. The vector contained residues 549-643 from the Ebola Zaire GP (6) with Cys-556 changed to Ala to avoid nonspecific disulfide bond formation. There were three other cysteine residues (Cys-601, Cys-608, and Cys-609) within a region having high similarity to the immunosuppressive-loop region of retroviral TM proteins (13). The analogous region in Mo-MLV has a disulfide connectivity from the first to the second cysteine residues (26). Based on homology, the Ebola virus would likely have a similar disulfide connectivity; however, we wanted to experimentally verify this hypothesis. To determine the correct disulfide connectivity for cysteine residues 601, 608, and 609, we created different disulfide permutations of Ebo-95. Each cysteine, 601, 608, and 609, was separately mutated to an alanine and will be referred to as Ebo-95.1, Ebo-95.2, and Ebo-95.3, respectively. For example, Ebo-95.1 has the first cysteine mutated to an alanine; therefore, the second and third cysteines should form a disulfide bond. To prevent structural inhibition of non-native disulfide bond formation, each protein was allowed to oxidize in 6 M guanidine hydrochloride (GuHCl). The proteins were then allowed to refold by the removal of GuHCl before their stability was tested through thermal denaturation and proteolysis (Figures 2 and 3). The mutant with the disulfide bond between 601 and 608, labeled Ebo-95.3, was the most stable with a T_m of 88°C. The other variants, Ebo-95.1 and Ebo-95.2, were significantly destabilized with T_m values of 78°C and 84°C, respectively. Furthermore, each Ebo-95 variant was subjected to subtilisin digestion at 25°C with aliquots removed at 20 minutes, 1 hour, 3 hours, 9 hours, and 27 hours. These digested samples were analyzed by SDS-PAGE. The Ebo-95.3 construct was the most resistant to proteolysis with a subspecies still present after 27 hours of digestion. This subspecies was identified as residues 555-633 of the full length GP by

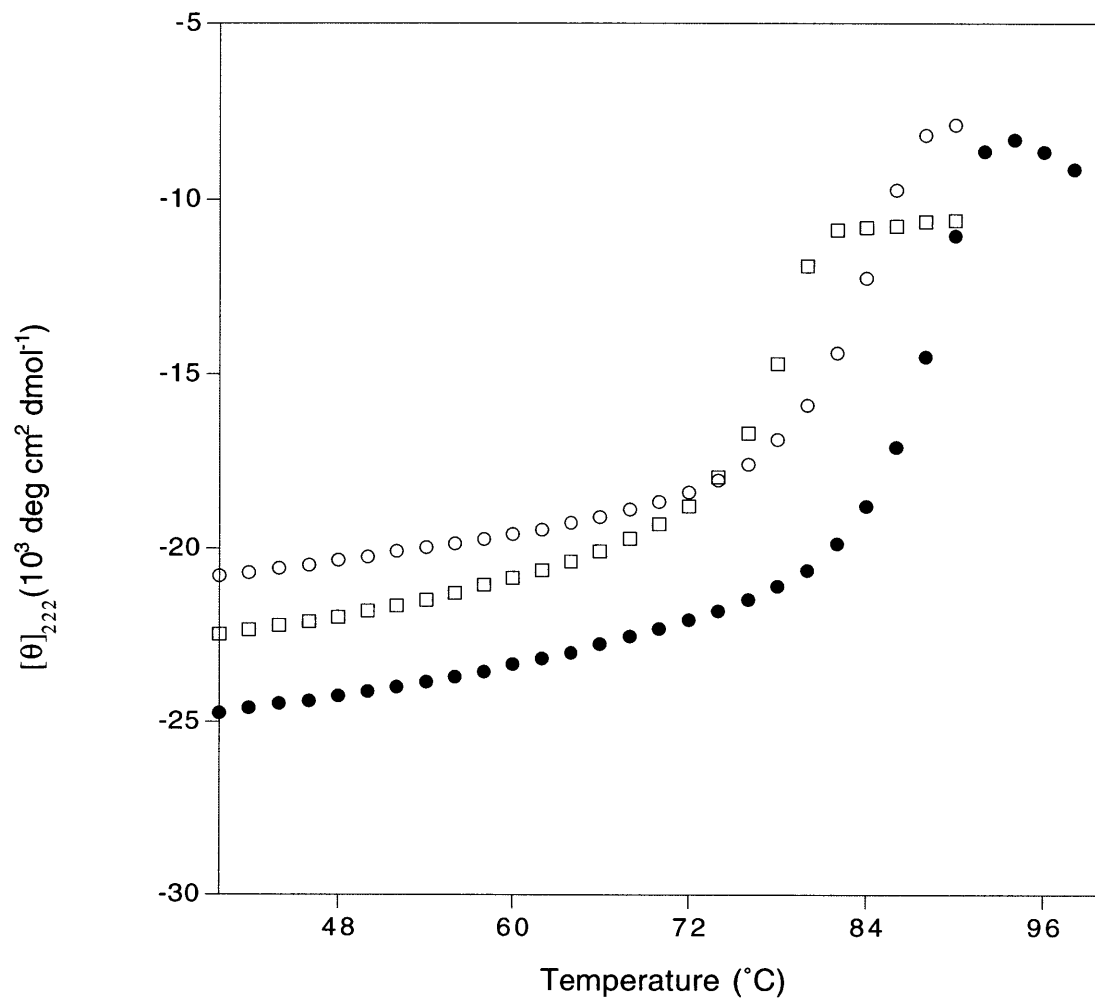


Figure 2. Thermostability of Ebo-95 cysteine variants. Circular dichroism of Ebo-95.1 (open squares), Ebo-95.2 (open circles), and Ebo-95.3 (closed circles) were measured at 222 nm at 2°C intervals. Data collected from 40°C to 100°C is shown.



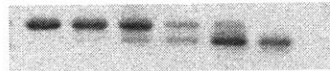
<u>Construct</u>	<u>Subtilisin Digest</u>	<u>Melting point</u>
	hours at 25°C 0 0.3 1 3 9 27	
EBO95.1		78°C
EBO95.2		84°C
EBO95.3		88°C

Figure 3. Stability of Ebo-95 cysteine variants. Middle column displays the SDS-PAGE results of 0.1 mg/mL subtilisin digesting 1 mg/mL of the respective construct for 0, 0.3, 1, 3, 9, and 27 hours at 25°C. Ebo-95.3 continues to have a significant subspecies present even after 27 hours, while Ebo-95.1 is rapidly degraded and Ebo-95.2 has no detectable subspecies present after 27 hours. This species was identified as fragment 555-633 by mass spectrometry. The melting point (T_m) for each protein was calculated as the highest point on the first derivative versus temperature plot.

mass spectrometry. However, the Ebo-95.2 construct had no remaining species after 27 hours, and the Ebo-95.1 construct was completely degraded in 20 minutes. As demonstrated by both the thermal denaturation and proteolytic digestion, the Ebo-95 variant with a Cys-601 to Cys-608 disulfide bond (Ebo-95.3) is the most stable. These results are comparable with similar experiments of Mo-MLV which also showed that the most stable disulfide connectivity was between the first two cysteine residues of the immunosuppressive loop region (26).

The model for the three dimensional structure of Ebo-95 based on homology and some experimental results is a trimeric helical hairpin confirmation in which two antiparallel helices are connected by a loop region (13, 14, 26). First, we determined the oligomerization state of Ebo-95.3 by analytical centrifugation. The data were fit to a single ideal species model and were consistent with a trimer model over a 10-fold concentration range (Figure 4). The residuals between the data and the linear fit were random and did not indicate a systematic trend. In order to test whether or not the construct contains a loop bringing the N- and the C-termini into proximity, an additional construct was created (Figure 7). As noted before, subtilisin digestion of Ebo-95.3 leads to a species that is truncated at the N-terminus by 7 residues and the C-terminus by 12 residues. The new construct, Ebo-81, has an additional 7 residues removed from the N-terminus of Ebo-95.3 and consists of residues 563 to 643 of the full length Ebola glycoprotein. By subjecting Ebo-81 to subtilisin digestion, we should be able to see whether there is additional C-terminal cleavage compared with the digestion of Ebo-95.3. Additional cleavage occurring at the C-terminus would support the hypothesis that the N- and C-termini interact with one another since shortening of the N-terminal coil would lead to an exposed region of the C-terminus. Ebo-95.3 was subjected to subtilisin digestion for 9 and 81 hours (Figure 5). Ebo-81 was subjected to subtilisin digestion for 4, 32, and 256 minutes (Figure 6). Each digestion was analyzed by HPLC and purified fragments were analyzed by electrospray mass spectrometry to identify masses of each peak on the HPLC chromatogram. These

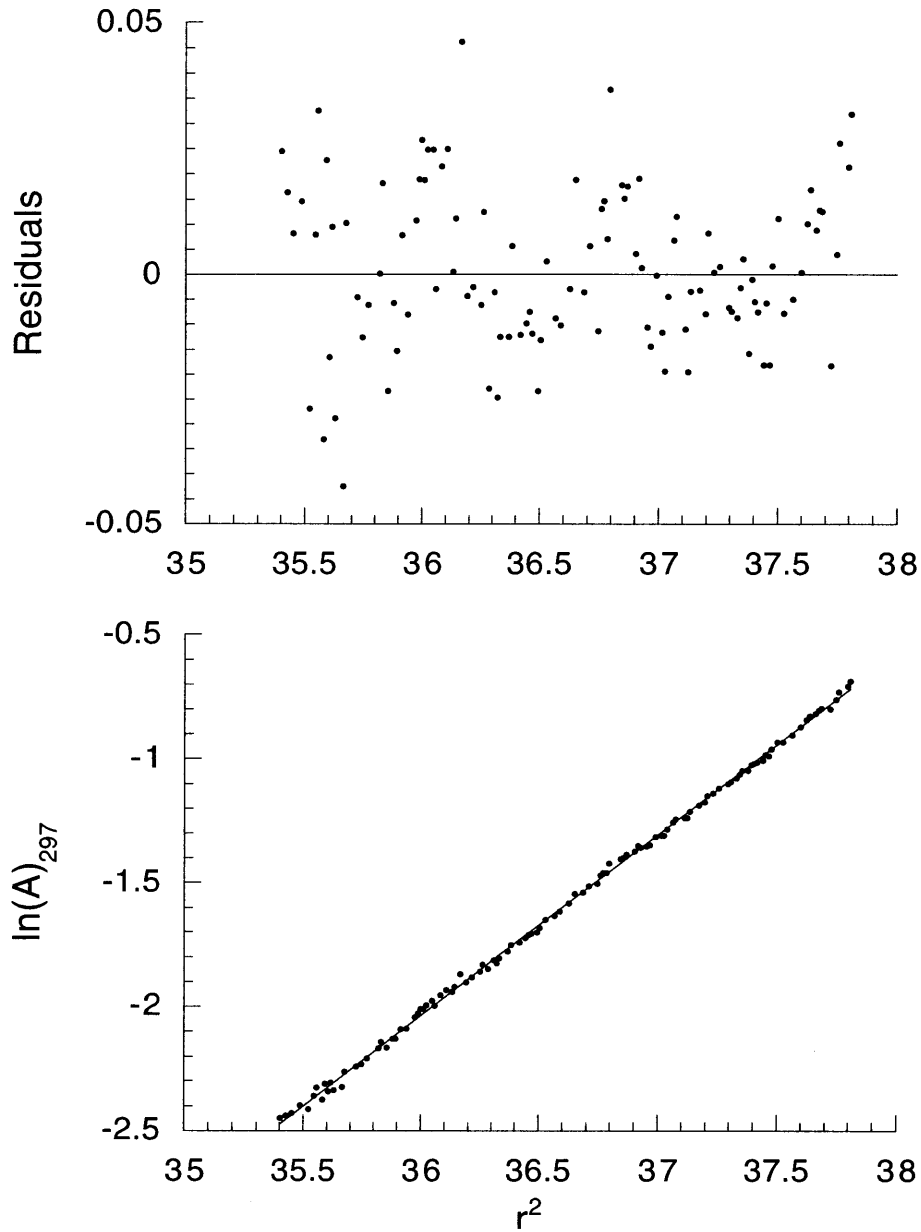


Figure 4. Equilibrium sedimentation of Ebo-95.3. Representative data is plotted as the $\ln(\text{absorbance})_{297}$ against the square of the radius from the axis of rotation. The slope is proportional to the molecular mass of the protein (see Materials and Methods). The line represents the calculated model for a trimeric species of Ebo-95.3. The residual plot displays the difference between the calculated model and observed values.

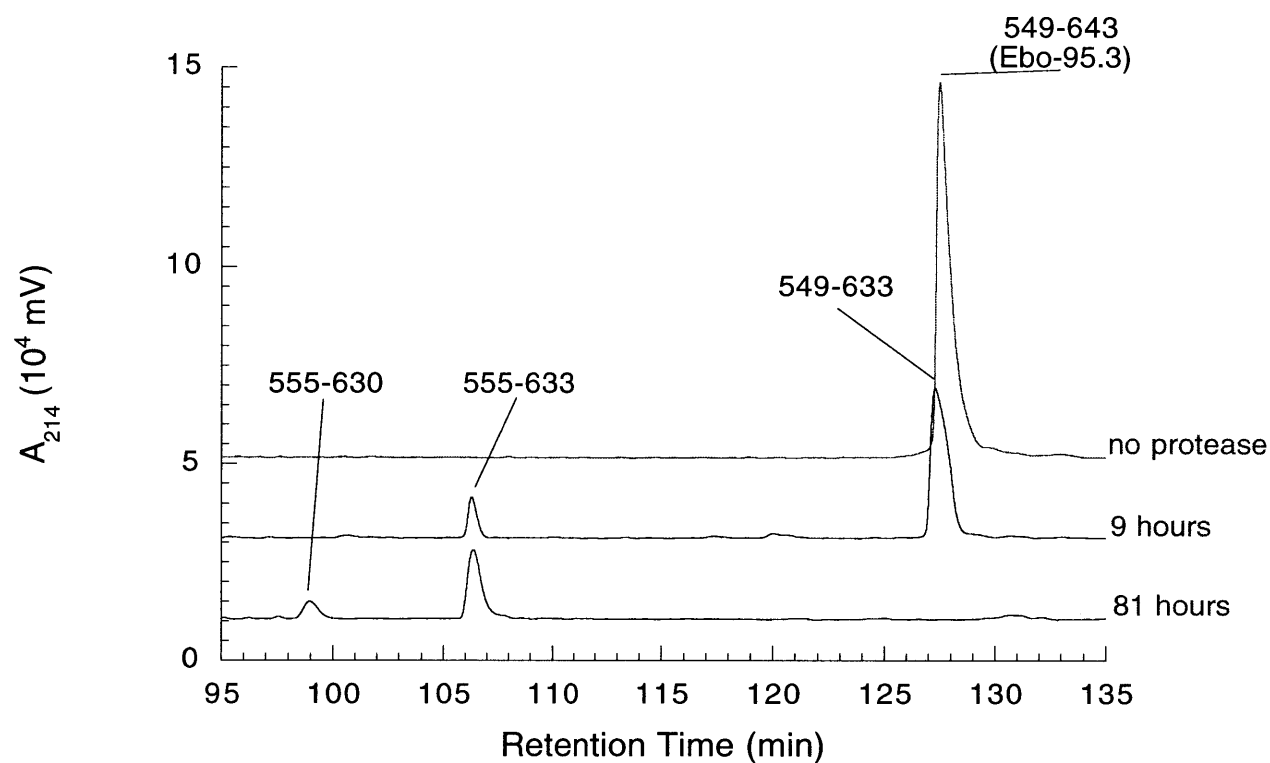


Figure 5. Subtilisin digestion of Ebo-95.3. This figure shows the elution of digested Ebo-95.3 from the reverse-phase HPLC. Ebo-95.3 was digested with subtilisin for 0 hours, 9 hours, and 81 hours in 20 mM Tris pH 8.0, 10 mM CaCl₂ at room temperature. The elutant's absorbance at 214 nm from 95 to 135 minutes is represented. Fragments were assigned by matching with possible fragment masses predicted by a computer algorithm (Wolf & Kim, unpublished). When assigning fragments, the next closest predicted mass was never closer than 3 daltons to the assigned mass. Labels represent the numbering of the full length Ebola glycoprotein (GP). Fragments corresponding to peptides mentioned in this paper are labeled with parentheses. Shorter fragments not depicted were present in much smaller quantities and could not be accurately assigned.

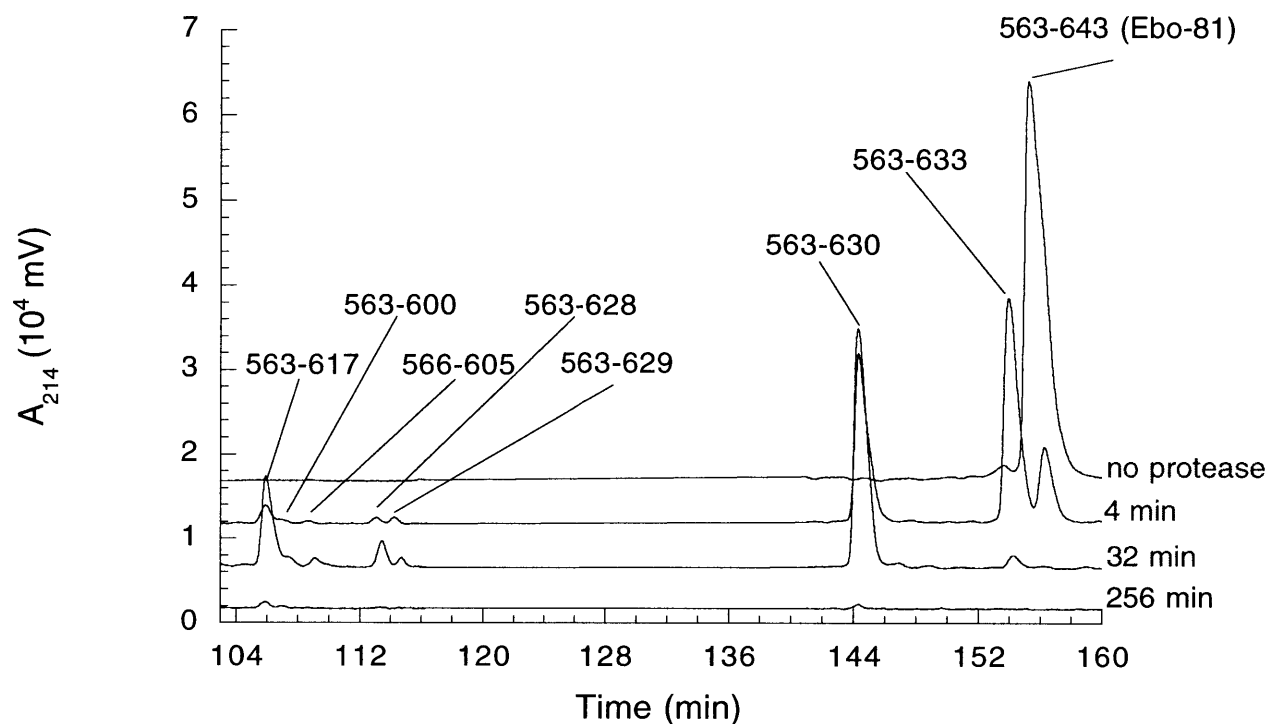
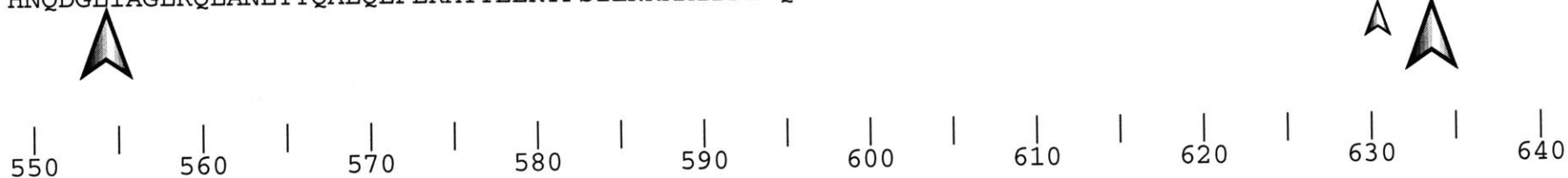


Figure 6. Subtilisin digestion of Ebo-81. This figure shows the elution of digested Ebo-81 from the reverse-phase HPLC. Ebo-81 was digested with subtilisin for 0 hours, 4 minutes, 32 minutes, and 256 minutes in 20 mM Tris pH 8.0, 10 mM CaCl₂ at room temperature. The elutant's absorbance at 214 nm from 104 to 160 minutes is represented. Fragments were assigned by matching with possible fragment masses predicted by a computer algorithm (Wolf & Kim, unpublished). When assigning fragments, the next closest predicted mass was never closer than 3 daltons to the assigned mass except for fragment 563-617 which has an identical mass to fragment 564-618. Labels represent the numbering of the full length Ebola glycoprotein (GP). Fragments corresponding to peptides mentioned in this paper are labeled with parentheses. Shorter fragments not depicted were present in much smaller quantities and could not be accurately assigned.

EBO95.3

HNQDGLIAGLRQLANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCAIEPHDWTKNITDKIDQIIHDFVDKTLPDQGDNDN



15

EBO81.3

NETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCAIEPHDWTKNITDKIDQIIHDFVDKTLPDQGDNDN



Figure 7. Location of subtilisin cleavage sites on Ebo-95.3 and Ebo-81. Arrow represent cleavage sites identified through proteolysis and assignment of identified masses. Large arrows represent fragments which areas were greater than 25% of the total area per digestion time, while smaller arrows represent fragments that were less than 25% of the total area per digestion time.

masses were assigned to fragment lengths using a computer algorithm (Wolf and Kim, unpublished). Table 1 indicates experimental mass values obtained for a particular peak along with the expected mass for the fragment length shown in parentheses. The identified cleavage sites are graphically displayed in Figure 7. The removal of 14 residues N-terminal residues significantly destabilizes Ebo-95.3 since the Ebo-81 construct is almost completely digested in 256 minutes, while the Ebo-95.3 construct still has stable fragments present after 4860 minutes (81 hours). Despite the rapid degradation there is a significant subspecies that occurs during the digestion of Ebo-81. This species has a cut after approximately residue 617. The exact location could not be assigned since proteolysis after residue 617 has the same mass as the fragment with cuts after both residue 564 and residue 618. The additional cleavage at the C-terminus of the shortened Ebo-81 construct supports that hypothesis that Ebo-95.3 folds back upon itself so that the N and C-termini interact.

An alternative hypothesis is that the protein becomes unfolded, especially since the digestion of Ebo-81 is at least an order of magnitude faster than Ebo-95.3. As demonstrated by the circular dichroism spectrum of Ebo-81 (Figure 8), the protein does contain significant amount of secondary structure, in particular, α -helix. The CD spectrum was measured for each construct from 260 nm to 200 nm (Figure 8). The $[\theta]_{222}$ values were -21,500 and -17,900 for Ebo-95.3 and Ebo-81 respectively. Assuming that a peptide that is 100% helical has a $[\theta]_{222}$ of -33,000 deg cm² dmol⁻¹ (29) with a baseline value of -2,000 deg cm² dmol⁻¹ (30), each had a significant helical characteristic with 59% and 48% α -helix or approximately 56 and 39 helical residues for Ebo-95.3 and Ebo-81, respectively. Furthermore, there is only a loss of approximately 17 residues that contribute to helical propensity. As noted before the N-terminus of the construct is expected to be a coiled coil so that a significant portion of the loss can be contributed to the removal of 14 helical residues from the N-terminus.

Since Ebo-95.3 did not form well diffracting crystals, we wanted to obtain a proteolytic fragment of Ebo-95.3 that eliminated unstructured regions that could inhibit

Table 1. Analysis of Fragments. Experimental values are the average of at least 5 separate mass determination. Expected values and fragments in parentheses are closest masses from prediction program (Wolf and Kim, unpublished). The next closest fragment and mass within 5 kDa is also shown. Note that mass 6339.3 could not be accurately assigned since fragments 563-617 and 564-618 have identical masses.

Ebo-81 Digestion

Experimental	Expected	Next Closest within 5 kDa
9035.7	9035.9 (563-643)	
8236.3	8235.8 (563-633)	
7893.0	7393.4 (563-630)	7891.5 (567-634)
7745.4	7746.2 (563-629)	7749.3 (565-631)
7631.1	7631.1 (563-628)	7632.1 (564-629)
6339.3	6339.6 (563-617, 564-618)	6637.7 (568-622)
4603.0	4602.7 (566-605)	4603.5 (585-624)
4424.1	4424.3 (563-600)	4421.3 (578-615)

Ebo-95.3 Digestion

Experimental	Expected	Next Closest within 5 kDa
10793.0	10793.7 (549-643)	
9722.7	9723.6 (549-633)	9726.7 (554-638)
9058.0	9058.9 (555-633)	9062.7 (565-643)
8715.6	8716.4 (555-630)	8719.4 (565-640)

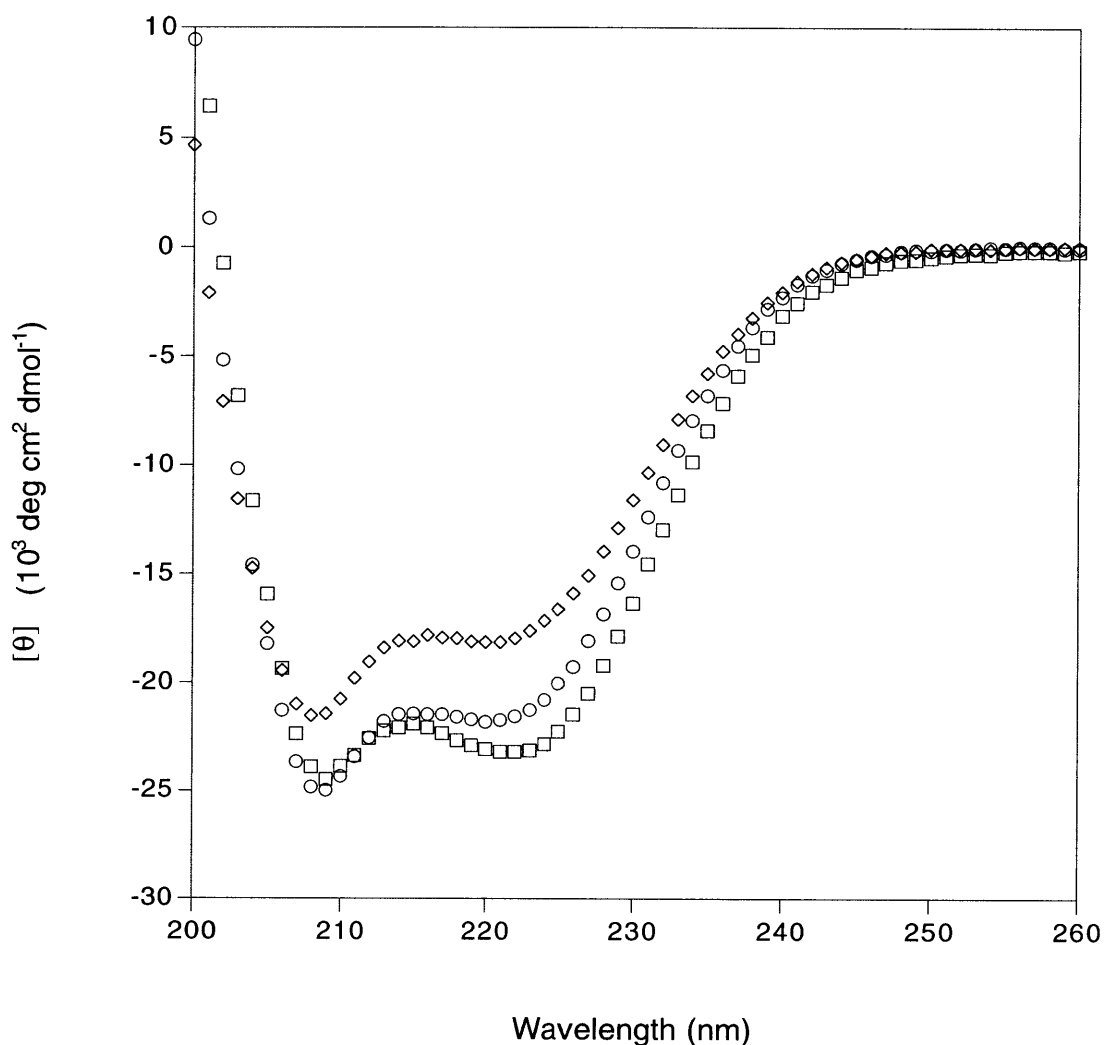


Figure 8. CD spectrum of Ebo-95.3 (open circles), Ebo-81 (open diamonds), and Ebo-74 (open squares). All protein concentrations are 25 μM in 100 mM PO_4^- pH 7.2, 150 mM NaCl at 25°C. Scans are from 260 to 200 nm, 5 second averaging time, averaged over 3 scans. The $[\theta]_{222}$ values for Ebo-95.3, Ebo-81, and Ebo-74 were -21,500, -17,900 and -23,200 $\text{deg cm}^2 \text{dmol}^{-1}$, respectively. Assuming -33,000 $\text{deg cm}^2 \text{dmol}^{-1}$ (29) for 100% α -helix content and a baseline of -2,000 $\text{deg cm}^2 \text{dmol}^{-1}$ (30) at $[\theta]_{222}$ the percentage of α -helix for Ebo-95.3 is 59%, 48% for Ebo-81, and 64% for Ebo-74. This corresponds to approximately 56, 39, and 48 helical residues for Ebo-95.3, Ebo-81 and Ebo-74, respectively.

crystal growth. Ebo-95.3 was subjected to proteolysis by subtilisin and protease K for long periods of time. Digestion led to the truncation of Ebo-95.3 at both the N- and C-termini. After 81 hours of digestion with subtilisin, two significant fragments remained representing residues 555-630 and 555-633 of the full length Ebola glycoprotein (Figure 5). These fragments represent a seven amino acid truncation from the N-terminus and either a ten or thirteen amino acid truncation from the C-terminus of Ebo-95.3. A two hour digestion with protease K (Figure 9), not only identified fragments 555-630 and 555-633, but also 557-630 as a significant fragment. Since the 557-630 fragment was two residues shorter than the other two fragments eliminating the cysteine to alanine mutation at 556, we wanted to determine whether or not the truncation eliminated unstructured regions that could be deleterious to crystal growth.

One method to determine a change in the overall secondary structure was to obtain the circular dichroism spectrum of both constructs (Figure 8). The CD spectra indicate that the truncation leads to an increase in α -helical content with 64% for Ebo-74 and 59% for Ebo-95.3. However, converting percentage to the approximate number of helical residues for each peptide, results in 56 helical residues for Ebo-95.3 and 48 helical residues for Ebo-74. While the elimination of 21 amino acids from Ebo-95.3 caused a reduction in the number of residues with helical contribution there was also a significant reduction in the number of unstructured residues.

The conserved asparagine at residue 586 is believed to bind to an anion due to its homology and alignment with Mo-55. To test this hypothesis, Ebo-74 was thermally denatured under different monovalent anions (Figure 10). Chloride and bromide ions result in substantial stabilization of Ebo-74, whereas fluoride has very little effect. The apparent $T_m = 84^\circ\text{C}$, 84°C , 79°C , or 77°C in the presence of 5 mM NaCl, 5 mM NaBr, 5 mM NaF, or no added salt, respectively. These results are very similar to those obtained by studies on Mo-55 (22), and suggest that the anion binding site is conserved in Ebo-74.

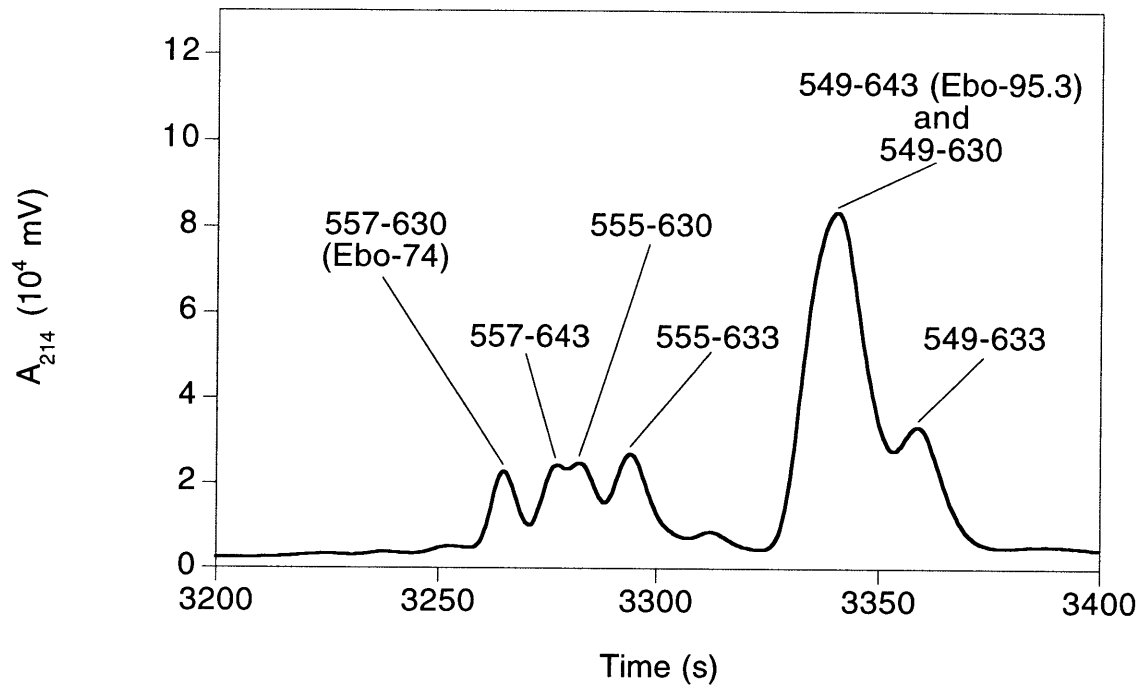


Figure 9. Protease K digest of Ebo-95.3. This figure shows the reverse-phase HPLC separation of Ebo-95.3 digested with protease K for 2 hours in 20 mM Tris pH 8.7. The elutant's absorbance at 214 nm from 3200 to 3400 s is represented. Fragments were assigned by matching with possible fragment masses predicted by a computer algorithm (Wolf & Kim, unpublished). When assigning fragments, the next closest predicted mass was never closer than 3 daltons to the assigned mass; most were greater than 5 daltons. Labels represent the numbering of the full length Ebola glycoprotein (GP). Fragments corresponding to peptides mentioned in this paper are labeled with parentheses. Shorter fragments not depicted were present in much smaller quantities and could not be accurately assigned.

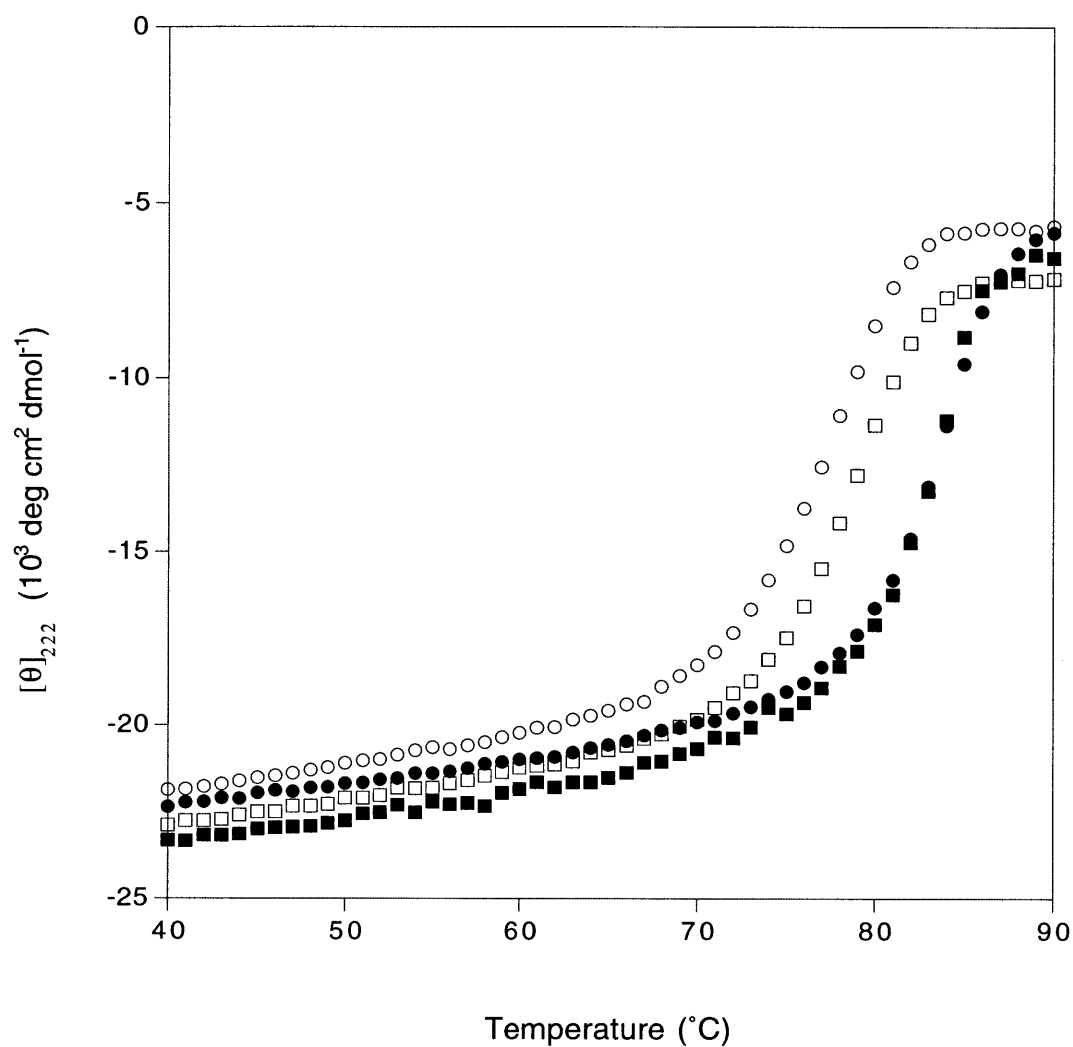


Figure 10. Thermostability of Ebo-74 in different monovalent anions. Circular dichroism of Ebo-74 in 10 mM acetate pH 4.5 and no additional salt (open circles), 5 mM NaF (open squares), 5 mM NaBr (closed squares) and 5 mM NaCl (closed circles) were measured at 222 nm at 1°C intervals. Data collected from 40°C to 90°C is shown.

The Ebo-74 construct formed well diffracting crystals and we were quickly able to obtain an electron density map of good quality (Figure 12) and reveals the positions of all the amino acid residues except for a few disordered side chains at the chain termini and on the protein surface (31).

As expected, Ebo-74 forms a homotrimer. Each of the three polypeptide chains folds into a helical-hairpin conformation, in which two antiparallel helices are connected by a loop region (Figure 11). The N-terminal helices from each monomer form a central, three-stranded coiled coil. This coiled-coil core includes approximately 35 residues (561 to 595) from each chain (the two most N-terminal residues are not well defined in the electron density maps and next two are in a random-coil conformation). Shorter C-terminal helices (residues 615 to 629) pack in an antiparallel manner into hydrophobic grooves on the surface of the coiled-coil core (the C-terminal Phe-630 residue is disordered in all chains). In the loop region connecting the N- and C-terminal helices, a disulfide between Cys-601 and Cys-608 links a short α -helix (residues 600 to 604) and a short 3_{10} -helix (residues 606 to 610). Additionally, the loop region between the 3_{10} -helix and C-helix (residues 611 to 614) are in an extended conformation (31).

Visual analysis and superposition of the Ebo-74 coiled-coil structure with the structures of regular three-stranded coiled coils (e.g., GCN4-pIQI, ref.(32); or Mo-55, ref.(22)) reveal clear irregularities at the N-terminal end of the Ebo-74 coiled coil (Figure 13). In terms of superhelix parameters (33, 34, 35), this deviation can be described as an unwinding of the superhelix (decrease in ω_0), an increase in the α -position orientation angle (ϕ), with a decrease in the superhelix crossing angle (χ). Indeed, in the first three helical turns of the Ebo-74 coiled coil, the helices are almost straight. As a result, the helices are further apart than in the regular coiled-coil region of Ebo-74 (supercoil radius $R_0= 7.0 \text{ \AA}$ vs. 6.3 \AA), and knobs-into-holes packing interactions, a hallmark of coiled coils, are lost (31).

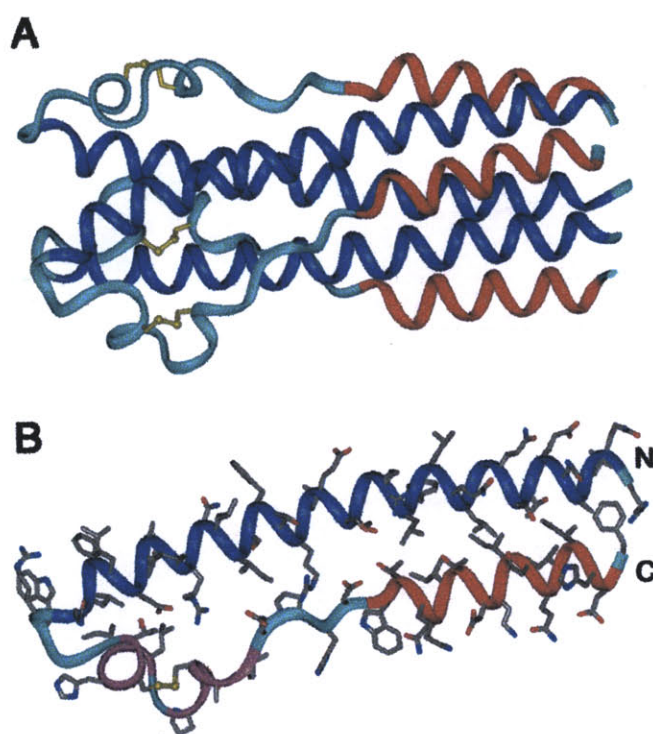


Figure 11. Ebo-74 forms a trimer-of-hairpins structure. (A) A side view of the Ebo-74 structure. N-helices (dark blue) constitute a central coiled coil and C-helices (red) pack into hydrophobic grooves on the surface of the coiled coil. Disulfides within the loop regions (light blue) are depicted in yellow. (B) Each monomer has an α -helical hairpin conformation. The last three turns of the N-helix, followed by the short α - and 310-helices (magenta), correspond to the immunosuppressive motif region in oncogenic retroviruses. Figures generated using Insight II (66).

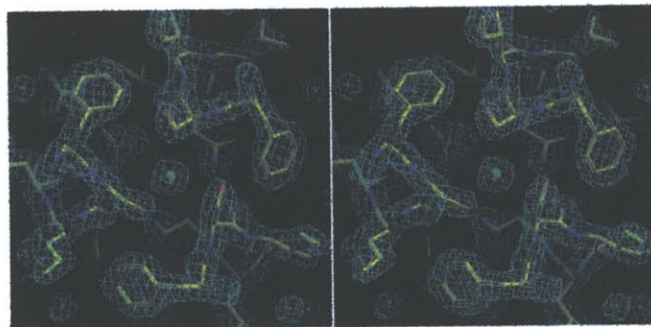


Figure 12. Chloride-binding site within the central coiled coil of Ebo-74 as visualized in a stereo view of the final 2Fo-Fc electron density map. A chloride ion and water molecules are shown as green and red spheres, respectively. The map was contoured at 1.5 σ .

An irregularity at the N-terminal end of the Ebola GP2 coiled coil also exists in the crystal structure of the GP2 fragment fused to an N-terminal GCN4 trimeric coiled coil (36). Our results indicate that this irregularity is not imposed by the GCN4 coiled coil. The N-terminal ends, however, are where the most significant differences exist between the structures of the Ebola-GCN4 fusion protein (36) and Ebo-74. Specifically, the Ebola-GCN4 fusion protein (36) has nine additional helical residues in an α -helical conformation at the N-terminus, as compared to Ebo-74 structure. (Four of these nine are present in Ebo-74 as disordered or non-helical residues, and the remaining five are not present). It is possible that the GCN4 coiled coil has forced these residues into an α -helical conformation in the Ebola-GCN4 fusion, or alternatively, it could be that end-effects in the Ebo-74 construct cause these residues to unfold. In addition, the irregularity at the N-terminal end of the coiled coil has been described as a “stutter” in the Ebola-GCN4 fusion protein (36, 37), but the lack of knobs-into-holes packing at the N-terminal end of Ebo-74 makes the definition of the heptad-repeat position irrelevant in this case (31).

Interestingly, the otherwise quite similar Mo-55 structure has a very regular structure throughout its coiled coil (22). The irregularity in the Ebo-74 coiled coil might be determined by the sequence of the coiled-coil region or may result from interactions of the C-terminal helices with the N-terminal coiled coil, since the C-terminal helices are not present in the Mo-55 fragment. However, as is demonstrated by the HIV-1 and SIV gp41 core structures, C-terminal helices can pack against a central N-terminal coiled coil without distorting the coiled-coil structure (16, 17, 19, 20).

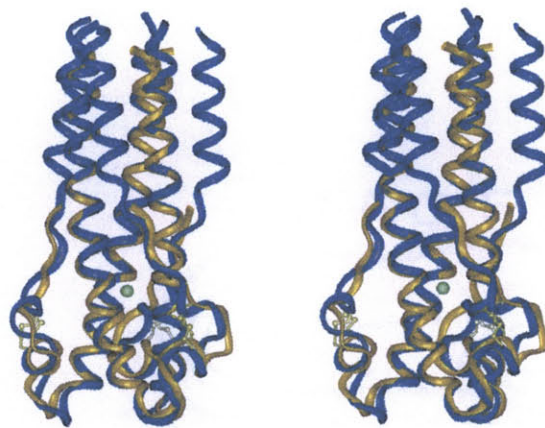


Figure 13. Remarkable similarity between filovirus and retrovirus membrane-fusion proteins. A stereo view of the superposition of the Ebo-74 (blue) and the Mo-55 (yellow) structures is shown. The location of the chloride ion is almost identical in both structures. The largest deviations occur at the N-termini and in the loop regions of the Ebo-74 and Mo-55 structures. The RMSD values for the Ca atoms of Ebo-74 and Mo-55 monomers (and values for trimers in parentheses) are as follows: entire structure 1.56Å (1.84Å); coiled-coil cores 0.49Å (0.94Å); loop regions 1.91Å (2.33Å). Figures generated using O (67) and Insight II (66).

Discussion

Homology (13, 26) and some experimental evidence (14) has suggested that the Ebola virus glycoprotein was trimeric helical hairpin confirmation in which two antiparallel helices are connected by a loop region with a disulfide connectivity similar to retroviruses. Analytical centrifugation has confirmed that the core of the Ebola virus GP2 is trimeric. Disulfide permutation experiments have shown that the region of Ebo-95 with sequence similarity to the retroviral immunosuppressive region has the same disulfide connectivity as Mo-MLV. Proteolysis experiments of Ebo-95.3 and Ebo-81 which is truncated by 14 residues at the N-terminus indicate that the N- and C-termini interact to form a hairpin structure. Since these biochemical experiments were completed, two crystal structures have been published. One was the structure of Ebo-74 solved to 1.9 Å presented above (31), while the other was of a similar construct fused to a trimeric coiled coil solved to 3.0 Å (36). Both of these structures confirm the conclusions of the biochemical experiments presented here. The core of the Ebola virus glycoprotein is a trimeric helical hairpin in which two antiparallel helices are connected by a loop region with a disulfide connectivity similar to retroviruses. The N-terminal helices form a trimeric coiled coil and the shorter C-terminal helices pack into the hydrophobic grooves along to the surface of the coiled coil. Furthermore, both structures contain a strong X-ray scatterer associated with Asn-586 indicating the presence of an anion.

Although there is only 22% sequence identity between residues in Ebo-74 from Ebola virus and Mo-55 from Mo-MLV, the location of these identical residues suggested very similar structures (Figure 1) (13, 22). The similarity between the Ebo-74 and Mo-55 structures is striking (Figure 5), since no obvious evolutionary relationship exists between the filoviruses and the retroviruses. In addition, the corresponding receptor-binding domains (GP1 from Ebola or SU from Mo-MLV) show no sequence similarity.

Ebo-74 contains C-terminal residues not present in Mo-55; they form C-terminal helices that pack against the central N-terminal coiled coil. Thus, the overall architecture of Ebo-74 is a trimer of helical hairpins. The core of the Mo-MLV TM protein likely also forms a trimer of helical hairpins, since CD studies indicate that Mo-92, a longer fragment of the Mo-MLV protein that includes Mo-55 as well as an additional 37 C-terminal residues, contains additional helical residues (26).

The trimeric helical hairpin of the Ebola glycoprotein core has become a general feature among viral envelope proteins (Figure 14). Similar structures have been identified in a wide variety of viral types, including orthomyxovirus (15), paramyxovirus (38, 39), oncoretrovirus (21, 22), and lentivirus (16, 17, 18, 19, 20). There are significant variations in the loop regions that connect the two helices; however, the overall structures are strikingly similar. The N-terminal central trimeric core is connected to a hydrophobic fusion peptide and the C-terminal end is connected to a transmembrane region inserted in the viral membrane. These two ends are brought together by the trimeric hairpin structure. Since the fusion peptide is believed to insert into the cellular membranes, the hairpin structure likely brings the viral and cellular membranes into close proximity. (17, 40, 41, 42). How the apposition of the cellular and viral membranes leads to complete fusion is still unclear. Likely theories involve the distortion of the membrane and the formation of membrane pores via bundles of viral fusion proteins (43, 44, 45). Interestingly, vesicular fusion has been shown to be mediated by coiled coil formation of SNAP/SNARE proteins that cause the apposition of membranes (46, 47, 48, 49).

The trimer of hairpin structures are believed to be the fusion active state distinct from the native (non-fusogenic). In the best studied examples, the native structure undergoes a conformational change that leads to the activated fusogenic state. At low pH, an extended loop in the Influenza HA1/HA2 complex converts to a trimeric coiled coil necessary for fusion (15, 50). In HIV-1, binding of the surface receptors to the gp120/41 causes an unspecified conformational change leading to the activated state (42).

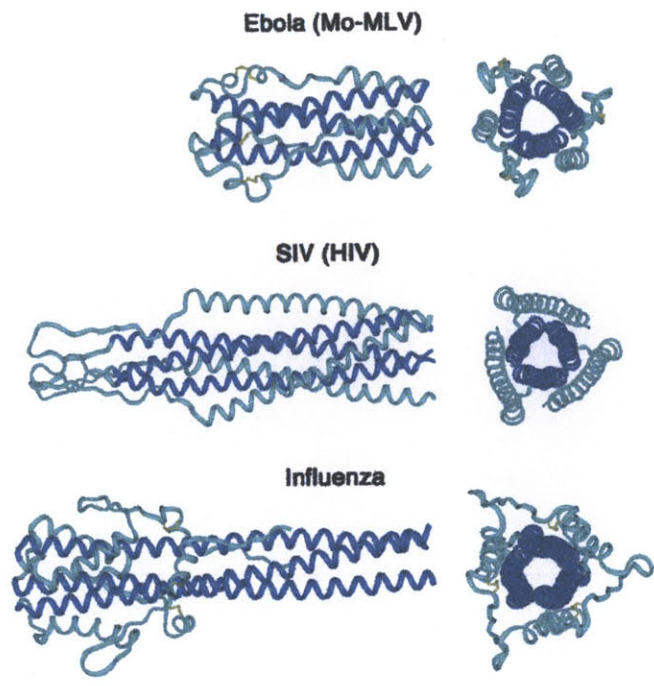


Figure 14. Trimer-of-hairpins structures in viral membrane-fusion proteins. Side and top views of the fusion proteins from three different groups of viruses are shown. Top; Ebola (31, 36) and Mo-MLV (22). Middle; SIV (19, 20) and HIV-1 (16, 17, 18). Bottom; Influenza (15). The N-terminal coiled-coil cores are depicted in dark blue. In all structures, the trimer-of-hairpins conformation brings the N- and C-termini, and therefore the fusion peptides and transmembrane helices, into close proximity. Figure generated using Insight II (66).

The structural model identified by these studies and confirmed by crystallographic structures of the Ebola virus GP2 core could lead to therapeutic agents against Ebola infection. Peptides corresponding to the C-terminal helices of HIV-1 (19, 28, 51, 52, 53) and paramyxoviruses (38, 54, 55, 56) have been shown to inhibit virus fusion. In particular, HIV-1 synthetic peptides seem to interact in a dominant negative manner with a transient “pre-hairpin” intermediate of gp41. This complex prevents the gp41 from reaching the fusogenic state (42, 53, 57). Likewise, peptides corresponding to similar region of the Ebola virus GP2 core that pack against the central coiled coil may lead to a defense against Ebola virus infection.

Materials and Methods

Gene Construction. A synthetic gene, Ebo-95 was constructed of residues 549-643 from the Ebola virus glycoprotein (Genebank accession U23187) (6) using optimal codons for E. coli expression (58). Cys-556 was changed to an alanine to avoid non-specific disulfide bond formation. The constructed gene was inserted into the HindIII-BamHI restriction site of pMMHb (59), provided by Michael Milhollen. This plasmid expresses proteins as a fusion protein containing a modified TrpLE leader sequence (60) and a 9-histidine tag. The cysteine to alanine mutations, Ebo-95.1, 95.2, and 95.3, were created with site-directed mutagenesis at Cys-601, Cys-608, and Cys-609, respectively (61). The N-terminal truncation, Ebo-81, was created with site-directed mutagenesis of Met-548 to alanine and Ala-562 to methionine in the Ebo-95.3 construct. The truncation, Ebo-74, was created by amplifying the nucleotides corresponding to residues 557-630 from the pMMHb-Ebo-95.3 using PCR with primers 5'-GACTAAGCTTATGGGTCTGCGTCAACTGGC-3' and 5'-ATAGGATCCCTAGAAGTCGTGAATGATCTGATCG-3'. The amplified region was inserted into the HindIII-BamHI restriction site of pMMHb.

Protein Purification. The recombinant proteins were expressed in E. coli strain BL21 (DE3) pLys E with the T7 system (62). Each protein was grown from an overnight culture of a single colony shaking at 37°C in Luria broth, 100 µg/mL ampicillin, 15 µg/mL chloramphenicol. Cultures were induced with 0.4 mM IPTG at $A_{600} = 0.4 - 0.6$. Cells were harvested after 3 hours and lysed with sonication in 50 mM Tris pH 8.7, 15% glycerol, 10 µM MgCl₂, 10 µM MnCl₂, 10 µg/mL DNaseI. After centrifugation, the pellet was resuspended in 50 mM Tris pH 8.7, 1% NP-40, 1% deoxycholic acid, 1 mM EDTA, sonicated and centrifuged. The resulting pellet was resuspended in 6 M GuHCl, 50 mM Tris pH 8.7, sonicated and incubated at room temperature overnight. The solution was passed over a 10 mL Ni²⁺-NTA column (QIAGEN). Bound protein was eluted with 6

M GuHCl, 0.2 M acetic acid. Fractions were pooled, dialyzed against water, and lyophilized. Material was cleaved with CnBr (~ 20 mg/mL, 70% formic acid) for 1 hour at room temperature. To remove excess CnBr, samples were centrifuged under vacuum in a Savant Speed Vac Concentrator 200H attached to a Savant Refrigerated Vapor Trap 4104 for 2 hours, diluted 6-fold with water, dialyzed against water, then lyophilized. Material was resuspended in 6 M GuHCl, 12 mM imidazole, 16 mM 2-mercaptoethanol and again passed over a Ni²⁺-NTA column. The uncleaved fusion and cleaved TrpLE leader sequences remained bound to the column, while the desired products eluted immediately. Fractions were pooled, dialyzed against water, and lyophilized. Material was resuspended in 50 mM Tris pH 8.7, 100 mM DTT. After at least 1 hour incubation at room temperature, the Ebo proteins were purified by reverse-phase HPLC, using a Vydac C18 preparative column with a water/acetonitrile gradient of 0.1%/min. and 0.1% trifluoroacetic acid. Fractions were pooled, lyophilized, resuspended in 50 mM Tris pH 8.7, and stirred at room temperature exposed to air for 48 hours. For the disulfide permutation experiment, Ebo-95.1, 95.2, and 95.3 were stirred in 50 mM Tris pH 8.7, 6 M GuHCl at room temperature exposed to air for 48 hours then dialyzed against 5% HOAc. The Ebo proteins were purified via HPLC, as above. The masses of each purified Ebo protein was verified by mass spectrometry on a Voyager Elite MALDI-TOF mass spectrometer (Perceptive Biosystems, Framingham, MA). All observed masses were within 3 Da of expected value.

Analytical Centrifugation. Equilibrium ultracentrifugation studies (63) were carried out with a Beckman XL-A analytical centrifuge with an An-60 Ti rotor at 25°C. Ebo-95.3 in 250 μ L H₂O was dialyzed overnight at 4°C against 250 mL PBS. Protein concentration was determined from the extinction coefficient of tryptophan, tyrosine, and cysteine residues and the absorbance at 280 nm in 6 M GuHCl, 20 mM NaPO₄ pH 6.5 (64). Solutions of 100 μ M, 50 μ M, and 10 μ M were created by diluting appropriate volume of dialysis with dialysate. Solutions were centrifuged in a six-sector cell at rotor speeds of 15000 and 20000. Three cells were allocated for each protein concentration,

while the remaining cells were filled with dialysate as a reference. Data were collected for each speed and protein concentration at 296 nm, 292 nm, and 232 nm and analyzed using the following equation: $M = [2RT/(1-v)\rho\omega^2][d(\ln(c))/dr^2]$. For Ebo-95.3, v was $0.7336 \text{ cm}^3 \text{ g}^{-1}$ and ρ was 1.013 g cm^{-1} .

Proteolysis. All proteolysis reactions were performed with 1 mg/mL Ebo protein and 0.1 mg/mL protease at room temperature. Individual aliquots at specific time intervals were terminated with 55 mM phenylmethylsulfonyl fluoride solution in methanol to a final concentration of 5 mM. Subtilisin reactions were performed in the presence of 20 mM Tris pH 8.0, 10 mM CaCl_2 , while protease K reactions included only 20 mM Tris pH 8.7. Proteolysis fragments were visualized on tricine-sodium dodecyl sulfate-polyacrylamide gels (65). Proteolysis samples were analyzed by reverse phase HPLC on a Vydac C18 microbore column with a water/acetonitrile gradient of 0.1%/min. and 0.1% trifluoroacetic acid. After HPLC separation, elutant was directly injected into an Finnigan MAT (San Jose, CA) LCQ electrospray mass spectrometer. Due to small variations in the elution profile, chromatograms were corrected to align peaks with identical masses. This correction was never more than 2% of the retention time. Fragments were assigned by matching with possible fragment masses predicted by a computer algorithm (Wolf & Kim, unpublished). All assigned mass fragments were within 1 Da of predicted values.

Circular Dichroism (CD) Spectroscopy. All CD spectroscopy was performed using an AVIV (Lakewood, NJ) 62 DS spectrometer. All reported data were collected a dynode voltage of less than 450 volts. Wavelength scans from 300 to 200 nm were performed in a 1 mm path length cell with 25 μM protein and 100 mM NaPO_4 pH 7.2, 150 mM NaCl (PBS) at 25°C. Thermal denaturation of Ebo-95.3 cysteine mutants contained 5 μM protein in PBS and a 1 cm pathlength cell. The circular dichroism at 222 nm was averaged over 30 s for each temperature from 0 to 100°C in 2°C steps equilibrating 1.5 minutes at each step. Thermal denaturation of Ebo-74 contained 2.5 μM protein in 10 mM acetate pH 4.5 with no additional salt, 5 mM NaCl, 5 mM NaBr, or 5 mM NaF in a 1

cm pathlength cell. Acetic acid was used to pH the solution to pH 4.5. The circular dichroism at 222 nm was averaged over 30 s for each temperature from 0 to 90°C in 1°C steps equilibrating 2 minutes at each step. The T_m for each protein was calculated as the highest point on the first derivative versus temperature plot. The $[\theta]_{222}$ values between 2.5 μ M Ebo-74 in 10 mM acetate pH 4.5, 5 mM NaCl and 25 μ M Ebo-74 in PBS at 25°C varied by less than 0.7%. In estimating the percentage of α -helical residues, the $[\theta]_{222}$ of -33,000 deg cm² dmol⁻¹ was assumed to correspond to 100% α -helix content (29) with a baseline of -2,000 deg cm² dmol⁻¹ (30). Protein concentrations were determined from the extinction coefficient of tryptophan, tyrosine, and cysteine residues and the absorbance at 280 nm in 6 M GuHCl, 20 mM NaPO₄ pH 6.5 (64).

References

1. K. M. Johnson, P. A. Webb, J. V. Lange, R. A. Murphy, *Lancet* **1**, 569-571 (1977).
2. Center for Disease Control and Prevention, *Morbid. Mortal. Week* **44**, 381-382 (1995).
3. L. H. Elliott, M. P. Kiley, J. B. McCormick, *Virology* **147**, 169-176 (1985).
4. A. Sanchez, M. P. Kiley, *Virology* **157**, 414-420 (1987).
5. V. E. Volchkov, et al., *Virology* **214**, 421-430 (1995).
6. A. Sanchez, S. G. Trappier, B. W. Mahy, C. J. Peters, S. T. Nichol, *Proc Natl Acad Sci U S A* **93**, 3602-3607 (1996).
7. Z. Yang, et al., *Science* **279**, 1034-1037 (1998).
8. V. E. Volchkov, H. Feldmann, V. A. Volchkova, H. D. Klenk, *Proc Natl Acad Sci U S A* **95**, 5762-7 (1998).
9. V. E. Volchkov, V. A. Volchkova, W. Slenczka, H. D. Klenk, H. Feldmann, *Virology* **245**, 110-119 (1998).
10. B. N. Fields, et al., Eds., *Fields Virology*, Third Edition, Lippincott, Philadelphia, (1996).
11. W. R. Gallaher, J. M. Ball, R. F. Garry, M. C. Fiffin, R. C. Montelaro, *AIDS Res. Hum. Retroviruses* **5**, 431-440 (1989).
12. P. Chambers, C. R. Pringle, A. J. Easton, *J. Gen. Virol.* **71**, 3075-3080 (1990).
13. W. R. Gallaher, *Cell* **85**, 477-478 (1996).
14. W. Weissenhorn, L. J. Calder, S. A. Wharton, J. J. Skehel, D. C. Wiley, *Proc Natl Acad Sci U S A* **95**, 6032-6036 (1998).
15. P. A. Bullough, F. M. Hughson, J. J. Skehel, D. C. Wiley, *Nature (London)* **371**, (1994).
16. D. C. Chan, D. Fass, J. M. Berger, P. S. Kim, *Cell* **89**, 263-273 (1997).
17. W. Weissenhorn, A. Dessen, S. D. Harrison, J. J. Skehel, D. C. Wiley, *Nature (London)* **387**, 426-430 (1997).
18. K. Tan, J. Liu, J. Wang, S. Shen, M. Lu, *Proc. Natl. Acad. Sci. USA* **94**, 12303-12308 (1997).
19. V. N. Malashkevich, D. C. Chan, C. T. Chutkowski, P. S. Kim, *Proc. Natl. Acad. Sci. USA* **95**, 9134-9139 (1998).
20. M. Caffrey, et al., *EMBO J.* **17**, 4572-4584 (1998).

21. B. Kobe, R. J. Center, B. E. Kemp, P. Poubouris, *Proc. Natl. Acad. Sci. USA* **Submitted**, (1999).
22. D. Fass, S. C. Harrison, P. S. Kim, *Nat. Struct. Biol.* **3**, 465-469 (1996).
23. M. B. Ruiz-Argüello, F. M. Goñi, F. B. Pereira, J. L. Nieva, *J Virol* **72**, 1775-1781 (1998).
24. M. Singh, B. Berger, P. S. Kim, *J. Mol. Biol.* **Submitted**, (1999).
25. V. E. Volchkov, V. M. Blinov, S. V. Netesov, *FEBS Lett* **305**, 181-184 (1992).
26. D. Fass, P. S. Kim, *Current Biology* **5**, 1377-1383 (1995).
27. S. C. Blacklow, M. Lu, P. S. Kim, *Biochemistry* **34**, 14955-14962 (1995).
28. M. Lu, S. C. Blacklow, P. S. Kim, *Nat. Struct. Biol.* **2**, 1075-1082 (1995).
29. E. K. O'Shea, R. Rutowski, P. S. Kim, *Science* **243**, 538-542 (1989).
30. Y.-H. Chen, J. T. Yang, K. H. Chau, *Biochemistry* **13**, 3350-3359 (1974).
31. V. N. Malashkevich, et al., *Proc. Natl. Acad. Sci. USA* **96**, 2662-2667 (1999).
32. D. M. Eckert, V. N. Malashkevich, P. S. Kim, *J. Mol. Biol.* **284**, 859-865 (1998).
33. F. H. C. Crick, *Acta Crystallogr.* **6**, 685-689 (1953).
34. P. B. Harbury, P. S. Kim, T. Alber, *Nature (London)* **371**, 80-83 (1994).
35. P. B. Harbury, B. Tidor, P. S. Kim, *Proc. Natl. Acad. Sci. USA* **92**, 8408-8412 (1995).
36. W. Weissenhorn, A. Carfi, K.-H. Lee, J. J. Skehel, D. C. Wiley, *Mol. Cell* **2**, 605-616 (1998).
37. J. H. Brown, C. Cohen, D. A. Parry, *Proteins* **26**, 134-145 (1996).
38. S. B. Joshi, R. E. Dutch, R. A. Lamb, *Virology* **248**, 20-34 (1998).
39. R. A. Lamb, *Cell* **Submitted**, (1999).
40. F. M. Hughson, *Curr. Biol.* **7**, R565-R569 (1997).
41. R. A. Furuta, C. T. Wild, Y. Weng, C. D. Weiss, *Nat. Struct. Biol.* **5**, 276-279 (1998).
42. D. C. Chan, P. S. Kim, *Cell* **93**, 681-684 (1998).
43. T. Danieli, S. L. Pelletier, Y. I. Henis, J. M. White, *J. Cell. Biol.* **133**, 559-569 (1996).

44. T. Kanaseki, K. Kawasaki, M. Murata, Y. Ikeuchi, S. Ohnishi, *J. Cell. Biol.* **133**, 559-569 (1997).
45. L. V. Chernomordic, V. A. Frolov, E. Leikina, P. Bronk, J. Zimmerberg, *J. Cell. Biol.* **140**, 1369-1382 (1998).
46. P. I. Hanson, R. Roth, H. Morisaki, R. Jahn, J. E. Heuser, *Cell* **90**, (1997).
47. R. C. Lin, R. H. Scheller, *Neuron* **19**, 1087-1094 (1997).
48. T. Weber, et al., *Cell* **92**, 759-772 (1998).
49. R. B. Sutton, D. Fasshauer, R. Jahn, A. T. Brünger, *Nature (London)* **395**, 347-353 (1998).
50. C. M. Carr, P. S. Kim, *Cell* **73**, 823-832 (1993).
51. S. Jiang, K. Lin, N. Strick, A. R. Neurath, *Nature (London)* **365**, 113 (1993).
52. C. T. Wild, D. C. Shugars, T. K. Greenwell, C. B. McDanal, M. T. J., *Proc. Natl. Acad. Sci. USA* **91**, 9770-9774 (1994).
53. D. C. Chan, C. T. Chutkowski, P. S. Kim, *Proc. Natl. Acad. Sci. USA* **95**, 15613-15617 (1998).
54. D. Rappaport, M. Ovadia, Y. Shai, *EMBO J.* **14**, 5524-5531 (1995).
55. D. M. Lambert, et al., *Proc. Natl. Acad. Sci. USA* **93**, (1996).
56. Q. Yao, R. W. Compans, *Virology* **223**, 103-112 (1996).
57. J. M. Kilby, et al., *Nat. Med.* **4**, 1302-1307 (1998).
58. G. Q. Chen, I. Choi, B. Ramachandran, J. E. Gouaux, *J Am Chem Soc* **116**, 8799-8800 (1994).
59. T. N. M. Schumacher, et al., *Science* **271**, 1854-1857 (1996).
60. J. P. Staley, P. S. Kim, *Prot Sci* **3**, 1822-1832 (1994).
61. T. A. Kunkel Roberts, J. D. & Zakour, R. A., *Methods Enzymol* **154**, 367-382 (1987).
62. F. W. Studier, A. H. Rosendberg, J. J. Dunn, J. W. Dubendorff, *Methods Enzymol* **185**, 60-89 (1990).
63. T. M. Laue, B. D. Shah, T. M. Ridgeway, S. L. Pelletier, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. S. E. Harding, A. J. Rowe, H. J.C., Eds. (Royal Society of Chemistry, Cambridge, 1992) pp. 90-125.
64. H. Edelhoch, *Biochemistry* **6**, 1948-1954 (1967).
65. H. Schägger, G. von Jagow, *Anal Biochem* **166**, 368-379 (1987).

66. Molecular Simulations. (Molecular Simulations, San Diego, CA, 1997),
67. T. A. Jones, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr. D* **47**, 110-119 (1991).