# Composition and Functions of the Influenza Fusion Peptide

Karen J. Cross[1], William A. Langley[2], Rupert J. Russell[3], John J. Skehel[1] and David A. Steinhauer[2,*]

[1]*National Institute for Medical Research, The Ridgeway, Mill Hill, London NW71AA, UK;* [2]*Department of Microbiology and Immunology, Emory University School of Medicine, 1510 Clifton Road, Atlanta, Georgia 30322, USA;* [3]*School of Biology, University of St. Andrews, Fife KY16 9ST, UK*

**Abstract:** Fusion of the influenza virus envelope with the endosomal membrane of host cells is mediated by the hemagglutinin glycoprotein (HA). The most conserved region of HA is at the N-terminus of the HA2 subunit, a relatively hydrophobic sequence of amino acids referred to as the fusion peptide. This domain is critical both for setting the trigger for fusion and for destabilizing target membranes during the fusion process. The "trigger" is set by cleavage of the HA precursor polypeptide, when the newly-generated HA2 N-terminal fusion peptide positions itself into the trimer interior and makes contacts with ionizable residues to generate a fusion competent neutral pH structure. This essentially "primes" the HA such that subsequent acidification of the endosomal environment can induce the irreversible conformational changes that result in membrane fusion. A key component of these acid-induced structural rearrangements involves the extrusion of the fusion peptide from its buried position and its relocation to interact with the target membrane. The role of the fusion peptide for both priming the neutral pH structure and interacting with cellular membranes during the fusion process is discussed.

## INTRODUCTION

During the initial stages of the replication cycle enveloped viruses must fuse their membranes with those of host cells to introduce their genetic material. Depending on the virus, fusion can take place at the plasma membrane or within internal organelles of the cell following endocytosis of virions. A number of diverse strategies have evolved to bring about the fusion process, which are all dependent on properties of viral fusion proteins (VFPs). These are oligomeric glycoproteins that are anchored in the viral membrane due to hydrophobic transmembrane sequences in the C-terminal region of the protein. A common feature of VFPs is that they can be triggered by external stimuli to undergo structural rearrangements that result in the exposure of a second relatively hydrophobic peptide domain that can interact with cellular target membranes as a pre-condition for the fusion event. These domains are referred to as fusion peptides, and for VFPs to be functional the fusion peptide domains need to satisfy several requirements. They need to be integrated into full-length polypeptide chains that fold into the native conformation, transport to the site of virus assembly, and incorporate into infectious virions. These proteins also need to be capable of being triggered at the appropriate time and place to undergo the conformational changes requisite for fusion. As a result of these conformational changes, the VFPs must be able to direct fusion peptide domains to the target membrane, whereupon they need to functionally integrate into the cellular membrane in a fashion that leads to fusion and delivery of virion contents to the cellular interior.

The VFPs of enveloped viruses share a number of common functional properties [1], but are often grouped into three classes, I, II, and III, based on structural and mechanistic considerations [2, 3]. For most class I VFPs, which include the fusion proteins of orthomyxoviruses, paramyxoviruses, and retroviruses among others, proteolytic cleavage of precursor polypeptide trimers generates membrane anchored fusogenic subunits that contain their fusion peptide domains directly at the newly created N-terminal ends. During fusion, these N-terminal fusion peptides interact with cellular membranes to physically link them to viral membranes, and associated conformational changes lead to the formation of highly thermostable α-helical rod-like structures that draw the two membranes into proximity with one another. The interior of these rod-like structures is characterized by long triple-stranded coiled coils containing α-helical components from each monomer of the trimer. Antiparallel polypeptide chains from each monomer pack against the central coiled coils on the outside of these structures to position the N-terminal fusion peptides at the same end as the viral transmembrane domains [4]. The influenza hemagglutinin glycoprotein is a prototype of the class I VFPs, belonging to a subset of these proteins that are triggered to mediate fusion in endosomes as they become acidified. An abundance of structural information exists for HA, and it has been well characterized both functionally and biochemically. There have also been a number of studies on the fusion activity of expressed HA mutants, virus variants with alternative fusion phenotypes, and viruses containing specifically mutated HAs generated by reverse genetics. In this review we concentrate on the role of the fusion peptide domain for HA-mediated membrane fusion. We consider the composition of this conserved sequence, how it integrates as a structural component of each of the three major conformations the HA assumes during the virus life cycle, and how these structural properties relate to the functional and biological data derived from the study of mutants.

*Address correspondence to this author at the Department of Microbiology and Immunology, Emory University School of Medicine, 1510 Clifton Road, Atlanta, Georgia 30322, USA; Tel: 404-712-8542;
E-mail: steinhauer@microbio.emory.edu

0929-8665/09 $55.00+.00 

## HA STRUCTURE

High resolution structures have been determined for proteolytic fragments of each predominant conformation of HA; the HA0 polypeptide precursor, the neutral pH structure that is generated by cleavage of HA0, and the highly thermostable rod-like structure that HA folds into during the fusion process [5-8]. The structures of these are shown in Fig. (**1**) as a lateral view and in Fig. (**2**) viewed down the three-fold axis of symmetry. The residues that constitute the fusion peptide have functional significance as components of each of these structures, which will be addressed in detail following an overall description of the HA structures and conformational transitions. Unless specifically noted, the descriptions here will relate to the HA of A/Aichi/2/68 virus, a prototype of the H3N2 subtype "Hong Kong" viruses that emerged in humans in 1968 and continue to circulate.

The Aichi HA is synthesized as precursor polypeptide chains (HA0), that fold and associate as non-covalently linked homotrimers in the endoplasmic reticulum (ER) prior to transport through the Golgi apparatus to the plasma membrane. HA is a classical type I membrane glycoprotein with an N-terminal signal sequence that is cleaved in the ER, a hydrophobic transmembrane anchor domain near its C-terminus, and a short cytoplasmic tail sequence. It is post-translationally modified by the addition of acyl chains at three cysteine residues in the cytoplasmic tail and its transmembrane interface [9, 10], and is N-glycosylated at seven asparagine residues in the ectodomain [8]. The trimeric ectodomain extends by over 130 angstroms from its membrane anchor to the tip of the distal globular head domains, where the receptor binding sites and major antigenic regions reside [8, 11]. The fibrous stem of the molecule features a long central coiled coil that is formed by the association of three helices of 50 residues, with each monomer contributing one of these. A hallmark of the HA0 precursor is the presence of a surface loop structure in the membrane-proximal third of the molecule [7]. Proteolytic cleavage at this site is required for virus infectivity as it allows the HA to assume the fusion competent conformation that can be triggered to mediate membrane fusion [12-14].

The proteolytic activation of each HA0 polypeptide yields subunits of 328 residues (HA1) and 221 residues (HA2), which are linked by a single disulfide bond between HA1 14 and HA2 137. The arginine at the cleavage site, residue 329, is subsequently removed by carboxypeptidase activity [15]. The structure of the HA0 precursor and that of cleaved, neutral pH HA are very similar, with only six residues at the C-terminus of the newly formed HA1 subunit and 12 residues at the N-terminus of HA2 actually relocating as a result [7]. However, the structural changes that result are not without significance for fusion. In particular, the highly conserved N-terminus of HA2, which occupies the membrane-proximal half of the cleavage loop in the HA0 structure, is repositioned to the interior of the trimer to fill a cavity that is present in HA0. The HA0 cavity is lined with ionizable residues that are exposed to solvent in the HA precursor struc-

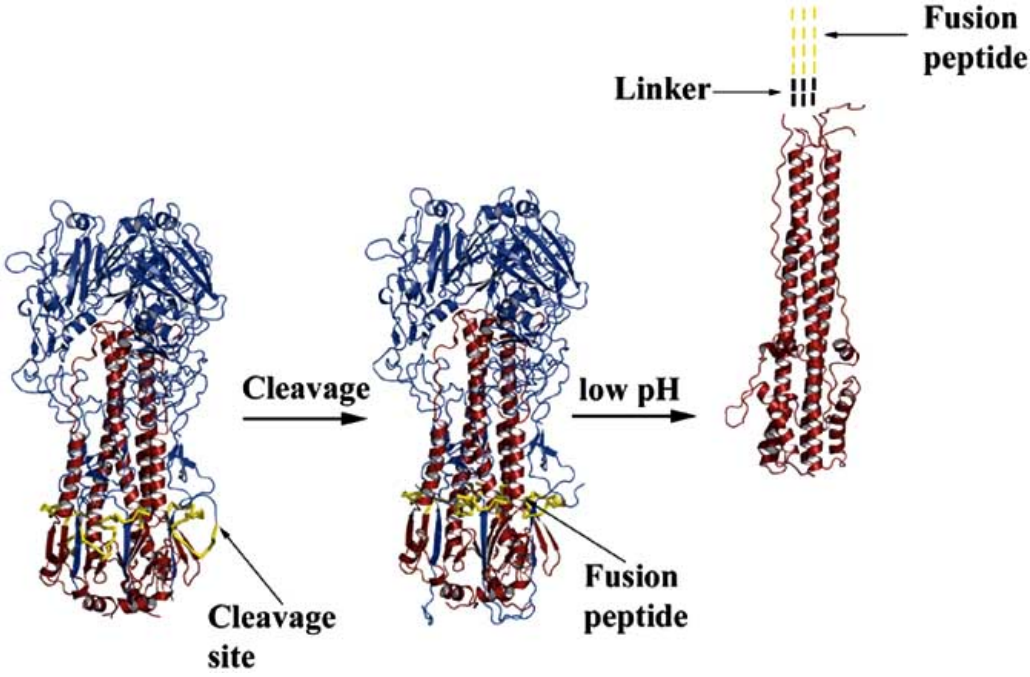


**Figure 1.** Ribbon diagram of the three conformations of the HA trimer from a lateral view. The left panel shows the HA0 precursor structure prior to proteolytic cleavage of the monomers. Residues that will constitute the HA1 subunit following cleavage are depicted in blue, and the HA2 subunit in red. Residues that ultimately form the fusion peptide at the N-terminus of HA2 are shown in yellow. The cleavage loop from one of the monomers can be seen extending from the trimer surface to the right. The center panel shows that most of the HA structure remains unchanged following cleavage, the major structural consequence being the relocation of N-terminal HA2 "fusion peptide" residues from the bottom of the cleavage loop to the interior of the trimer. The right panel shows the structure of the HA2 trimer following the acid-induced conformational changes required for membrane fusion, which are described in detail in the text. This thermostable rod-like structure locates the fusion peptide and the viral transmembrane domain at the same end. Hashed lines indicate that the structure is unknown for of both the fusion peptide and the 10-residue peptide that links it to HA2 residues of known structure.

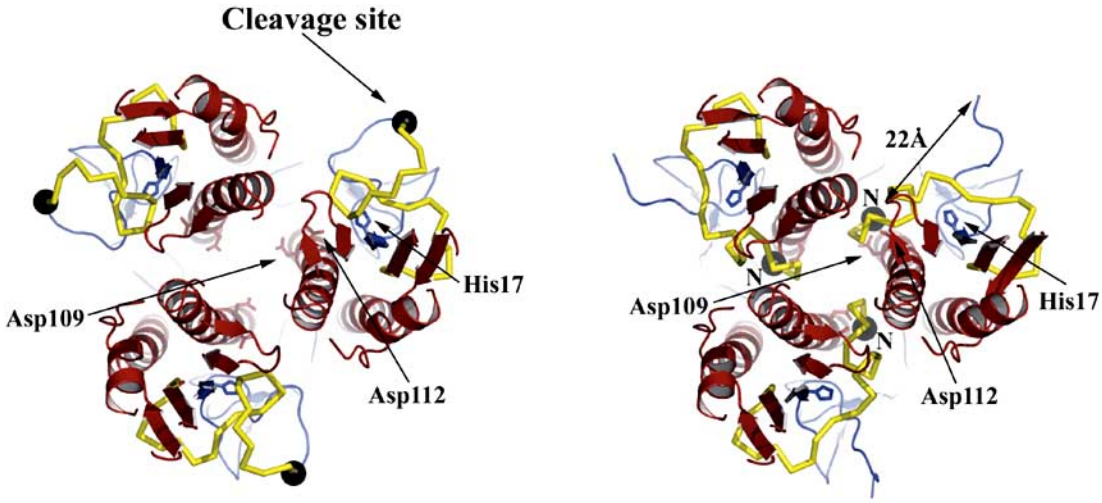


**Figure 2.** Structure of HA0 and cleaved neutral pH HA viewed down the three-fold axis of symmetry. Colors are as described for figure 1. This view illustrates more clearly the relocation of fusion peptide residues at the N-terminus of HA2 from the surface loop of uncleaved HA0 to the interior of the trimer, whereupon the N-terminus of HA2 and the C-terminus of HA1 are separated by 22Å. The location in HA0 of the glycine residues that will constitute the HA2 N-termini following cleavage are shown as a filled black circles in the left panel, and as filled gray circles labeled N in the cleaved HA structure on the right. Also depicted are the locations of ionizable residues HA1 His 17 and HA2 Asp 109 and Asp112, which are buried by fusion peptide residues following cleavage.

ture, but the relocation of HA2 amino acids 1-12 effectively buries these residues. This appears to prime the HA for fusion, as HA0 is relatively unresponsive to reduction in pH whereas cleaved HA can subsequently be triggered by acidification to mediate fusion. It is thought that a selection of the ionizable residues that are buried by fusion peptide residues may be important for triggering the fusion process [7, 16, 17], and this is discussed in more detail below.

## CONFORMATIONAL CHANGES REQUIRED FOR FUSION

As endosomes become acidified, a critical pH is attained at which membrane fusion is initiated. This normally occurs between pH 5.0 and 6.0 depending on the viral strain, but mutants have been identified which fuse membranes at higher or lower pH than this [18, 19]. The HA conformational changes that lead to fusion have been defined in detail by X-ray crystallography studies [5, 6, 20]. In essence, four separate aspects of these molecular rearrangements can be defined, each involving the relocation of intact structural domains relative to one another.

1) The membrane distal monomeric head domains become de-trimerized, but remain tethered to the trimeric core structure of low pH HA by HA1 residues 28-43, which are disordered in the fusion pH structure [5]. This suggests that this linking polypeptide is flexible, possibly allowing head domains to distance themselves from the HA2 fusion subunits during the fusion process.

2) The HA2 N-terminal fusion peptide domain is extruded from its buried position in the trimer interior, whereupon it is presumably relocated to interact with target membranes due to the other structural transitions.

3) The N-terminal segment of the neutral pH coiled coil becomes extended due to the transition into a helical structure of the peptide chain that links the long and short alpha helices of native HA. As a result, the triple-stranded coiled coil of neutral pH HA extends N-terminally to include residues of the linking polypeptides as well as the short helices. In essence, this functions to recruit the fusion peptide to the N-terminal end of this extended coiled coil.

4) Amino acids at HA2 positions 106-112 in the central coiled coil of neutral pH HA undergo a helix-to-loop transition. This functions to invert by 180 degrees all of the structural entities C-terminal to this. This jack knife-like structural rearrangement functions to position the membrane anchor domain at the same end of the low pH rod-like structure as the fusion peptide [20], and the proximity of the two membrane associating regions that result from this are thought to be crucial for the fusion process.

The order in which these structural rearrangements occur, the extent to which certain individual domains unfold and refold, and the high-resolution structure of any intermediates in the pathway, all remain to be determined. However, these conformational changes provide a clear basis for the mechanism by which a bridge is formed between the viral and cellular membranes and how they might be drawn into proximity with one another. These suggest a general model for membrane fusion as depicted in Fig. (**3**). A number of variations on this theme have been proposed to provide more details and justify or explain the large collection of results and observations related to fusion. However, many of the specifics involved remain the topic of study and a unifying detailed model for fusion is not currently available. Despite this, there is general agreement on the broad concept by which type I VFPs such as HA bring membranes together as a result of conformational changes, and related mechanisms have been proposed for class II and class III VFPs as well [1-3]. Among the accepted premises is the idea that fusion proceeds through a hemifusion intermediate, in which the outer leaflets of the membranes merge and lipid mixing occurs prior to full fusion and content mixing [21]. It is also possible that the composition and structure of the fusion peptides, and indeed the transmembrane anchor domains as well, play

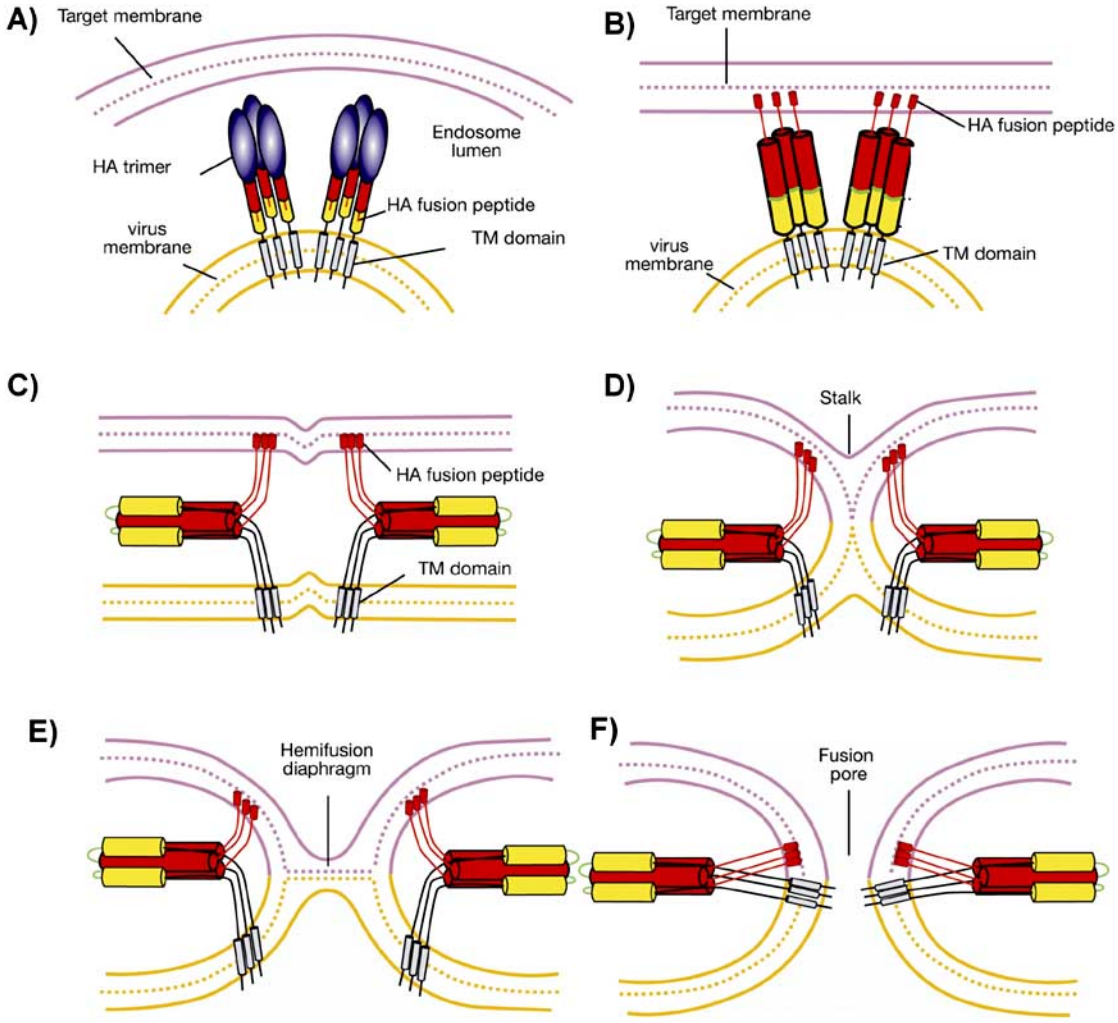


**Figure 3.** Schematic diagram of a basic model for membrane fusion by influenza HA. A) Viral membrane with two representative cleaved neutral pH HA molecules and endosomal target membrane. B) Acid-induced conformational changes lead to fusion peptide insertion into target membranes. C) Packing of the inverted C-terminal portion of HA2 against central coiled coil draws membranes together. D) Formation of pre-fusion stalk intermediate. E) Hemifusion intermediate with lipid mixing but no content mixing. F) Formation of fusion pore leading to content mixing and transfer of viral genome.

a critical role in the disruption of membranes in a fashion that proceeds to fusion.

## COMPOSITION OF FUSION PEPTIDES

All fusion peptides are reasonably hydrophobic and are capable of associating with membranes, but additional attributes are required for biologically relevant fusion function. In general, the lengths of such fusion peptides have not been rigorously defined by experimental means, but are arbitrarily based on considerations such as hydrophobicity and other features of the amino acid sequences. Fig. (**4A**) shows a comparison of fusion peptide sequences from the class I VFPs of several viruses demonstrating that there is little or no direct identity among the more distantly related taxonomic groups. As might be expected, these domains are rich in hydrophobic residues such as Phe, Leu, Ile, and Val, and most fusion peptides also contain a number of alanine residues. Concordantly, ionizable residues are very sparsely represented among these sequences. It is also notable that these fusion peptides are characterized by the presence of several interspersed glycine residues. These are often spaced at intervals that might suggest a structural significance, but they do not necessarily align directly or demonstrate definitive motifs.

On the other hand, Fig. (**4B**) shows that the fusion peptides of influenza A viruses are extremely well conserved among the 16 recognized subtypes, particularly for HA2 residues 1-11 in the N-terminal half [22]. For HA the fusion peptide is often defined as the N-terminal 23 residues of HA2, in part because residues 20-23, GWYG, are completely conserved, include two large hydrophobic residues, and extend the interspersed glycine motif. In fact the HA fusion peptide domains are by far the most highly conserved domains of this notably variable glycoprotein. Presumably, this sequence conservation is due in large part to the prominent role of this domain in the final stages of membrane fusion. However, as can be noted from the variability among viral fusion peptides illustrated in Fig. (**4A**), sequence require-

| | |
|---|---|
| Influenza A virus | **G**LF**G**AIA**G**FIEN**G**WE**G**MID**G**WY**G** |
| Influenza B virus | **G**FF**G**AIA**G**FLE**GG**WE**G**MIA**G**WH**G** |
| Influenza C virus | IF**G**IDDLII**G**LLFVAIVEA**G**I**GG**YLL**G**S |
| Parainfluenza 1 virus | FF**G**AVI**G**TIAL**G**VATAAQITA**G**IALA |
| Sendai virus | FF**G**AVI**G**TIAL**G**VATSAQITA**G**IALA |
| Respiratory syncitial virus | FL**G**FLL**G**V**G**SAIAS**G**IAVSKVLHL |
| Measles virus | FA**G**VVLA**G**AAL**G**VATAAQITA**G**IALH |
| Ebola virus | **G**AAI**G**LAWIPYF**G**PAAE |
| Simian virus 5 | FA**G**VVI**G**LAAL**G**VATAAQVTAAVALVK |
| Human immunodeficiency virus I | AV**G**IV**G**AMFL**G**FL**G**AA**G**STM**G**AVSLTLTVQA |
| Human immunodeficiency virus II | **G**VFVL**G**FL**G**FLATA**G**SAM**G**ARSLTLSA |
| Simian immunodeficiency virus | VPFVL**G**FL**G**FL**G**AA**G**TAM**G**AAATALTV |
| Human T-cell leukemia virus | AVPVAVWLVSALAM**G**A**G**VA**GG**IT**G**SMSLAS**G** |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **H1** | G | L | F | G | A | I | A | G | F | I | E | G | G | W | T | G | M | I | D | G | W | Y | G |
| **H2** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | Q | - | - | V | - | - | - | - | - |
| **H3** | - | - | - | - | - | - | - | - | - | - | - | N | - | - | E | - | - | - | - | - | - | - | - |
| **H4** | - | - | - | - | - | - | - | - | - | - | - | N | - | - | Q | - | L | - | - | - | - | - | - |
| **H5** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | Q | - | - | V | - | - | - | - | - |
| **H6** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| **H7** | - | - | - | - | - | - | - | - | - | - | - | N | - | - | E | - | L | V | - | - | - | - | - |
| **H8** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | S | - | - | - | - | - | - | - | - |
| **H9** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | P | - | L | V | A | - | - | - | - |
| **H10** | - | - | - | - | - | - | - | - | - | - | - | N | - | - | E | - | - | V | - | - | - | - | - |
| **H11** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | P | - | L | - | N | - | - | - | - |
| **H12** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | P | - | L | V | A | - | - | - | - |
| **H13** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | P | - | L | - | N | - | - | - | - |
| **H14** | - | - | - | - | - | - | - | - | - | - | - | N | - | - | Q | - | L | - | - | - | - | - | - |
| **H15** | - | - | - | - | - | - | - | - | - | - | - | N | - | - | E | - | L | - | - | - | - | - | - |
| **H16** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | P | - | L | - | N | - | - | - | - |

**Figure 4. A)** Representative fusion peptide sequences from various Class I viral fusion proteins. Glycine residues are highlighted in bold. **B)** Fusion peptide sequences from influenza A virus strains representing each of the 16 known antigenic subtypes. The H1 subtype is used as the reference sequence for this figure to keep the list in numerical order, and glycine residues are highlighted in bold. Dashes for other fusion peptide sequences indicate direct homology to the H1 sequence, and changes from the H1 reference sequence are noted with the single letter amino acid code.

ments for fusion peptide function do not appear to be particularly rigorous. Therefore, the constraints on the HA fusion peptide sequences are also indicative of a functional relevance of these residues within the structure of the HA0 precursor and the neutral pH cleaved HA, and for transitional events as one conformation of HA proceeds to the next.

## STRUCTURE OF THE FUSION PEPTIDE IN HA0

Within the context of the HA0 precursor polypeptide, the 12 N-terminal residues of the fusion peptide form the bottom half of the cleavage loop, which orients away from the trimer stalk as depicted in Figs. (**1**) and (**2**) [7]. The most obvious functional constraints for the residues that constitute the fusion peptide within HA0 involve its folding into the native, cell surface-expressed oligomeric structure, and the capacity for recognition of the cleavage loop by activating proteases. The amino acid composition of sequences directly N-terminal to conserved GLFGAIAGFI fusion peptide sequences of the loop have been shown to be particularly relevant for influenza pathogenicity [7, 23-27]. For human and most nonpathogenic avian virus strains, the HA cleavage site contains a single arginine residue, and is cleaved extracellularly by trypsin-like proteases following the transport of HA to the plasma membrane. However, certain avian strains, most notably represented by subsets of the H5 and H7 serotype viruses, contain polybasic sequences in the cleavage loop that can be recognized by ubiquitous furin-like proteases that reside in the trans-Golgi of most cell types [28, 29]. As such, protease activation of these viruses can occur intracellularly in a variety of cell and tissue types allowing them to spread more readily in systemic fashion. The upstream cleavage loop sequences of such HAs are often characterized by insertions of basic residues containing the protease recognition motifs R-X-R/K-R or derivatives of this. The recognition motifs, the greater accessibility provided by a larger cleavage loop, and the presence or absence of proximal residues that are glycosylated can all have relevance for protease activation. Clearly, any mutations to fusion peptide sequences that inhibit HA0 cleavage would be selected against in wild isolates, and therefore such mutants have not been documented. However, cleavage activation mutants with changes in fusion peptide sequences have been selected in laboratory by passaging influenza in the presence of proteases with alternative cleavage site specificity [30, 31], and an expressed HA with an alanine insertion at the N-terminus of the fusion peptide has been shown to be resistant to trypsin activation of fusion potential [32].

## STRUCTURE OF THE FUSION PEPTIDE DOMAIN IN NEUTRAL PH HA

In the HA0 structure a cavity lined with ionizable residues lies adjacent to the cleavage loop. Following proteolytic activation, the highly conserved N-terminal domain of HA2 inserts into the trimer interior and buries these ionizable residues, which are also highly or partially conserved (Fig. **2**). The newly formed contacts resulting from this structural

relocation are thought to be important for priming of the neutral pH HA such that it can subsequently be activated by acidification to mediate fusion [7, 16, 17]. This may in part explain the conserved nature of both fusion peptide and cavity-lining amino acids. Among the contacts formed when the fusion peptide is inserted into the cavity are a series of hydrogen bonds that link main chain atoms of HA2 residues 2, 3, 4, 5, and 6 to the highly conserved aspartic acids located at HA2 positions 109 and 112, which reside in a segment of the long helix of the central coiled coil that undergoes a helix-to-loop transition during fusion. Another ionizable residue, HA1 histidine 17, forms hydrogen bonds with fusion peptide residues 6 and 10 via a water molecule in H3 group HA subtypes and is thought to play a role in the initiation of acid-induced conformational changes [16, 17]. The close association between the residues in the N-terminal half of the fusion peptide and conserved ionizable residues that are buried following HA0 cleavage suggests that the HA2 N-terminal domain may be conserved not only for the purpose of functional association with target membranes during fusion, but for making appropriate contacts to allow for the triggering of the metastable neutral pH HA. It is noteworthy in this regard that nearly all single residue substitution mutations that have been examined within the first 10 amino acids of the fusion peptide lead to an elevated pH at which conformational changes are initiated [18, 32-35], and many of these are perfectly functional for membrane fusion. Therefore, strong selective pressure may exist on conserved fusion peptide residues not only for fusion capacity, but also for maintaining a neutral pH cleaved HA structure that is energetically favorable for triggering membrane fusion at an optimal pH. Such a structure would ideally balance requirements for stabilizing HA against premature triggering of the irreversible conformational changes with those for the capacity to be induced to cause fusion within the endosomal pathway. This balance may have an influence on the stability of viruses in the environment, and might account, in part, for the differences in fusion pH and kinetics that can be observed among differing HA subtypes [36, 37].

## STRUCTURE OF THE FUSION PEPTIDE DURING FUSION

Nearly all mutant HAs with single substitutions in the fusion peptide domain that have been examined as expressed proteins or as components of infectious viruses have been found capable of transporting to the cell surface and responding to acidification with the characteristic structural changes. However, a significant number of these fusion peptide mutants are inhibited for the capacity to mediate fusion, even though they retain the capacity to associate with target membranes. Therefore, it is clear that the structure adopted by the fusion peptides of influenza, and other VFPs for that matter, in association with target membranes, is critical for function. The hydrophobic nature of fusion peptides has proven to be a formidable obstacle for obtaining structural information on these domains. For example, in all X-ray crystal structures of low pH HAs determined to date, the fusion peptides have been removed in order to generate soluble proteolytic fragments capable of forming crystals. Therefore, our current knowledge on fusion peptide structures is largely based on alternative structural and biophysical techniques.

Hydrophobic photolabeling studies have shown that only the first 22 residues of HA2 are responsible for interaction with target membranes and it was suggested by the periodicity of labeling that this segment adopts an amphipathic α-helical structure [38, 39]. Biophysical studies using circular dichroism and FTIR spectroscopy suggest that membrane-bound fusion peptide analogs contain 40% to 60% α-helical structure, with some studies also showing evidence of β-structure [40-45]. In some such studies the ability of mutant HA peptides to mediate liposome fusion showed a loose correlation between helical content and fusogenicity [42, 45, 46]. The HA and other viral fusion proteins, as well as apolipoproteins, proteins of neurodegenerative diseases (β-amyloid, human PrP), signal peptides, toxins, and fusion proteins of spermatozoids, can be modeled as oblique oriented or tilted helices that possess an asymmetric gradient of hydrophobicity, which can be hypothesized to cause membrane destabilization by perturbation of regular lipid acyl chain packing [47].

Tamm and colleagues have used a combination of NMR structural data and EPR depth analysis to analyze the structure of the influenza fusion peptide in DPC micelles [48, 49]. This line of work has involved the utilization of synthetic fusion peptide analogs containing the seven residue peptide sequence GCGKKKK attached C-terminal to a 20-residue HA2 N-terminal domain to mitigate against solubility problems, and these 27 residue peptides were shown to possess pH-dependent activity in lipid mixing and erythrocyte hemolysis assays. Analysis of its structure by NMR showed the peptide to adopt an inverted V shaped structure in detergent micelles at both neutral pH and pH 5.0, with the apex of the V located at the aqueous interface [48]. The neutral pH structure comprises an N-terminal helix (residues L2 to F9) followed by a turn, stabilized by backbone hydrogen bonds (H-bonds) between residues E11, N12, G8 and F9. The C-terminal segment is less ordered with residues W14 - G20 forming an extended structure with a bend between residues G16 and M17. At pH 5, residue I10 extends the N-terminal helix and is followed by a turn, which is stabilised by a side chain H-bond between residues N12 and G8 and a backbone H-bond between residues W14 and F9. Additionally, the C-terminal half of the fusion segment folds into a short $3_{10}$ helix with rotation of residues E15 and D19, and formation of main chain H-bonds among residues 13 through 18 [48]. This allows for complete alignment of apolar residues at the bottom face of the kinked peptide, creating a hydrophobic pocket and giving the peptide a more closed and deeply inserted structure within the membrane. A conserved glycine ridge exists at the top face of the N-terminal α-helix (residues G1, G4 and G8) whereas the C-terminal segment contains charged residues at positions 11, 15 and 19. More recent NMR structural studies on mutant fusion peptide analogues that demonstrate either reduced or negative fusogenic activity suggest a structural significance for the hinge region which lies in the middle of the fusion peptide sequence [50-53]. These studies allow for the comparison of structural features displayed by mutant peptide analogs at different pH, and for the correlation of these features with the functional fusion properties of intact HAs. However, some of the interpretations derived from such studies are dependent on the assumption that fusion peptides are active in monomeric

form, and the oligomeric state of fusion peptide domains during the fusion process remains unknown at present. The effects of pH on fusion peptide structures and their relevance for membrane destabilization are also not clear.

Tryptophan fluorescence quenching and NMR spectroscopy analyses of a 25 residue HA fusion peptide analog in SDS micelles indicate that residues 2-14 exist predominately as an α-helix with residues 16–18 residing at the micelle water interface and no observable changes in insertion depth at low pH [54]. The evidence suggests that residue E11 is located in the interior of the SDS micelle and the peptides tend to self-associate during SDS gel electrophoresis, consistent with a trimeric assembly model in which fusion peptide residues E11, E15 and D19 at the hydrophilic face lie within the oligomeric interior, while hydrophobic residues I6, F9, W14 and M17 are exposed to the bilayer. Results with infectious viruses within the first 10 fusion peptide residues [33] are also compatible with a model in which helical fusion peptides orient such that the relatively polar glycine residues form a trimeric interface and the large hydrophobic side chains of residues 2, 6, and 10 reside on the surface to interact with membrane lipids [55]. Furthermore, in experiments with HA fusion peptide analogs linked to coiled coil domains to promote their trimerization, DeGrado and coworkers demonstrated that trimeric peptides promoted a greater degree of liposomal content leakage and membrane mixing than the monomeric form [56].

The issues of pH effects on fusion peptide structure and the relevance of this for fusion function are also unclear. Studies with fusion peptide analogs show that fusion activity is often higher at low pH compared to neutral pH [43-45, 49]. However, most of these comparative studies are carried out at specific pH such as 7.4 and 5.0, and are not extended to intermediate pH. This may be critical for the interpretation of the relevance of fusion peptide structures if they are considered as separate entities from the HA as a whole. For example, HA mutants that mediate fusion at pH as high as 6.5 have been characterized, and this correlates with the pH of HA ectodomain conformational changes [18, 57]. It is not clear whether the structure of fusion peptide domains at pH 6.5 more closely resembles that at lower or neutral pH, but this should be taken into consideration, as should the effects of pH and temperature on the membrane. HA-mediated membrane fusion can be promoted at neutral pH by increasing temperature or by the addition of chemical denaturants such as urea [58, 59]. In these studies the conditions under which fusion activity was detected again correlated with those at which conformational changes of HA ectodomains were observed, although it was not possible to assess structural changes that may have taken place in the fusion peptide domains.

## STUDIES OF MUTANT HA PROTEINS

There have been numerous studies on the phenotypes of mutant HAs with changes in the fusion peptide domain, particularly for the N-terminal 10 amino acids, and these provide some guidelines regarding the requirements for particular classes of amino acids at given positions for folding and fusion function. Table **1** summarizes the fusion results for a number of mutants that have been examined in our laboratory and by other investigators. This list encompasses numerous studies and observations based on the analysis of expressed HAs and mutant viruses selected in the laboratory by various means. Viruses generated by reverse genetics will be addressed separately below in order to discuss highly debilitated viruses with poor fusion activity and the selection of revertants and pseudorevertants. In Table **1**, a fusion positive phenotype is designated for HAs that demonstrate any level of full fusion activity with content mixing properties, unless stated. This compendium is limited, to a degree, in that some of the assignments of fusion positive or negative may not be directly comparable with one another when taking into account factors such as HA expression levels, the sensitivity of various assay systems employed, or differences associated with results obtained in separate laboratory settings. However, the table is meant to provide a relatively comprehensive overall indicator of which types of changes are generally tolerated at particular positions for fusion activity, and which positions are more sensitive to changes with regard to function.

One observation suggested by the data is that fusion peptide length may be important, as deletion of the N-terminal glycine residue results in a fusion negative phenotype, as did deletion of the leucine at position 2. This is consistent with results on cleavage activation mutants selected for growth in the presence of the protease thermolysin [31], which cleaves HA between HA2 Gly1 and Leu2. Mutants selected for growth in the presence of thermolysin contained single residue insertions with leucine at the HA2 N-terminus, but had the authentic fusion peptide length restored. Although these studies and those on expressed HAs showed that LFL, LIL, and LLL (amino acid codes designate three N-terminal residues) are competent for fusion [31, 32], expressed HA with the single Leu substitution at the N-terminus was shown to display very low fusion activity [32]. In fact, with the exception of alanine, all mutants with single residue substitutions for the N-terminal glycine have been shown to be negative or significantly impaired for fusion activity [31-33, 35, 60]. However, despite the low levels of full fusion activity displayed by the serine substitution mutant at HA2 position 1, it is quite efficient at promoting the lipid mixing, characteristic of a hemifusion phenotype [33, 35]. The conserved glycine at position 4 appears to accept changes reasonably well while maintaining fusion function, although the mutants that have been characterized lead to a significant increase in the pH at which this takes place [32, 34, 61]. Among expressed HAs and mutant viruses, fusion activity is evident for HAs with alanine substitution or conservative changes for the large hydrophobic residues at positions 2, 3, 6, 9, and 10 [18, 32-34, 52, 61-63]. However, alanine is not functional when substituted for the conserved glycine at position 8 [32]. On the other hand, expressed HAs with glycine substitutions for these large hydrophobic residues are significantly inhibited for fusion activity.

Not surprisingly, the less conserved positions between residues 12 and 23 seem reasonably tolerant of substitutions, whereas the highly conserved tryptophan residues at HA2 positions 14 and 21, and the conserved tyrosine at HA2 22 are less tolerant of mutations. Our data show that alanine and glycine substitution mutants at these positions were consistently negative for polykaryon formation, and this has been

**Table 1. Fusion Activity of Expressed HAs or Mutant Viruses**

| Mutations to $HA_2$ Residues | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** | **21** | **22** | **23** | **Fusion** | **References** |
| **G** | **L** | **F** | **G** | **A** | **I** | **A** | **G** | **F** | **I** | **E** | **G/N** | **G** | **W** | **T/E** | **G** | **M** | **I** | **D** | **G** | **W** | **Y** | **G** | | |
| L | F | L | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 31 |
| L | I | L | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 31 |
| L | – | L | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 31,32 |
| A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 32,35,60 |
| S | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + ($low_a$, hemifusion) | 32,33,35 |
| L | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + $(low)_a$ | 32 |
| F | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + $(low)_a$ | 32 |
| H | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + $(low)_a$ | 32 |
| I | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + $(low)_a$ | 32 |
| V | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 35 |
| E | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 34,35 |
| Q | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 35 |
| del | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 32 |
| – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 33 |
| – | I | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 62 |
| – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 33 |
| – | del | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 32 |
| – | – | I | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 63 |
| – | – | L | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 18 |
| – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | $UP_c$ |
| – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | $UP_c$ |
| – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 32 |
| – | – | – | E | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 34 |
| – | – | – | S | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 61 |
| – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 41 |
| – | – | – | – | V | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | +/– [b] | 41,$UP_c$ |
| – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 33 |
| – | – | – | – | – | M | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 18 |
| – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 33 |
| – | – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | +/– [b] | 41,$UP_c$ |
| – | – | – | – | – | – | V | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 41 |
| – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 32 |
| – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 33,52 |
| – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 33 |
| – | – | – | – | – | – | – | – | L | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 18,61 |

**(Table 1) contd….**

| Mutations to $HA_2$ Residues | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** | **21** | **22** | **23** | **Fusion** | **References** |
| **G** | **L** | **F** | **G** | **A** | **I** | **A** | **G** | **F** | **I** | **E** | **G/N** | **G** | **W** | **T/E** | **G** | **M** | **I** | **D** | **G** | **W** | **Y** | **G** | | |
| – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | + | 33,52 |
| – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 33 |
| – | – | – | – | – | – | – | – | A | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 52 |
| – | – | – | – | – | – | – | – | – | – | K | – | – | – | – | – | – | – | – | – | – | – | – | + | 65 |
| – | – | – | – | – | – | – | – | – | – | V | – | – | – | V | – | – | – | – | – | – | – | – | + | 32,64 |
| – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | + | 34,65 |
| – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – (hemifusion) | 51 |
| – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | – | G | – | – | – | – | – | 65 |
| – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | + | 51 |
| – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | 52,$UP_c$ |
| – | – | – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | $UP_c$ |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | V | – | – | – | – | – | – | – | – | + | 32 |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | + | $UP_c$ |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | + | $UP_c$ |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | + | $UP_c$ |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | + | $UP_c$ |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | + | 65 |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | K | – | – | – | – | + | 65 |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | + | $UP_c$ |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | $UP_c$ |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | $UP_c$ |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | $UP_c$ |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | G | – | – | $UP_c$ |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | + | $UP_c$ |

a: Fusion activity was detectable at low levels following more prolonged incubation at acidic pH.
b: Fusion was detected in these mutants only when a high level expression vector was utilized.
c: Unpublished previously, this data is reported here for the first time.

confirmed by others for the position 14 alanine mutant [52]. Substitutions for the conserved glutamic acid at position 11 have yielded quite variable results for fusion activity, depending on the amino acid substituted and the assay system employed [32, 34, 51, 64, 65].

## STUDIES ON VIRUSES GENERATED BY REVERSE GENETICS

Some noteworthy insights on the requirements and functional constraints of fusion peptide residues derive from studies on attempts to generate infectious viruses containing specific mutations using reverse genetics. In our laboratory we have used highly efficient methods in efforts to make several such mutants from cDNA clones [66], and have been able to rescue viruses that are clearly debilitated for replication on many occasions [33, 67, 68]. The results for experiments on the generation of mutant viruses with substitutions in the fusion peptide are summarized in Table **2**. In many cases the experiments were done with both Aichi HA (H3 subtype) and WSN HA (H1 subtype), always with the genetic background of the WSN virus. For simplicity, the table scores a mutant as positive if the virus was rescued for either HA during any experiment, although in most cases the results were the same for each HA. In some cases the efficiency of rescue was low and virus titers were reduced relative to WT [33].

**Table 2. Rescue of Infectious Viruses and Selection of Pseudorevertants**

| Mutations to $HA_2$ Residues | | | | | | | | | | | | | | | | | | | | | | | Virus Rescue[a] | Reversions |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | | |
| G | L | F | G | A | I | A | G | F | I | E | G/N | G | W | T/E | G | M | I | D | G | W | Y | G | | |
| L | – | L | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | FLL |
| S | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | G |
| L | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | F |
| – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | V |
| – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | –[b] |
| – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | V |
| – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | V |
| – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | I,V |
| – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | –[b] |
| – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | NONE |
| – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | + | A |
| – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | – | – | + | V |
| – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | – | – | – | – | –[b] |
| – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | – | + | NONE |
| – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | – | + | D |
| – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | + | NONE |
| – | – | – | – | – | – | – | – | – | – | – | – | – | F | – | – | – | – | – | – | – | – | – | + | NONE |
| – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | – | – | – | –[b] |
| – | – | – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | – | – | – | – | –[b] |
| – | – | – | – | – | – | – | – | – | – | – | – | – | V | – | – | – | – | – | – | – | – | – | – | –[b] |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | – | + | NONE |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | – | + | NONE |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | + | NONE |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | V | – | – | – | – | – | – | + | L |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | F | – | – | – | – | – | – | + | NONE |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | – | – | + | NONE |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | – | – | –[b] |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | – | + | NONE |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | + | NONE |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | – | –[b] |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | F | – | – | + | NONE |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | G | – | – | – | –[b] |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | – | – | –[b] |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | G | – | – | –[b] |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | F | – | + | NONE |
| – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | A | + | NONE |

a: Data for residues 1-10 published in reference 33. Data for residues 12-23 reported here for the first time. Data for residue 3 also reported here for the first time.
b: Original mutant not recovered.

In general, the results with rescued viruses correlate with fusion activity displayed by expressed mutant HAs, although the capacity to generate viruses would appear to be a more sensitive indicator for suboptimal levels of fusion. For example, fusion peptide mutants of HA that display activity using standard assays for fusion such as polykaryon formation, erythrocyte hemolysis, or assays of HA-mediated content mixing, can be rescued efficiently as infectious viruses. On the other hand, it has also been possible to generate a number of viruses that demonstrate little or no detectable fusion using standard assays with expressed HAs. Studies with these mutants provide us with additional data to consider from a biological perspective, as a number of them initially replicate quite poorly, and subsequently select for reversions or pseudoreversions upon passage. For example, the G1S and G1L substitutions at the HA2 N-terminus were found to promote heterokaryon formation very poorly [32] and using some assays display only hemifusion activity [33, 35]. The G1S virus reverted to WT almost immediately in one rescue experiment, but was still the predominant genotype following 3 passages in another experiment before subsequent reversion. Thus, the hemifusion phenotype ascribed to this mutant might be "leaky" enough for a low level of biological activity that is readily selected against once WT reversions are generated. Similarly, the G1L mutant was observed to mutate readily upon passage, but in this case the change involved a pseudoreversion to phenylalanine. For such examples, the possibility exists that minor representatives of virus populations could be responsible for the apparent capacity of the mutants to replicate, even as sequence analysis of total viral RNA shows that a majority of the population contains the mutation designated in the table. However, in most cases it takes at least two or three passages reflecting several rounds of replication for the reversions or pseudoreversions to appear at the RNA sequence levels, and for other examples viruses are never rescued despite many attempts using the same or different "master stocks" of parental cDNAs. This suggests at least some basal level of initial fusion activity for the mutant HA is probably required to generate virus.

It was also possible to rescue viruses with phenylalanine to glycine substitutions at either positions 3 or 9, even though fusion activity was not detected with these expressed HAs. In each case the viruses replicated poorly and pseudoreversions to valine and alanine respectively, were selected upon passage. The viruses in which large hydrophobic residues at positions 2, 3, 6, 9, and 10 were changed to alanine also proved interesting. When expressed, these mutants were positive for heterokaryon formation and were rescued as components of infectious viruses, but upon passage, pseudoreversion mutations to alternative large hydrophobic residues were selected (true reversion to WT would have required two separate changes at the RNA level for most of these). This was observed for the alanine mutants at L2, F3, I6, and I10, all of which mutated to valine after several rounds of replication.

Among the viruses rescued with substitutions at positions 12-23, changes to conserved positions W14, W21, and Y22 were notably difficult to rescue. Only the W14F mutant for WSN HA was rescued among position 14 mutants, and this replicated to titers approximately 10-fold lower than virus with WT HA. At positions 21 and 22, only the large hydrophobic substitutions of phenylalanine were rescued for tryptophan and tyrosine respectively. Among the viruses generated with changes in the C-terminal half of the fusion peptide, the only examples of pseudoreversion observed involved N12G changing to aspartic acid and M17V changing to leucine.

The major inferences derived from these studies on mutants relate to the 10 N-terminal fusion peptide residues, for which glycine appears to be operative only at the positions 1, 4, and 8, where they reside in the WT sequence, but not at other positions. Similarly, the large hydrophobic residues at positions 2, 3, 6, 9, and 10 confer a clear selective advantage. These results are consistent with models for the structure of the fusion peptide in membranes with the glycines oriented on one face of the helix and the large hydrophobic side chains interacting with the membrane bilayer. These properties can be incorporated into models of monomeric fusion peptides [48], or in those of fusion peptide oligomers with glycines in the trimer interior. The cumulative results also point to the conserved W14 as being important for fusion, based on expressed HAs and mutant viruses as reported here, and on structural and functional considerations with peptide analogs [51]. The data on mutants reported here also suggests that the conserved W21 and Y22 residues may be significant for functional association with membranes. Alanine and glycine substitutions at these positions were negative for both fusion activity and virus rescue. Reverse genetics studies on mutants with a phenylalanine substitution at position 21 shows that this appears to be well tolerated for virus replication in cell culture, but viruses with phenylalanine at position 22 displayed titers reduced by approximately two logs. The effects of changes at conserved positions W21 and Y22 might relate directly to fusion function, or alternatively, may reflect a structural significance for the initiation of fusion within the cleaved neutral pH HA. In this structure these residues are in close proximity to HA1 position 17 and HA2 position 111, where group specific histidine residues thought to be important for initiating fusion reside [16, 17]. Unfortunately, many of the studies using fusion peptide analogs utilized shorter versions of the fusion peptide domain that do not extend through conserved residues W21 and Y22.

## SUMMARY

Although a great deal of structural information exists for HA, and for the fusion peptide residues in the pre-fusion states of the molecule, the functional structure of HA fusion peptides during fusion and their oligomeric form within biological membranes remain to be elucidated. The viral transmembrane domain of HA should also be taken into consideration, as in the final fusion state following membrane merger this domain is available to interact with fusion peptides, and such interactions may play a role in the final stages of the fusion process [69, 70]. Additionally the presence of the complete HA molecule may impose constraints on the structure that the fusion peptide adopts in membrane that may not be apparent based on studies of synthetic peptide analogs. The determination of complete HA structures including transmembrane and/or fusion peptides associated with membranes will be required to fully appreciate fusion

mechanism. Indeed, the question of how many complete HA trimers participate in biologically relevant membrane fusion has been addressed but not defined with precision [71-73]. If and when a consensus emerges, the structural data will have to reconcile with the large quantity of biological information that has accumulated regarding HA fusion peptide mutants.

## ACKNOWLEDGMENTS

We thank Byeong-Jae Lee, Konrad Bradley, and Summer Galloway for critical comments on the manuscript. Support for our work has been provided by NIH Public Service Grant AI66870 (D.A.S.), NIH/NIAID contract HHSN266200 700006C, the Medical Research Council, UK (K.C., J.J.S., and R.J.R.), and the Scottish Funding Council (R.J.R.).

## REFERENCES

[1] Harrison, S. C. Viral membrane fusion. *Nat. Struct. Mol. Biol*., **2008**, *15*, 690-8.

[2] Kielian, M.; Rey, F. A. Virus membrane-fusion proteins: more than one way to make a hairpin. *Nat. Rev. Microbiol.*, **2006**, *4*, 67-76.

[3] Weissenhorn, W.; Hinz, A.; Gaudin, Y. Virus membrane fusion. *FEBS Lett.*, **2007**, *581*, 2150-5.

[4] Skehel, J. J.; Wiley, D. C. Coiled coils in both intracellular vesicle and viral membrane fusion. *Cell*, **1998**, *95*, 871-4.

[5] Bizebard, T.; Gigant, B.; Rigolet, P.; Rasmussen, B.; Diat, O.; Bosecke, P.; Wharton, S. A.; Skehel, J. J.; Knossow, M. Structure of influenza virus haemagglutinin complexed with a neutralizing antibody. *Nature*, **1995**, *376*, 92-4.

[6] Bullough, P. A.; Hughson, F. M.; Skehel, J. J.; Wiley, D. C. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature*, **1994**, *371*, 37-43.

[7] Chen, J.; Lee, K. H.; Steinhauer, D. A.; Stevens, D. J.; Skehel, J. J.; Wiley, D. C. Structure of the hemagglutinin precursor cleavage site, a determinant of influenza pathogenicity and the origin of the labile conformation. *Cell*, **1998**, *95*, 409-17.

[8] Wilson, I. A.; Skehel, J. J.; Wiley, D. C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. *Nature*, **1981**, *289*, 366-73.

[9] Kordyukova, L. V.; Serebryakova, M. V.; Baratova, L. A.; Veit, M. S-acylation of HA of influenza viruses: Mass-spectrometry reveals site-specific attachment of stearic acid to a transmembrane cysteine. *J. Virol.*, **2008**, *82*, 9288-9292.

[10] Schmidt, M. F. Acylation of viral spike glycoproteins: a feature of enveloped RNA viruses. *Virology*, **1982**, *116*, 327-38.

[11] Wiley, D. C.; Wilson, I. A.; Skehel, J. J. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature*, **1981**, *289*, 373-8.

[12] Appleyard, G.; Maber, H. B. Plaque formation by influenza viruses in the presence of trypsin. *J. Gen. Virol.*, **1974**, *25*, 351-7.

[13] Klenk, H. D.; Rott, R.; Orlich, M.; Blodorn, J. Activation of influenza A viruses by trypsin treatment. *Virology*, **1975**, *68*, 426-39.

[14] Lazarowitz, S. G.; Choppin, P. W. Enhancement of the infectivity of influenza A and B viruses by proteolytic cleavage of the hemagglutinin polypeptide. *Virology*, **1975**, *68*, 440-54.

[15] Garten, W.; Klenk, H. D. Characterization of the carboxypeptidase involved in the proteolytic cleavage of the influenza haemagglutinin. *J. Gen. Virol.*, **1983**, *64*, 2127-37.

[16] Russell, R. J.; Gamblin, S. J.; Haire, L. F.; Stevens, D. J.; Xiao, B.; Ha, Y.; Skehel, J. J. H1 and H7 influenza haemagglutinin structures extend a structural classification of haemagglutinin subtypes. *Virology*, **2004**, *325*, 287-96.

[17] Thoennes, S.; Li, Z. N.; Lee, B. J.; Langley, W. A.; Skehel, J. J.; Russell, R. J.; Steinhauer, D. A. Analysis of residues near the fusion peptide in the influenza hemagglutinin structure for roles in triggering membrane fusion. *Virology*, **2008**, *370*, 403-14.

[18] Daniels, R. S.; Downie, J. C.; Hay, A. J.; Knossow, M.; Skehel, J. J.; Wang, M. L.; Wiley, D. C. Fusion mutants of the influenza virus hemagglutinin glycoprotein. *Cell*, **1985**, *.40*, 431-9.

[19] Steinhauer, D. A.; Wharton, S. A.; Skehel, J. J.; Wiley, D. C.; Hay, A. J. Amantadine selection of a mutant influenza virus containing an acid-stable hemagglutinin glycoprotein: evidence for virus-specific regulation of the pH of glycoprotein transport vesicles. *Proc. Natl. Acad. Sci. USA*, **1991**, *88*, 11525-9.

[20] Chen, J.; Skehel, J. J.; Wiley, D. C. N- and C-terminal residues combine in the fusion-pH influenza hemagglutinin HA2, subunit to form an N cap that terminates the triple-stranded coiled coil. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 8967-72.

[21] Kemble, G. W.; Danieli, T.; White, J. M. Lipid-anchored influenza hemagglutinin promotes hemifusion, not complete fusion. *Cell*, **1994**, *76*, 383-91.

[22] Nobusawa, E.; Aoyama, T.; Kato, H.; Suzuki, Y.; Tateno, Y.; Nakajima, K. Comparison of complete amino acid sequences and receptor-binding properties among 13 serotypes of hemagglutinins of influenza A viruses. *Virology*, **1991**, *182*, 475-85.

[23] Bosch, F. X.; Garten, W.; Klenk, H. D.; Rott, R. Proteolytic cleavage of influenza virus hemagglutinins: primary structure of the connecting peptide between HA1 and HA2 determines proteolytic cleavability and pathogenicity of Avian influenza viruses. *Virology*, **1981**, *113*, 725-35.

[24] Bosch, F. X.; Orlich, M.; Klenk, H. D.; Rott, R. The structure of the hemagglutinin, a determinant for the pathogenicity of influenza viruses. *Virology*, **1979**, *95*, 197-207.

[25] Klenk, H. D.; Garten, W. Host cell proteases controlling virus pathogenicity. *Trends Microbiol.*, **1994**, *2*, 39-43.

[26] Steinhauer, D. A. Role of hemagglutinin cleavage for the pathogenicity of influenza virus. *Virology*, **1999**, *258*, 1-20.

[27] Webster, R. G.; Rott, R. Influenza virus A pathogenicity: the pivotal role of hemagglutinin**.** *Cell*, **1987**, *50*, 665-6.

[28] Stieneke-Grober, A.; Vey, M.; Angliker, H.; Shaw, E.; Thomas, G.; Roberts, C.; Klenk, H. D.; Garten, W. Influenza virus hemagglutinin with multibasic cleavage site is activated by furin, a subtilisin-like endoprotease**.** *EMBO J.*, **1992**, *11*, 2407-14.

[29] Walker, J. A.; Molloy, S. S.; Thomas, G.; Sakaguchi, T.; Yoshida, T.; Chambers, T. M.; Kawaoka, Y. Sequence specificity of furin, a proprotein-processing endoprotease, for the hemagglutinin of a virulent avian influenza virus. *J. Virol.*, **1994**, *68*, 1213-8.

[30] Orlich, M.; Linder, D.; Rott, R. Trypsin-resistant protease activation mutants of an influenza virus. *J. Gen. Virol.*, **1995**, *76*(Pt. 3), 625-33.

[31] Orlich, M.; Rott, R. Thermolysin activation mutants with changes in the fusogenic region of an influenza virus hemagglutinin. *J. Virol.*, **1994**, *68*, 7537-9.

[32] Steinhauer, D. A.; Wharton, S. A.; Skehel, J. J.; Wiley, D. C. Studies of the membrane fusion activities of fusion peptide mutants of influenza virus hemagglutinin. *J. Virol.*, **1995**, *69*, 6643-51.

[33] Cross, K. J.; Wharton, S. A.; Skehel, J. J.; Wiley, D. C.; Steinhauer, D. A. Studies on influenza haemagglutinin fusion peptide mutants generated by reverse genetics.. *EMBO J.*, **2001**, *20*, 4432-42.

[34] Gething, M. J.; Doms, R. W.; York, D.; White, J. Studies on the mechanism of membrane fusion: site-specific mutagenesis of the hemagglutinin of influenza virus. *J. Cell. Biol.*, **1986**, *102*, 11-23.

[35] Qiao, H.; Armstrong, R. T.; Melikyan, G. B.; Cohen, F. S.; White, J. M. A specific point mutant at position 1 of the influenza hemagglutinin fusion peptide displays a hemifusion phenotype. *Mol. Biol. Cell*, **1999**, *10*, 2759-69.

[36] Korte, T.; Ludwig, K.; Booy, F. P.; Blumenthal, R.; Herrmann, A. Conformational intermediates and fusion activity of influenza virus hemagglutinin. *J. Virol.*, **1999**, *73*, 4567-74.

[37] Puri, A.; Booy, F. P.; Doms, R. W.; White, J. M.; Blumenthal, R. Conformational changes and fusion activity of influenza virus hemagglutinin of the H2 and H3 subtypes: effects of acid pretreatment. *J. Virol*., **1990**, *64*, 3824-32.

[38] Durrer, P.; Galli, C.; Hoenke, S.; Corti, C.; Gluck, R.; Vorherr, T.; Brunner, J. H+-induced membrane insertion of influenza virus hemagglutinin involves the HA2 amino-terminal fusion peptide but not the coiled coil region. *J. Biol. Chem.*, **1996**, *271*, 13417-21.

[39] Harter, C.; James, P.; Bachi, T.; Semenza, G.; Brunner, J. Hydrophobic binding of the ectodomain of influenza hemagglutinin to membranes occurs through the "fusion peptide". *J. Biol. Chem*., **1989**, *264*, 6459-64.

[40] Gray, C.; Tatulian, S. A.; Wharton, S. A.; Tamm, L. K. Effect of the N-terminal glycine on the secondary structure, orientation, and interaction of the influenza hemagglutinin fusion peptide with lipid bilayers. *Biophys. J*., **1996**, *70*, 2275-86.

[41] Han, X.; Steinhauer, D. A.; Wharton, S. A.; Tamm, L. K. Interaction of mutant influenza virus hemagglutinin fusion peptides with

lipid bilayers: probing the role of hydrophobic residue size in the central region of the fusion peptide. *Biochemistry*, **1999**, *38*, 15052-9.

[42] Lear, J. D.; DeGrado, W. F. Membrane binding and conformational properties of peptides representing the NH2 terminus of influenza HA-2. *J. Biol. Chem.*, **1987**, *262*, 6500-5.

[43] Murata, M.; Sugahara, Y.; Takahashi, S.; Ohnishi, S. pH-dependent membrane fusion activity of a synthetic twenty amino acid peptide with the same sequence as that of the hydrophobic segment of influenza virus hemagglutinin. *J. Biochem.*, **1987**, *102*, 957-62.

[44] Rafalski, M.; Ortiz, A.; Rockwell, A.; van Ginkel, L. C.; Lear, J. D.; DeGrado, W. F.; Wilschut, J. Membrane fusion activity of the influenza virus hemagglutinin: interaction of HA2 N-terminal peptides with phospholipid vesicles. *Biochemistry*, **1991**, *30*, 10211-20.

[45] Wharton, S. A.; Martin, S. R.; Ruigrok, R. W.; Skehel, J. J.; Wiley, D. C. Membrane fusion by peptide analogues of influenza virus haemagglutinin. *J. Gen. Virol.*, **1988**, *69*, 1847-57.

[46] Burger, K. N.; Wharton, S. A.; Demel, R. A.; Verkleij, A. J. The interaction of synthetic analogs of the N-terminal fusion sequence of influenza virus with a lipid monolayer. Comparison of fusion-active and fusion-defective analogs. *Biochim. Biophys. Acta*, **1991**, *1065*, 121-9.

[47] Brasseur, R. Tilted peptides: a motif for membrane destabilization. hypothesis, . *Mol. Membr. Biol*, **2000**, *17*, 31-40.

[48] Han, X.; Bushweller, J. H.; Cafiso, D. S.; Tamm, L. K. Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. *Nat. Struct. Biol.*, **2001**, *8*, 715-20.

[49] Han, X.; Tamm, L. K. A host-guest system to study structure-function relationships of membrane fusion peptides. *Proc. Natl. Acad. Sci. USA*, **2000**, *97*, 13097-102.

[50] Hsu, C. H.; Wu, S. H.; Chang, D. K.; Chen, C. Structural characterizations of fusion peptide analogs of influenza virus hemagglutinin. Implication of the necessity of a helix-hinge-helix motif in fusion activity. *J. Biol. Chem.*, **2002**, *277*, 22725-33.

[51] Lai, A. L.; Park, H.; White, J. M.; Tamm, L. K. Fusion peptide of influenza hemagglutinin requires a fixed angle boomerang structure for activity. *J. Biol. Chem.*, **2006**, *281*, 5760-70.

[52] Lai, A. L.; Tamm, L. K. Locking the kink in the influenza hemagglutinin fusion domain structure. *J. Biol. Chem.*, **2007**, *282*, 23946-56.

[53] Li, Y.; Han, X.; Lai, A. L.; Bushweller, J. H.; Cafiso, D. S.; Tamm, L. K. Membrane structures of the hemifusion-inducing fusion peptide mutant G1S and the fusion-blocking mutant G1V of influenza virus hemagglutinin suggest a mechanism for pore opening in membrane fusion. *J. Virol.*, **2005**, *79*, 12065-76.

[54] Chang, D. K.; Cheng, S. F.; Deo Trivedi, V.; Yang, S. H. The amino-terminal region of the fusion peptide of influenza virus hemagglutinin HA2 inserts into sodium dodecyl sulfate micelle with residues 16-18 at the aqueous boundary at acidic pH. Oligomerization and the conformational flexibility. *J. Biol. Chem.*, **2000**, *275*, 19150-8.

[55] Skehel, J. J.; Cross, K.; Steinhauer, D.; Wiley, D. C. Influenza fusion peptides. *Biochem. Soc. Trans*, **2001**, *29*, 623-6.

[56] Lau, W. L.; Ege, D. S.; Lear, J. D.; Hammer, D. A.; DeGrado, W. F. Oligomerization of fusogenic peptides promotes membrane fusion by enhancing membrane destabilization. *Biophys. J.*, **2004**, *86*, 272-84.

[57] Wharton, S. A.; Skehel, J. J.; Wiley, D. C. Studies of influenza haemagglutinin-mediated membrane fusion. *Virology*, **1986**, *149*, 27-35.

[58] Carr C.M.; Chaudhry, C.; Kim P.S. Influenza hemagglutinin is spring-loaded by a metastable native conformation. *Proc. Natl. Acad. Sci. USA*, **1997**, *94*, 14306-14313.

[59] Ruigrok, R. W.; Martin, S. R.; Wharton, S. A.; Skehel, J. J.; Bayley, P. M.; Wiley, D. C. Conformational changes in the hemagglutinin of influenza virus which accompany heat-induced fusion of virus with liposomes. *Virology*, **1986**, *155*, 484-97.

[60] Walker, J. A.; Kawaoka, Y. Importance of conserved amino acids at the cleavage site of the haemagglutinin of a virulent avian influenza A virus. *J. Gen. Virol.*, **1993**, *74*, 311-4.

[61] Lin, Y. P.; Wharton, S. A.; Martin, J.; Skehel, J. J.; Wiley, D. C.; Steinhauer, D. A. Adaptation of egg-grown and transfectant influenza viruses for growth in mammalian cells: selection of hemagglutinin mutants with elevated pH of membrane fusion. *Virology*, **1997**, *233*, 402-10.

[62] Verhoeyen, M.; Fang, R.; Jou, W. M.; Devos, R.; Huylebroeck, D.; Saman, E.; Fiers, W. Antigenic drift between the haemagglutinin of the Hong Kong influenza strains A/Aichi/2/68 and A/Victoria/3/75. *Nature*, **1980**, *286*, 771-6.

[63] Yewdell, J. W.; Taylor, A.; Yellen, A.; Caton, A.; Gerhard, W.; Bachi, T. Mutations in or near the fusion peptide of the influenza virus hemagglutinin affect an antigenic site in the globular region. *J. Virol.*, **1993**, *67*, 933-42.

[64] Korte, T.; Epand, R. F.; Epand, R. M.; Blumenthal, R. Role of the Glu residues of the influenza hemagglutinin fusion peptide in the pH dependence of fusion activity. *Virology*, **2001**, *289*, 353-61.

[65] Nobusawa, E.; Hishida, R.; Murata, M.; Kawasaki, K.; Ohnishi, S.; Nakajima, K. The role of acidic residues in the "fusion segment" of influenza A virus hemagglutinin in low-pH-dependent membrane fusion. *Arch. Virol.*, **1995**, *140*, 865-75.

[66] Neumann, G.; Watanabe, T.; Ito, H.; Watanabe, S.; Goto, H.; Gao, P.; Hughes, M.; Perez, D. R.; Donis, R.; Hoffmann, E.; Hobom, G.; Kawaoka, Y. Generation of influenza A viruses entirely from cloned cDNAs. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 9345-50.

[67] Li, Z. N.; Lee, B. J.; Langley, W. A.; Bradley, K. C.; Russell, R. J.; Steinhauer, D. A. Length requirements for membrane fusion of influenza virus hemagglutinin peptide linkers to transmembrane or fusion peptide domains. *J. Virol.*, **2008**, *82*, 6337-48.

[68] Li, Z. N.; Mueller, S. N.; Ye, L.; Bu, Z.; Yang, C.; Ahmed, R.; Steinhauer, D. A. Chimeric influenza virus hemagglutinin proteins containing large domains of the Bacillus anthracis protective antigen: protein characterization, incorporation into infectious influenza viruses, and antigenicity. *J. Virol.*, **2005**, *79*, 10003-12.

[69] Melikyan, G. B.; Markosyan, R. M.; Roth, M. G.; Cohen, F. S. A point mutation in the transmembrane domain of the hemagglutinin of influenza virus stabilizes a hemifusion intermediate that can transit to fusion. *Mol. Biol. Cell*, **2000**, *11*, 3765-75.

[70] Chang, D. K.; Cheng, S. F.; Kantchev, E. A.; Lin, C. H.; Liu, Y. T. Membrane interaction and structure of the transmembrane domain of influenza hemagglutinin and its fusion peptide complex. *BMC Biol.*, **2008**, *6*, 2.

[71] Bentz, J. Minimal aggregate size and minimal fusion unit for the first fusion pore of influenza hemagglutinin-mediated membrane fusion. *Biophys. J.*, **2000**, *78*, 227-45.

[72] Blumenthal, R.; Sarkar, D. P.; Durell, S.; Howard, D. E.; Morris, S. J. Dilation of the influenza hemagglutinin fusion pore revealed by the kinetics of individual cell-cell fusion events. *J. Cell. Biol.*, **1996**, *135*, 63-71.

[73] Danieli, T.; Pelletier, S. L.; Henis, Y. I.; White, J. M. Membrane fusion mediated by the influenza virus hemagglutinin requires the concerted action of at least three hemagglutinin trimers. *J. Cell. Biol.*, **1996**, *133*, 559-69.

**Received: October 12, 2008 Revised: December 11, 2008 Accepted: December 11, 2008**