

# Designed protein mimics of the Ebola virus glycoprotein GP2 $\alpha$ -helical bundle: stability and pH effects

Joseph S. Harrison,<sup>1,3</sup> Chelsea D. Higgins,<sup>1,3</sup> Kartik Chandran,<sup>2</sup> and Jonathan R. Lai<sup>1\*</sup>

<sup>1</sup> Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461

<sup>2</sup> Department of Microbiology and Immunology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461

<sup>3</sup> These two authors contributed equally to this work.

Grant sponsor: National Institutes of Health (NIH); Grant numbers: R01-AI090249 and R01-AI088027; Grant sponsor: Albert Einstein College of Medicine; Grant sponsor: NIH Molecular Biophysics Training Grant; Grant number: T32-GM008572.

*Correspondence to:* Jonathan R. Lai, Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461. Phone: 718-430-8641. Fax: 718-430-8565. E-mail: jon.lai@einstein.yu.edu

Running Title: Mimicry of the Ebola virus GP2  $\alpha$ -helical bundle.

**(This manuscript contains 25 total pages, 1 Table, and 5 Figures)**

**Abstract**

Ebolavirus (EboV) belongs to the *Filoviridae* family of viruses that causes severe and fatal hemorrhagic fever. Infection by EboV involves fusion between the virus and host cell membranes mediated by the virus's envelope glycoprotein GP2. Similar to the envelope glycoproteins of other viruses, the central feature of the GP2 ectodomain post-fusion structure is a six-helix bundle formed by the protein's N- and C-heptad repeat regions (NHR and CHR, respectively). Folding of this six-helix bundle provides the energetic driving force for membrane fusion; in other viruses, designed agents that disrupt formation of the six-helix bundle act as potent fusion inhibitors. To interrogate determinants of EboV GP2-mediated membrane fusion, we designed model proteins that consist of the NHR and CHR segments linked by short protein linkers. Circular dichroism and gel filtration studies indicate these proteins adopt stable  $\alpha$ -helical folds consistent with design. Thermal denaturation indicated the GP2 six-helix bundle is highly stable at pH 5.3 (melting temperature,  $T_m$ , of  $86.8 \pm 2.0$  °C and van't Hoff enthalpy,  $\Delta H_{VH}$ , of  $-28.2 \pm 1.0$  kcal/mol) and comparable in stability to other viral membrane fusion six-helix bundles. We found the stability of our designed  $\alpha$ -helical bundle proteins was dependent on buffering conditions with increasing stability at lower pH. Small pH differences (5.3 – 6.1) had dramatic effects ( $\Delta T_m = 37$  °C) suggesting a mechanism for conformational control that is dependent on environmental pH. These results suggest a role for low pH in stabilizing six-helix bundle formation during the process of GP2-mediated viral membrane fusion.

**Keywords**

Viral membrane fusion; protein design; protein folding; circular dichroism

**Broad statement**

Model proteins for studying factors that drive ebolavirus membrane fusion were designed and characterized. The resulting information provided insight into the mechanism by which this virus infects cells and may aid in the development of ebolavirus entry inhibitors.

## Introduction

The transmembrane envelope glycoproteins of many viruses employ simple  $\alpha$ -helical bundle motifs to facilitate fusion between the virus and host membranes during infection.<sup>1-3</sup> (Historically, envelope glycoproteins have been categorized according to structural features with  $\alpha$ -helical bundle proteins comprising the so-called 'class I' family.<sup>2</sup>) In the general model for membrane fusion, folding of a six-helix bundle by the glycoprotein ectodomain provides the driving force for bringing the two bilayers together.<sup>1-3</sup> The prototypic six-helix bundle formed by the N- and C-heptad repeat (NHR and CHR) regions from the HIV-1 transmembrane envelope glycoprotein gp41 has been particularly instructive for understanding this mechanism (Figure 1).<sup>4-6</sup> The primary sequence of gp41 contains a hydrophobic N-terminal fusion peptide, the ectodomain containing the NHR and CHR, and a transmembrane domain that anchors the glycoprotein to the virus. Following receptor binding by the surface glycoprotein (gp120), the fusion peptide of gp41 embeds in the membrane of the host cell giving rise to a transient trimeric intermediate known as the 'extended' or 'prehairpin' intermediate that spans both the virus and cell membranes and in which the NHR and CHR segments are exposed (Figure 1A). Next, folding of the NHR and CHR into the six-helix bundle forces the two anchored membranes into proximity which promotes mixing and, ultimately, formation of a fusion pore. This process almost certainly proceeds via a hemifusion intermediate, in which the outer leaflets are fused but the inner leaflets are not, and it is thought that the transition from this hemifusion intermediate to the fusion pore represents the rate-limiting step during entry.<sup>1</sup> Transmembrane glycoproteins from other viruses such as ebolavirus (GP2) and influenza A virus (HA2) have similar

architectural features and by analogy are thought to promote membrane fusion via similar mechanisms.<sup>1-3,7-9</sup> For ebolavirus (EboV), experimental evidence suggests that GP2 also forms an extended intermediate that ultimately collapses into the six-helix bundle and that degradation of the surface subunit (GP1) by host endosomal proteases cathepsins L and B (CatL and CatB) is an obligatory step in the pathway toward this intermediate.<sup>7-12</sup> However, the precise molecular events that lead to triggering of the envelope spike to form the extended intermediate are still under investigation.

Crystal structures of the EboV GP2 ectodomain in the putative post-fusion conformation were reported independently by the Kim and Wiley groups (Figure 1B).<sup>8,9</sup> The structure described by Kim and coworkers was of a protease stable 74-residue fragment encompassing residues 557-630 (*Zaire ebolavirus* numbering, 'Ebo-74').<sup>9</sup> The structure reported by Wiley and coworkers contained residues 552-650 of the ectodomain linked to a N-terminal trimeric GCN4 coiled-coil domain to promote solubility and stability ('pII Gp2(552-650)').<sup>8</sup> The GP2 ectodomain regions in these two structures were in good agreement and revealed several unique features of the  $\alpha$ -helical bundle. The EboV GP2 six-helix bundle is much shorter than that of HIV-1 gp41, consisting of segments containing just three  $\alpha$ -helical turns each in the NHR and CHR regions compared to seven turns for these regions in gp41. The region linking the NHR and CHR in GP2 contains a short helix-turn-helix motif that is stabilized by an intramolecular disulfide bond. Interestingly, this motif is structurally similar to that found in the moloney murine leukemia virus (MoMLV) glycoprotein post-fusion core despite the fact that these viruses are not phylogenetically related.<sup>8,9,13</sup> Sequence homology suggests that analogous regions in the envelope glycoproteins of avian sarcoma/leukosis virus

(ASLV) and human T-cell leukemia virus type I (HTLV-1) adopt similar structures, and mutational studies suggest a histidine residue in this region plays a role in pH-dependent ASLV viral entry.<sup>14</sup> In EboV GP2, the helix-turn-helix motif appears to contribute to overall stability as the native disulfide bond pairing is required for high thermostability.<sup>9</sup> Another feature of EboV GP2 is the presence of a 'stutter' in the heptad repeat pattern in the NHR core, noted by both the Kim and Wiley groups. Typical  $\alpha$ -helical bundle proteins contain a characteristic **abcdefg** repeat where **a** and **d** positions are occupied by hydrophobic residues (the '3-4 repeat').<sup>15</sup> In the NHR region of GP2, there is a stutter in this hydrophobic repeat at position T565 that results in a 3-4-4-3 repeat pattern; a similar stutter exists at the N-terminal end of the low-pH induced influenza A virus HA2 coiled-coil.<sup>16</sup> Finally, the central NHR coiled-coil contains an anion binding pocket that also appears to contribute to folding stability. These features make the EboV GP2  $\alpha$ -helical bundle an interesting system to study in the context of factors that are required to drive membrane fusion.

The evidence for the extended intermediate in HIV-1 gp41 was obtained from designed peptides and proteins that bound the extended intermediate and prevented rearrangement to the six-helix bundle ('fusion inhibitors', Figure 1A).<sup>3,4</sup> Proteins, peptides, and small molecules that bind protein mimics of the HIV-1 gp41 extended intermediate have antiviral activity and have been exploited for design of viral entry inhibitors.<sup>17-20</sup> Furthermore, appropriately designed protein mimics of the gp41 extended intermediate serve as targets for neutralizing antibodies and thus have potential as immunogens for vaccine discovery.<sup>21-24</sup> In theory, similar approaches could be adopted for EboV GP2, given the similarity in their proposed fusion mechanisms.<sup>25</sup>

Our goal was to design protein mimics of the  $\alpha$ -helical bundle segment of EboV GP2 that could be used as model systems to study folding stability and potentially aid in development of inhibitors of EboV membrane fusion. To this end, we prepared model proteins consisting of the alternating NHR and CHR segments from EboV GP2 linked by short, flexible loops. Biophysical characterization reveals that these proteins adopt stable  $\alpha$ -helical structures consistent with their design. Further analysis provided interesting insights into factors that influence stability of the EboV GP2 six-helix bundle.

## Results and Discussion

### *Design of Protein Mimics of the EboV GP2 $\alpha$ -Helical Bundle.*

We sought to analyze the folding properties of the  $\alpha$ -helical bundle segment of the EboV GP2 ectodomain. The GP2 post-fusion structure is distinct from that of HIV-1 gp41 in that GP2 contains a relatively small six-helix bundle segment but a long, triple stranded core NHR coiled-coil; much of the GP2 post-fusion out layer consists of the helix-turn-helix and loop segments (the six-helix bundle is boxed in Figure 1B).<sup>8,9</sup> By contrast, the long, central NHR trimer of HIV-1 gp41 is surrounded by CHR  $\alpha$ -helices of essentially equal length.<sup>5,6,26-28</sup> Previously, Wiley and coworkers performed denaturation studies with pII<sub>GP2</sub>(552-650) found it to be highly thermostable (complete unfolding was not observed upon heating to 98 °C at pH 8).<sup>29</sup> Kim and coworkers performed denaturation studies with two GP2 ectodomain constructs, a larger 95-residue construct containing essentially the entire GP2 ectodomain ('Ebo-95'), and the protease-resistant Ebo-74.<sup>9</sup> These proteins were also found to be thermostable (thermal midpoint,  $T_m$ , of 88 °C at pH 7.2 for Ebo-95 and 77 °C at pH 4.5 for Ebo-74). We wondered what contributions were provided by the GP2 six-helix bundle segment alone (i.e., in the absence of the

helix-turn-helix and loop segments). Analysis of folding energetics for six-helix bundle constructs from HIV-1 and SIV gp41 has previously been performed using short hairpin proteins (i.e., a single NHR and a single CHR segment linked by a short linker that form a trimer).<sup>27,28</sup> However, such proteins typically display concentration-dependent behavior and therefore aggregation complicates thermodynamic analysis. We reasoned that a single chain construct consisting of the EboV GP2  $\alpha$ -helical NHR and CHR segments alone would provide a useful system to examine folding stability of the GP2 six-helix bundle.

We designed two proteins consisting of alternating NHR and CHR segments of *Zaire ebolavirus* GP2: 'Ebo6H' and 'Ebo5H' (Figure 2). Ebo6H contains the segments each of the CHR region encompassing residues 615-633 (this segment is referred to as 'C19') linked to the NHR segment consisting of residues 559-577 ('N18', which includes the unusual heptad repeat stutter) via short, flexible protein linkers. When properly folded, this polypeptide is predicted to adopt a six-helix bundle that mimics that of the GP2 post-fusion core.<sup>8,9</sup> Ebo5H is similar to Ebo6H but lacks one of the CHR regions. When folded, Ebo5H should adopt a structure similar to that of Ebo6H but lacking a single CHR segment. Comparison of stabilities for Ebo5H and Ebo6H should therefore provide insight into contribution of stability of a single CHR segment. In addition, similar '5-Helix' constructs based on HIV-1 gp41 that have been used as entry inhibitors and mimics of the extended intermediate.<sup>18-24</sup> Therefore, Ebo5H may be useful for development of EboV entry inhibitors. Ebo6H and Ebo5H proteins containing hexahistidine tags were expressed in *E. coli*, purified from inclusion bodies, and then



subjected to refolding conditions. Both proteins were obtained in good yield and purity (Figure 1C).

***Ebo6H and Ebo5H Adopt Stable  $\alpha$ -Helical Conformations.***

Circular dichroism spectra for Ebo6H (2  $\mu$ M) and Ebo5H (0.4  $\mu$ M) in 20 mM sodium acetate (pH 5.3) exhibited double minima at 208 nm and 222 nm, characteristic of  $\alpha$ -helical secondary structure and consistent with the intended design (Figure 3A). The CD signature (as judged by ratio of signal intensity at 208 nm and 222 nm) for both proteins was invariant across a 0.2 – 10  $\mu$ M concentration range, indicating that  $\alpha$ -helix formation was not driven by aggregation. Gel filtration analysis at mid-micromolar protein concentrations (80  $\mu$ M for Ebo6H and 50  $\mu$ M for Ebo5H) yielded molecular weight estimates consistent with monomer for each protein, although a small amount of aggregate was observed in both cases (Figure 3B). Ebo6H underwent a cooperative unfolding transition upon heating when monitored by CD at 222 nm (Figure 3) with  $T_m$  of  $86.8 \pm 2.0$  °C at pH 5.3 and  $77.8 \pm 1.6$  °C at pH 5.8. Ebo5H was less stable than Ebo6H at pH 5.3 ( $T_m = 80.6 \pm 0.7$  °C); therefore, the additional CHR segment in Ebo6H provided 6.2 °C of thermal stability under these conditions. The reversibility of unfolding was confirmed by cooling heated samples to room temperature over 5 minutes; we found the intensity of the 222 nm signal returned to initial values during this process (data not shown).

To obtain insight into thermodynamic parameters of unfolding, we performed van't Hoff analysis of the thermal denaturation data.<sup>30-32</sup> Assuming a simple two-state unfolding model ( $N \rightleftharpoons D$ ) with no intermediates, the equilibrium constant for the unfolding reaction ( $K_{unf}$ ) is related to the fraction unfolded ( $F_{unf}$ ) by equation 1 below.

$$K_{\text{unf}} = F_{\text{unf}} / (1 - F_{\text{unf}}) \quad (1)$$

The van't Hoff enthalpy ( $\Delta H_{\text{vH}}$ ) can be derived from non-linear regression of  $K_{\text{unf}}$  as a function of temperature (T) accordance with equation 2 (where R is the gas constant).

$$K_{\text{unf}} = \exp [\Delta H_{\text{vH}}/R (1/T - 1/T_m)] \quad (2)$$

While  $\Delta H_{\text{vH}}$  does not provide direct information about the free energy of unfolding ( $\Delta G_{\text{unf}}$ ), it nonetheless provides a rough parameter to assess folding stability. This analysis yielded  $\Delta H_{\text{vH}}$  of  $-28.2 \pm 1.0$  kcal/mol for Ebo6H and  $-27.6 \pm 0.4$  kcal/mol for Ebo5H. These values are lower than unfolding enthalpies obtained by differential scanning calorimetry of model six-helix bundle proteins of SIV gp41 ( $\Delta H_{\text{unf}} = -42.3$  kcal/mol).<sup>28</sup> The SIV gp41 six-helix bundle segment is much longer than that of EboV GP2; therefore it is not surprising that the unfolding enthalpy of SIV gp41 is larger than that of EboV GP2.

### ***Effects of pH on Stability.***

The envelope glycoproteins of many viruses that enter via the endosome undergo large structural transitions upon exposure to acidic pH.<sup>1-3,16,33-37</sup> These pH-dependent conformational changes are generally thought to control timing of the membrane fusion reaction such that it occurs only in endosomal compartments. For example, low pH triggers dissociation of the surface subunit (HA1) and HA2 subunits of influenza and extension of the central HA2 coiled-coil.<sup>1,3</sup> A pH-dependent coil-to-helix transition in the midsection of the HA2 coiled-coil effectively jackknives the N-terminal fusion peptide from the center of the envelope spike out toward the cell membrane.<sup>16,33,34</sup> This rearrangement projects the fusion peptide into the cell membrane, leading to the extended intermediate. In the envelope glycoproteins of viruses such as alphaviruses

and flaviviruses that contain mostly  $\beta$ -sheet structures (the 'class II' envelope glycoproteins), similar pH effects destabilize the inactive prefusion conformation and promote oligomerization of fusion-active components.<sup>35,37</sup> In EboV entry, proteolytic disassembly of GP1 by CatL and CatB, whose optimal catalytic activity is known to be at pH ~5, is an essential step in early infection events.<sup>10,11</sup> However, pH-dependent structural transitions have not been reported for GP1 or the GP2 ectodomain. Very recently, Gregory et al. demonstrated that the GP2 fusion loop undergoes pH-dependent conformational changes that promote fusogenic activity.<sup>38</sup>

We determined the effect of pH on  $\alpha$ -helical bundle stability of Ebo6H and Ebo5H by performing spectroscopic analysis under various buffer conditions. In conditions above pH 6.5, both proteins were poorly behaved and prone to aggregation. However, in range of pH 4.8 to pH 6.1, we found that both proteins retained their well-behaved, native-like properties. We performed thermal denaturation of Ebo6H and Ebo5H under several buffering conditions and found that the  $T_m$  was lower at higher pH for both proteins (Figure 4 and Table 1). The folding stability of small, globular proteins can be significantly influenced by pH; however, most proteins are most stably folded at neutral (physiologic) pH and become less stable in conditions below pH 6 (converse to the behavior observed here).<sup>39</sup> The effects of pH on  $T_m$  were most pronounced in Ebo6H, where the  $T_m$  differed by 37 °C between pH 5.2 ( $T_m = 86.8 \pm 2.0$  °C) and pH 6.1 ( $T_m = 49.8 \pm 1.1$  °C), corresponding to a  $\Delta\Delta H_{VH}$  of 7.8 kcal/mol over this range. In addition, heating to 100 °C at pH 4.8 did not result in complete unfolding of Ebo6H (Figure 3C). Similar stability effects were observed for Ebo5H although they were less dramatic. The  $T_m$  values ranged from  $82.2 \pm 1.2$  °C at pH 4.8 to  $72.6 \pm 1.9$  °C at pH 6.1; this

corresponded to a  $\Delta\Delta H_{\text{vH}}$  of 9.6 kcal/mol over this range. Both pII GP2(552-650), the Wiley construct, and Ebo-95, the longer of the two Kim constructs, had high thermostability at neutral pH or higher.<sup>9,29</sup> However, the shorter Ebo-74 construct was crystallized and characterized by CD at pH 4.5 – 4.6.<sup>9</sup> Ebo6H and Ebo5H contain even less of the ectodomain than does Ebo-74, therefore the  $\alpha$ -helical stability of shorter constructs may be more sensitive to environmental pH.

An examination of the six-helix bundle segment of the GP2 post-fusion structure suggests two side chain-side chain ionic interactions may be responsible for this pH-dependent behavior. On the CHR  $\alpha$ -helix, an intrahelical  $i \rightarrow i+3$  interaction between D621 and D624 could destabilize CHR  $\alpha$ -helix formation at neutral pH due to repulsion between the carboxylate anions (Figure 5A). In other short  $\alpha$ -helical bundle systems, overall folding stability can be greatly influenced by propensity of individual segments to adopt  $\alpha$ -helical fold.<sup>40</sup> Electrostatic  $i \rightarrow i+3$  interactions can strongly influence  $\alpha$ -helix formation and therefore  $\alpha$ -helical bundle formation in systems containing segments as short as the GP2 CHR (three  $\alpha$ -helical turns).<sup>41,42</sup> In addition, a repulsive interaction between E564 of the NHR and D629 of the CHR may disrupt interhelical association at neutral pH (Figure 5A). To test this hypothesis, we prepared an Ebo6H variant with the mutations D621N, D624N, D629N on the CHR segments, and E564Q on the NHR segments ('Ebo6H-QNNN'). Since Ebo6H contains a total of three NHR and CHR segments each, the GP2 NHR/CHR segments of Ebo6H-QNNN have twelve total substitutions relative to the analogous segments in Ebo6H. We found Ebo6H-QNNN to be  $\alpha$ -helical by CD; thermal denaturation indicates the  $T_{\text{ms}}$  were similar at pH 4.8 and 5.3 ( $43.1 \pm 0.9$  °C and  $46.2 \pm 1.0$  °C, respectively, Figure 5B) but was lower than Ebo6H

under these conditions. Furthermore, unfolding transition for Ebo6H-QNNN was much broader than that of Ebo6H. These differences may reflect the fact that asparagine has lower  $\alpha$ -helical propensity in coiled-coil proteins than does aspartic acid and therefore the nine Asp→Asn substitutions in the GP2 NHR/CHR regions have a cumulative effect on folding stability.<sup>40</sup> Ebo6H-QNNN had higher stability at pH 6.1 ( $T_m = 53.5 \pm 1.1$  °C) than at the lower pHs. This behavior contrasts with that of Ebo6H, which was significantly more stable under acidic conditions than at pH 6.1. Furthermore, Ebo6H-QNNN had 3.7 °C higher  $T_m$  than did Ebo6H at pH 6.1, indicating that the  $\alpha$ -helical fold of Ebo6H-QNNN is less sensitive to conditions near neutral pH. In addition, Ebo6H was much less stable than Ebo5H at pH 6.1 ( $T_m$ s of  $49.8 \pm 1.1$  °C for Ebo6H and  $72.6 \pm 1.9$  °C for Ebo5H) suggesting that repulsive interactions within the additional CHR segment in Ebo6H substantially affects global  $\alpha$ -helical bundle fold stability under these conditions. Taken together, these results suggest that anionic side chain-side chain interactions mediate pH-dependent  $\alpha$ -helical bundle stability in Ebo6H.

Both Ebo6H and Ebo5H were purified with two hexahistidine tags, one at each terminus. Since the pKa of the imidazole side chain is ~6, it is possible that the histidine tags contribute to the pH-sensitive behavior observed here. Unfortunately, attempts to isolate variants containing fewer than two histidine tags were unsuccessful and therefore we could not directly compare thermostability with and without these segments. However, increased cationic charge density by protonation of histidine tags at the N- and C-terminal ends is predicted to disfavor folding due to repulsion between the N- and C-terminal ends in a compact, globular state. While we cannot rule out the possibility that the histidine tags play a role in the pH-dependent stability, it seems

unlikely that the observed effects could be explained by these regions alone. Although the pH range in which we see the most dramatic changes in Ebo6H and Ebo5H structural stability (pH 4.8 – 6.1) falls above the pKa of carboxylic acid side chains (pH ~4), it is not unusual for side chain pKas to be elevated in situations where their deprotonation disfavors formation of a highly stable fold.<sup>43-46</sup> Such phenomena have been documented in other  $\alpha$ -helical bundle proteins as well.<sup>43-46</sup>

### ***Conclusions and Implications for Membrane Fusion***

Our analysis with Ebo6H and Ebo5H provide new insight into folding and stability for the six-helix bundle of EboV GP2. We found that Ebo6H, a designed protein mimic of the GP2  $\alpha$ -helical bundle, adopts a highly thermostable fold. Among class I viral membrane fusion glycoproteins, the central ectodomain core NHR trimeric coiled-coil of EboV GP2 and analogous proteins from the structurally related MarV, MoMLV HTLV-1, and ASLV are relatively short in length.<sup>1,2,14</sup> The core NHR trimers of influenza A virus HA2 and severe acute respiratory syndrome (SARS) coronavirus S2 are longer than EboV GP2, but these ectodomains also contain loop or non- $\alpha$ -helical segments and therefore their six-helix bundle segments are also small in comparison to HIV-1 gp41.<sup>3</sup> Our studies demonstrate that even the small  $\alpha$ -helical bundle of EboV GP2 (consisting of  $\alpha$ -helical segments containing just 3-4 turns) provides a large amount of energy to overcome barriers associated with membrane fusion. Furthermore, we found that Ebo5H was less stable than Ebo6H ( $\Delta T_m = 6.2$  °C at pH 5.3) thereby demonstrating that a single CHR segment can enhance thermostability of the  $\alpha$ -helical bundle by a significant amount. Similar designed '5-Helix' proteins based on HIV-1 gp41 act as potent inhibitors of entry by sequestering the CHR segment on the extended intermediate during infection.<sup>18-24</sup>

We tested Ebo5H for entry inhibition using a recombinant vesicular stomatitis virus containing the EboV GP in place of the native envelope glycoprotein (VSV-GP). We found that Ebo5H was not able to inhibit infection of VSV-GP (E. H. Miller and K. Chandran, unpublished results), although this result is likely because the extended intermediate is exposed only in the endosome.<sup>7</sup> We recently reported that antiviral activity of peptides corresponding to the GP2 CHR segment ('C-peptides') can be enhanced by inclusion of an endosome-targeting sequence.<sup>25</sup>

Further studies indicated that the folding stabilities of Ebo6H and Ebo5H were pH-sensitive. The  $\alpha$ -helical bundle stability as judged by  $T_m$  and  $\Delta H_{vH}$  values was significantly lower in buffering conditions closer to neutral pH than in acidic conditions. Analysis with Ebo6H-QNNN suggested that anionic side chain-side chain interactions mediate this pH-dependent stability in Ebo6H; this feature of the GP2 six-helix bundle may have implications in control of the fusion reaction. In the prevailing model for viral membrane fusion, it is thought that the transmembrane subunit is prevented from adopting the thermodynamically stable  $\alpha$ -helical bundle structure by interactions with the surface subunit in the prefusion conformation.<sup>47-48</sup> Triggering events cause dissociation of the surface subunit and releases constraints on the transmembrane subunit ectodomain, allowing the eventual formation of the  $\alpha$ -helical bundle that drives membrane fusion (the 'spring loaded clamp' model). In HIV-1 and influenza A virus, the events that trigger release the transmembrane subunits (gp41 and HA2) from the prefusion conformation are well-established. For HIV-1, where membrane fusion is thought to occur at the cellular surface, a conformational change in gp120 upon CD4 and coreceptor binding induces its dissociation from the spike.<sup>49</sup> In influenza A virus,

which enters via the endosome, low pH destabilizes the prefusion conformation and leads to the extended intermediate.

A role for pH in structural transitions of GP1 or the GP2 ectodomain has not been established and the precise trigger for membrane fusion is still under investigation. Comparison of the NHR and CHR segments in the prefusion GP1-GP2 assembly demonstrates that the protein segment corresponding to the long  $\alpha$ -helical NHR coiled-coil is interrupted into three shorter  $\alpha$ -helices in the prefusion structure.<sup>7,12</sup> Electron density for the CHR segments was not observed in the prefusion structure, suggesting this region is dynamic.<sup>12</sup> Recent mutational studies on the helix-turn-helix/loop segment of the ASLV transmembrane envelope glycoprotein (EnvA) have implicated a histidine residue in the pH dependence of EnvA-mediated viral entry.<sup>14</sup> Mutation of residues in the analogous segment in EboV GP2 have an effect on viral entry but, in the absence of a fusion assay directed at the cell surface, it is not possible to determine whether these residues are involved in pH-dependent structural changes that promote membrane fusion.<sup>14</sup> Furthermore, a recent report suggests that the fusion loop of GP2 also undergoes a conformational transition between pH 7 and pH 5.5 that activates fusogenic activity under the lower pH condition.<sup>38</sup> The results presented here suggest that pH-dependent  $\alpha$ -helical bundle stability could also play a role in controlling the timing of membrane fusion. At neutral pH, the six-helix bundle is less stable and therefore the prefusion conformation is preferred. However, once the virus particle is taken up into an endosome that undergoes acidification, the post-fusion six-helix bundle become more stable and therefore the post-fusion state becomes preferred. The pH range in which we see the most significant changes in stability (pH 4.8 to pH 6.1)



corresponds to the pH of late endosomes and correlates closely with the optimal pH for acid-dependent cysteine cathepsin activity. This proposed mechanism for conformational control is distinct from that of other endosomal viruses: in most cases, the pH change destabilizes the prefusion conformation rather than stabilizing the post-fusion conformation.<sup>1-4</sup> Further studies will be required to determine if this pH-dependent stability is relevant in the virus and whether other pH-induced structural changes occur.

## Materials and Methods

### ***Cloning, Expression, Purification, and Refolding of Ebo5H and Ebo6H.***

Synthetic DNA encoding the genes for Ebo6H and Ebo5H codon optimized for *E. coli* were obtained from a commercial supplier (DNA 2.0, Menlo Park, CA) and cloned into pET28 using the NdeI and XhoI restriction sites. Constructs of Ebo5H were prepared containing His<sub>6</sub> tags at the N-terminus, at the C-terminus, or at both the N- and C-termini and protein expression assessed in small-scale cultures. We found the construct containing His<sub>6</sub> tags at both the N- and C-termini was most abundantly expressed and therefore we used this format for large-scale expression of both Ebo6H and Ebo5H. *E. coli* BL21(DE3) (Invitrogen, Madison, WI) harboring expression plasmids for Ebo6H or Ebo5H were grown in LB broth at 37 °C to an OD<sub>600</sub> of 0.6 and then expression of the protein induced by addition of 1 mM IPTG. The cells were incubated overnight at 37 °C, harvested by centrifugation, and then lysed in a French Pressure cell. The insoluble fraction was separated by ultracentrifugation, dissolved in 6 M guanidine HCl (GdnHCl) and applied to Ni-NTA resin (Qiagen, Valencia, CA). The resin was washed with 10 column volumes of 15 mM imidazole containing 6 M GdnHCl

and the protein eluted with 250 mM imidazole containing 6 M GdnHCl. The fractions containing purified protein were pooled and refolded by two-step dialysis first into 100 mM glycine HCl (pH 3.5), then into 20 mM NaOAc pH 5.3 containing 150 mM NaCl. Precipitated material was removed by centrifugation and the protein solution used immediately for analysis or flash frozen and stored at  $-80^{\circ}\text{C}$ . Preparation of Ebo6H-QNNN was similar.

### ***Circular Dichroism Spectroscopy.***

Protein samples were prepared in the appropriate buffer and measurements performed on a Jasco J-815 spectrometer with a 1 cm quartz cuvette. Protein concentrations ranged from 0.4 – 10  $\mu\text{M}$  as determined by absorbance at 280 nm. CD wavelength scans were obtained with a 0.2 nm or 1 nm step size and a 2 second averaging time. The signal was corrected for blank and then converted to mean residue ellipticity ( $\theta$ ). Thermal denaturation data were obtained with an  $3^{\circ}\text{C}$  step size and 1 minute equilibration at each temperature. The raw ellipticity signal was corrected for folded and unfolded baselines and then non-linear least squares regression was performed using a standard four-parameter logistic equation. The  $T_m$  values were derived from the inflection point of the curve. For van't Hoff analysis, the baseline-corrected ellipticity was converted to  $F_{\text{unf}}$  and then  $K_{\text{unf}}$  was calculated at each temperature (in degrees Kelvin) according to equation 1. Non-linear least-squares regression of a  $K_{\text{unf}}$  vs.  $T$  curve was then performed in accordance with equation 2 to obtain  $\Delta H_{\text{vH}}$ .

### ***Gel Filtration Analysis.***

Analysis was performed on a Sephadex S75 column (10/300) at  $4^{\circ}\text{C}$  in 20 mM NaOAc pH 5.3 containing 150 mM NaCl. Protein samples were loaded at concentrations of 50

– 80  $\mu$ M and elution was monitored by absorbance at 280 nm. Both Ebo6H and Ebo5H eluted as a single peak. Protein molecular weight standards (Bio-Rad Laboratories, Hercules, CA) were used to generate a standard curve and the Ebo6H and Ebo5H molecular weights estimated by elution volume.

### **Acknowledgements**

We thank Michael Brenowitz and Margaret Kielian (Albert Einstein College of Medicine) for helpful discussions and critical reading of this manuscript.

## References

1. Harrison SC (2008) Viral Membrane Fusion. *Nat Struct Mol Biol* 15:690-698.
2. White JM, Delos SE, Brecher M, Schornberg K (2008) Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit Rev Biochem Mol Biol* 43:189-219.
3. Eckert DM, Kim PS (2001) Mechanisms of viral membrane fusion and its inhibition. *Annu Rev Biochem* 70:777-810.
4. Chan DC, Kim PS (1998) HIV entry and its inhibition. *Cell* 93:681-684.
5. Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, Wiley DC (1997) Atomic structure of the ectodomain from HIV-1 gp41. *Nature* 387:426-430.
6. Chan DC, Fass D, Berger JM, Kim PS (1997) Core structure of gp41 from the HIV envelope glycoprotein. *Cell* 89:263-273.
7. Lee JE, Saphire EO (2009) Neutralizing ebolavirus: structural insights into the envelope glycoprotein and antibodies targeted against it. *Curr Opin Struc Biol* 19:408-417.
8. Weissenhorn W, Carfi A, Lee KH, Skehel JJ, Wiley DC (1998) Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain. *Mol Cell* 2:605-616.
9. Malashkevich VN, Schneider BJ, McNally ML, Mihollen MA, Pang JX, Kim PS (1999) Core structure of the envelope glycoprotein GP2 from Ebola virus at 1.9Å resolution. *Proc Natl Acad Sci USA* 96:2662-2667.

10. Chandran K, Sullivan NJ, Felbor U, Whelan SP, Cunningham JM (2005) Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. *Science* 308: 1643-1645.
11. Schornberg K, Matsuyama S, Kabsch K, Delos S, Bouton A, White J (2006) Role of endosomal cathepsins in entry mediated by the Ebola virus glycoprotein. *J Virol* 80:4174-4178.
12. Lee JE, Fusco ML, Hessel AJ, Oswald WB, Burton DR, Saphire EO (2008) Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature* 454:177-182.
13. Fass D, Harrison SC, Kim PS (1996) Retrovirus envelope domain at 1.7Å resolution. *Nat Struct Biol* 3:465-469.
14. Delos SE, La B, Gilmartin A, White JM (2010) Studies of the “chain reversal regions” of the Avian Sarcoma/Leukosis Virus (ASLV) and Ebolavirus fusion proteins: Analogous residues are important, and a His residue unique to EnvA affects the pH dependence of ASLV entry. *J. Virol.* 84: 5687-5694.
15. Woolfson DN (2005) The design of coiled-coil structures and assemblies. *Adv Protein Chem* 70: 79-112.
16. Bullough PA, Hughson FM, Skehel JJ, Wiley DC (1994) Structure of the influenza haemagglutinin at the pH of membrane fusion. *Nature* 371: 37-43.
17. Eckert DM, Malashkevich VN, Hong LH, Carr PA, Kim PS (1999) Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. *Cell* 99:103-115.

18. Root MJ, Kay MS, Kim PS (2001) Protein design of an HIV-1 entry inhibitor. *Science* 291:884-888.
19. Frey G, Rits-Volloch S, Zhang XQ, Schooley RT, Chen B, Harrison SC (2006) Small molecules that bind the inner core of gp41 and inhibit HIV envelope-mediated fusion. *Proc Natl Acad Sci USA* 103:13938-13943.
20. Steger HK, Root MJ (2006) Kinetic dependence to HIV-1 entry inhibition. *J Biol Chem* 281: 25813-25821.
21. Luftig MA, Mattu M, Di Giovine P, Geleziunas R, Hrin R, Barbato G, Bianchi E, Miller MD, Pessi A, Carfi A (2006) Structural basis for HIV-1 neutralization by a gp41 fusion intermediate-directed antibody. *Nat Struct Mol Biol* 13:740-747.
22. Miller MD, Geleziunas R, Bianchi E, Lennard S, Hrin R, Zhang H, Lu M, An Z, Ingallinella P, Finotto M, Mattu M, Finnefrock AC, Bramhill D, Cook J, Eckert DM, Hampton R, Patel M, Jarantow S, Joyce J, Ciliberto G, Cortese R, Lu P, Strohl W, Schleif W, McElhaugh M, Lane S, Lloyd C, Lowe D, Osbourn J, Vaughan T, Emini E, Barbato G, Kim PS, Hazuda DJ, Shiver JW, Pessi A (2005) A human monoclonal antibody neutralizes diverse HIV-1 isolates by binding a critical gp41 epitope. *Proc Natl Acad Sci USA* 102: 14759-14764.
23. Sabin C, Corti D, Buzon V, Seaman MS, Lutje Hulsik D, Hinz A, Vanzetta F, Agatic G, Silacci C, Mainetti L, Scarlatti G, Sallusto F, Weiss R, Lanzavecchia A, Weissenhorn W (2010) Crystal structure and size-dependent neutralization properties of HK20, a human monoclonal antibody binding to the highly conserved heptad repeat 1 of gp41. *PLoS Pathog* 6: e1001195.

24. Gustchina E, Li M, Louis JM, Anderson DE, Lloyd J, Frisch C, Bewley CA, Gustchina A, Wlodawer A, Clore GM (2010) Structural basis of HIV-1 neutralization by affinity matured Fabs directed against the internal trimeric coiled-coil of gp41. *PLoS Pathog* 6: e1001182.
25. Miller EH, Harrison JS, Radoshitzky SR, Higgins CD, Chi X, Dong L, Kuhn JH, Bavari S, Lai JR, Chandran K (2011) Inhibition of Ebola virus entry by a C-peptide targeted to endosomes. *J Biol Chem* 286:15854-15861.
26. Frey G, Chen J, Rits-Volloch S, Freeman MM, Zolla-Pazner S, Chen B (2010) Distinct conformational states of HIV-1 gp41 are recognized by neutralizing and non-neutralizing antibodies. *Nat Struct Mol Biol* 17:1486-1491
27. Jelesarov I, Lu M (2001) Thermodynamics of trimer-of-hairpins formation by the SIV gp41 envelope protein. *J Mol Biol* 307:637-656.
28. Marti DN, Bjelić S, Lu M, Bosshard HR, Jelesarov I (2004) Fast folding of the HIV-1 and SIV gp41 six-helix bundles *J Mol Biol* 336: 1-8.
29. Weissenhorn W, Calder LJ, Wharton SA, Skehel JJ, Wiley DC (1998) The central feature of the membrane fusion subunit from the Ebola virus glycoprotein is a long triple-stranded coiled coil *Proc Natl Acad Sci USA* 95: 6032-6036.
30. Greenfield NJ (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions *Nat Protoc* 1: 2527-2535.
31. John DM, Weeks KM (2000) van't Hoff enthalpies without baselines. *Protein Sci* 9:1416-1419.

32. Allen DL, Pielak GJ. (1998) Baseline length and automated fitting of denaturation data. *Protein Sci* 7: 1262-1263.
33. Carr CM, Kim PS (1993) A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell* 73: 823-832.
34. Wilson IA, Skehel JJ, Wiley DC (1981) Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3Å resolution. *Nature* 289: 366-373.
35. Sánchez-San Martín C, Liu CY, Kielian M (2009) Dealing with low pH: entry and exit of alphaviruses and flaviviruses. *Trends Microbiol* 17: 514-521.
36. Fritz R, Stiasny K, Heinz FX (2008) Identification of specific histidines as pH sensors in flavivirus membrane fusion. *J Cell Biol* 183: 353-361.
37. Qin ZL, Zheng Y, Kielian M (2009) Role of conserved histidine residues in the low-pH dependence of the Semliki Forest virus fusion protein. *J Virol* 83: 4670-4677.
38. Gregory SM, Harada E, Liang B, Delos SE, White JM, Tamm LK (2011) Structure and function of the complete internal fusion loop from Ebolavirus glycoprotein 2. *Proc. Natl. Acad. Sci.* Epub Jun 20 DOI 10.1073
39. Yang AS, Honig B (1993) On the pH dependence of protein stability. *J Mol Biol* 231: 459-474.
40. O'Neil KT, DeGrado WF (1990) A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science* 250: 646-651.
41. Kohn WD, Cyril MK, Hodges RS (1995) Protein destabilization by electrostatic repulsions in the two-stranded  $\alpha$ -helical coiled-coil/leucine zipper. *Protein Sci* 4: 237-250.



42. Scholtz JM, Qian H, Robbins VH, Baldwin RL (1993) The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide. *Biochemistry* 32: 9668-9676.
43. O'Shea EK, Lumb KJ, Kim PS (1993) Peptide 'Velcro': design of a heterodimeric coiled-coil. *Curr Biol* 3: 658-667.
44. O'Shea EK, Rutkowski R, Kim PS (1992) Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. *Cell* 68: 699-708.
45. Lumb KJ, Kim PS (1995) Measurement of interhelical electrostatic interactions in the GCN4 leucine zipper. *Science* 268: 436-439.
46. Lau WL, Degrado WF, Roder H (2010) The effects of pK(a) tuning on the thermodynamics and kinetics of folding: design of a solvent-shielded carboxylate pair at the  $\alpha$ -position of a coiled-coil. *Biophys J* 99: 2299-2308.
47. Baker D, Agard DA (1994) Influenza hemagglutinin: kinetic control of protein function. *Structure* 2: 907-910.
48. Carr CM, Chaudhry C, Kim PS (1997) Influenza hemagglutinin is spring-loaded by a metastable native conformation. *Proc Natl Acad Sci USA* 94: 14306-14313.
49. Chen B, Vogan EM, Gong H, Skehel JJ, Wiley DC, Harrison SC. (2005) Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature* 433: 834-841.

**Table 1 – Folding and Stability of Ebo6H and Ebo5H under Various Buffering Conditions.**

Buffering condition	$T_m$ (°C) <sup>A</sup>	$\Delta H_{vH}$ (kcal/mol) <sup>A</sup>
<b><i>Ebo6H</i></b>		
20 mM NaOAc, pH 5.3	$86.8 \pm 2.0$	$-28.2 \pm 1.0$
20 mM NaOAc, pH 5.8	$77.8 \pm 1.6$	$-24.0 \pm 0.4$
20 mM NaHPO <sub>4</sub> , pH 6.1	$49.8 \pm 1.1$	$-20.4 \pm 0.2$
<b><i>Ebo5H</i></b>		
20 mM NaOAc, pH 4.8	$82.2 \pm 1.2$	$-27.7 \pm 0.6$
20 mM NaOAc, pH 5.3	$80.6 \pm 0.7$	$-27.6 \pm 0.4$
20 mM NaOAc, pH 5.8	$77.2 \pm 0.8$	$-28.9 \pm 0.3$
20 mM NaHPO <sub>4</sub> , pH 6.1	$72.6 \pm 1.9$	$-18.1 \pm 0.2$

<sup>A</sup> Errors listed here represent 95% confidence interval from data fitting.

Accepted Article

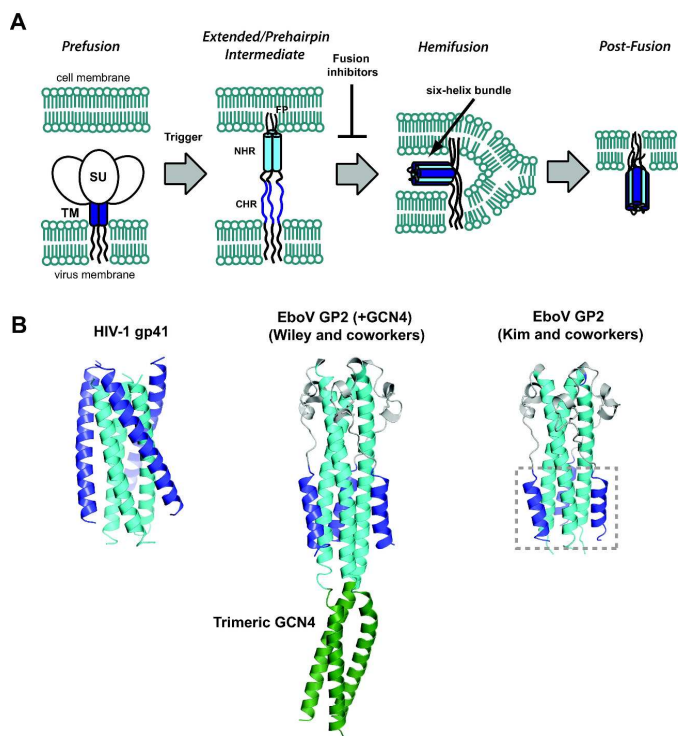


Figure 1. (A) General mechanism of viral membrane fusion catalyzed by  $\alpha$ -helical bundle-containing ectodomains. In both HIV-1 and EboV, the prefusion assembly consists of a dimer of trimers containing a surface subunit (SU, gp120 for HIV-1 and GP1 for EboV) and the transmembrane subunit that contains the NHR and CHR regions (TM, gp41 for HIV-1 and GP2 for EboV). In both cases a 'trigger' leads to insertion of the N-terminal fusion peptide (FP) into the cell membrane resulting in the 'extended' or 'prehairpin' intermediate in which the NHR (cyan) and CHR (blue) are exposed to the environment. Folding of the six-helix bundle juxtaposes the two membranes and ultimately leads to their coalescence. Agents that bind the extended/prehairpin intermediate and prevent six-helix bundle formation have been shown to act as potent antiviral agents in HIV-1 ('fusion inhibitors'). (B) Post-fusion conformations of HIV-1 gp41 (PDB ID 1AIK) and EboV GP2 ectodomains solved by the Wiley (PDB ID 1EBO) and Kim groups (PDB ID 2EBO). The NHR and CHR elements are colored as in panel A; the Wiley GP2 construct contained an N-terminal trimeric GCN4 segment which is shown in green. The six-helix bundle of EboV GP2 is boxed on the structure reported by Kim and coworkers.

328x292mm (300 x 300 DPI)

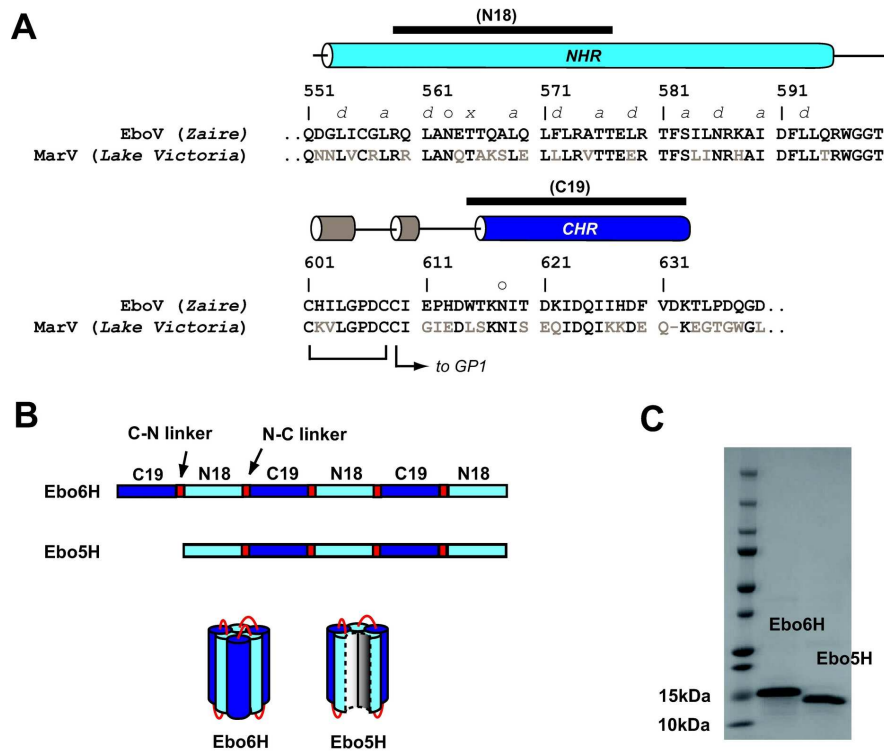


Figure 2. Protein design of mimics of the EboV GP2  $\alpha$ -helical bundle. (A) Sequence of GP2 ectodomain from EboV (Zaire strain). The regions corresponding to the NHR, CHR, and helix-turn-helix regions are indicated with cylinders whose colors match the structural elements depicted in Figure 1B. The helix-turn-helix region contains three cysteines, two of which form an intramolecular disulfide bond and the third tethers GP1 to GP2. The ectodomain contains two glycosylation sites, these are indicated with an 'o'. The T565 residue that gives rise to the unusual 3-4-4-3 hydrophobic stutter in the NHR is indicated with an 'x', other residues in the NHR that form core positions are denoted a and d. For comparison, the sequence of the related marburgvirus (MarV, Lake Victoria strain) is also shown; conserved residues are colored black. The segments used in design of Ebo6H and Ebo5H (N18 and C19) are indicated with a black bar above. (B) Schematic representation of Ebo6H and Ebo5H design. Ebo6H consists of alternating C19 and N18 segments linked by short linkers (C-N linker sequence -GSSGG- and N-C linker sequence -GGSGG-). Ebo5H is similar but lacking the first C19 sequence. (C) SDS-PAGE analysis of purified, refolded Ebo6H and Ebo5H. In both cases, a single band was observed of expected molecular weight (16 kDa for Ebo6H and 15 kDa for Ebo5H).

179x135mm (300 x 300 DPI)

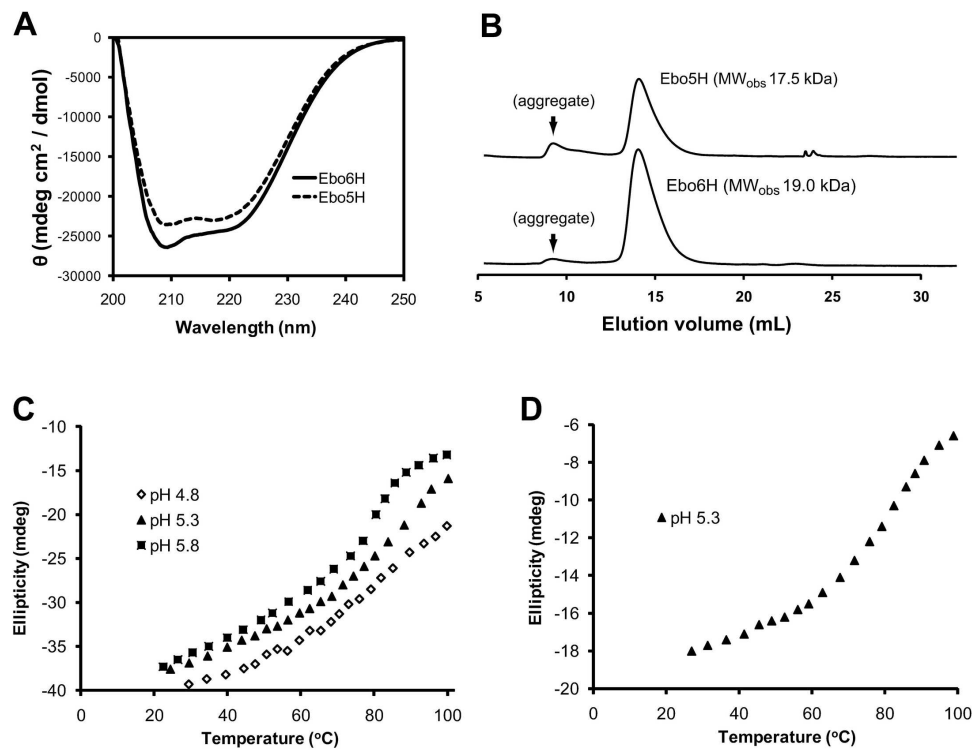


Figure 3. Characterization of Ebo6H and Ebo5H. (A) CD spectra of Ebo6H and Ebo5H in 20 mM sodium acetate, pH 5.3. (B) Gel filtration analysis under similar conditions. Observed molecular weight estimates (19.0 kDa and 17.5 kDa for Ebo6H and Ebo5H, respectively) were consistent with monomer in each case. A small amount of aggregate was observed in each case. (C and D) Thermal denaturation of Ebo6H (C, pH 5.8, pH 5.3, and pH 4.8) and Ebo5H (D, pH 5.3).

192x143mm (300 x 300 DPI)

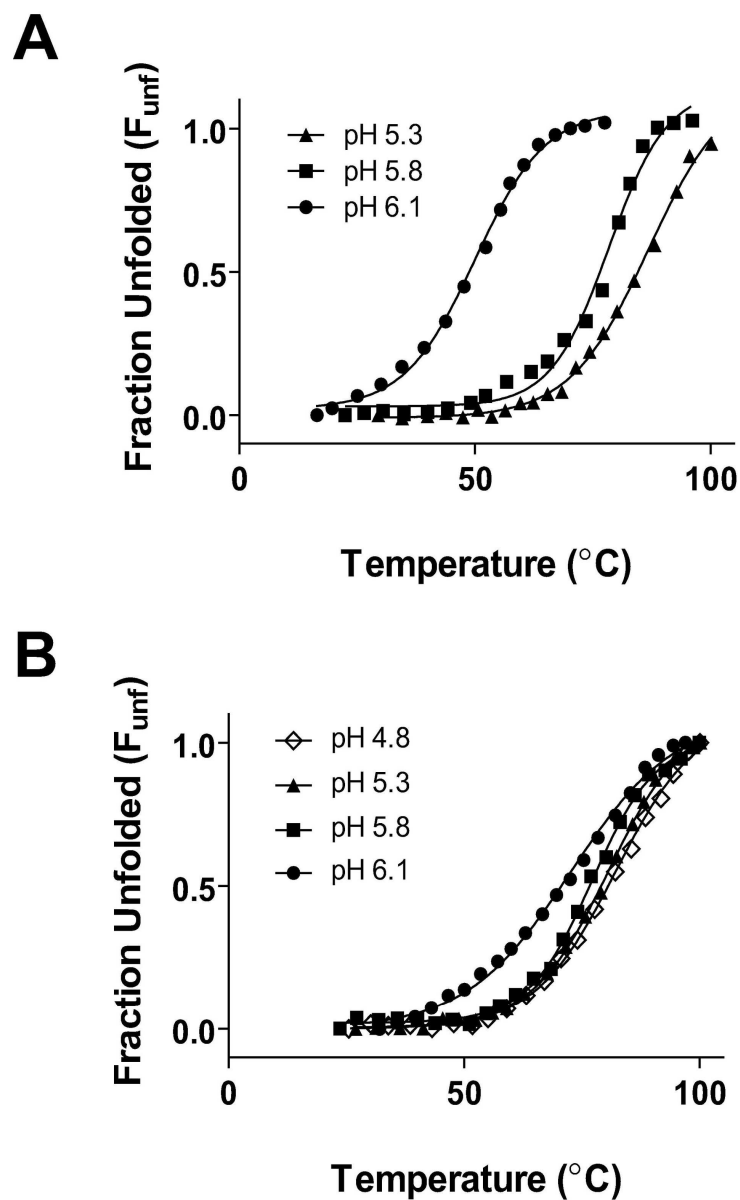


Figure 4. Thermal denaturation of Ebo6H and Ebo5H under various buffering conditions. Corresponding  $T_m$  and  $\Delta H_{\text{vH}}$  values are listed in Table 1.

161x266mm (300 x 300 DPI)

Accepted Article

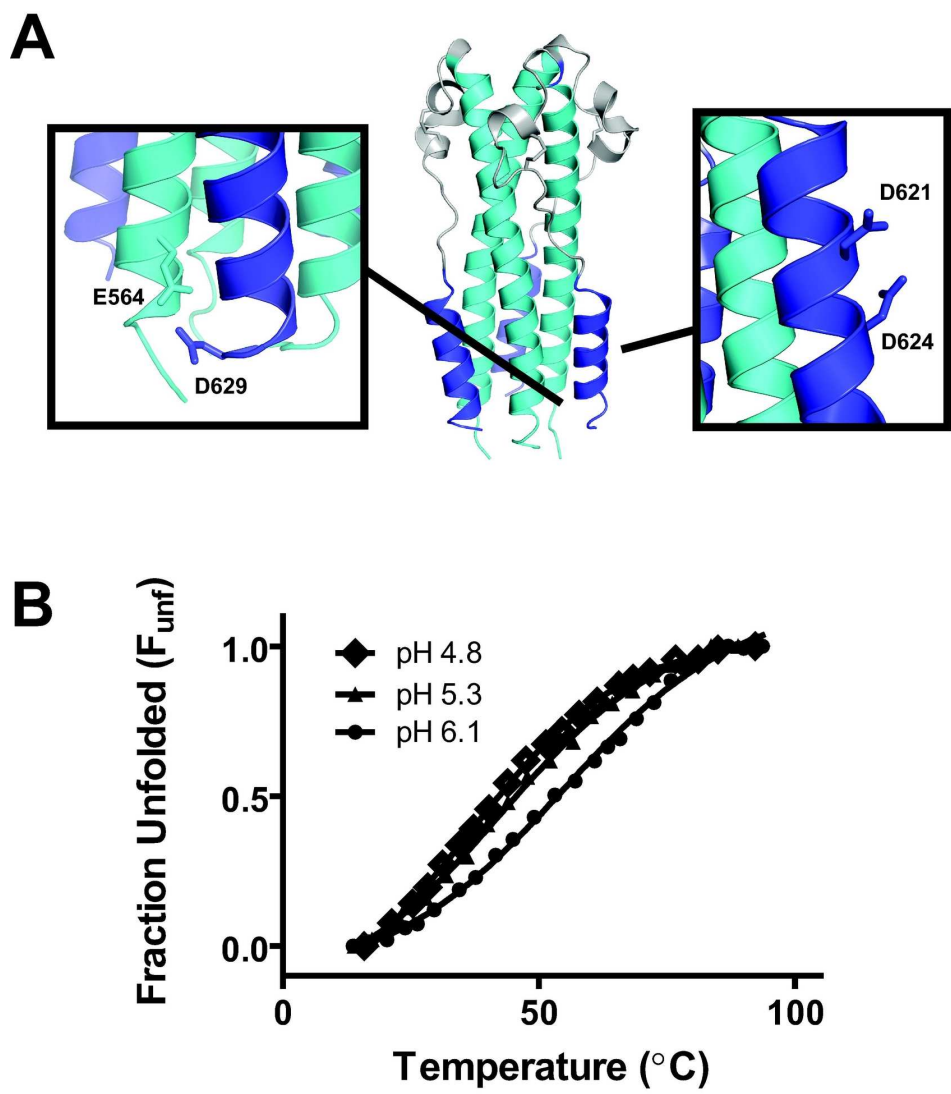


Figure 5. (A) Side chain-side chain interactions that may contribute to pH-dependent  $\alpha$ -helical fold stability. (B) Thermal denaturation of Ebo6H-QNNN in 20 mM acetate at pH 4.8, 5.3, and 6.1. The  $T_m$ s were  $43.1 \pm 0.9$  oC,  $46.2 \pm 1.0$  oC, and  $53.5 \pm 1.1$  oC, respectively.