Structural biology of the influenza virus fusion peptide

Remigiusz Worch

Institute of Physics, Polish Academy of Sciences, Warsaw, Poland

The release of influenza RNA inside the host cell occurs through the fusion of two membranes, the viral envelope and that of the cellular endosome. The fusion is mediated by the influenza hemagglutinin protein (HA), in particular by the fusion peptide (HAfp) located in the N-terminal fragment of HA2 subunit. This protein fragment anchors in the internal endosomal membrane, whereas the C-terminal HA2 part comprises a transmembrane domain (TMD) embedded in the viral envelope. A drop of pH in the endosome acts as the main trigger for HA2 large conformational change that leads to anchoring of the fusion peptide, close contact of the membranes and the subsequent fusion. Throughout the years the major research effort was focused on a 20-aminoacid fragment (HAfp1-20), shown by NMR to adopt a ‘boomerang’-like structure. However, recent studies showed that extending HAfp1-20 by three highly conserved residues W21-Y22-G23 leads to formation of a unique, tight helical hairpin structure. This review summarizes recently discovered structural aspects of influenza fusion peptides and their relations with the membrane fusion mechanism.

Key words: influenza virus, viral entry, membrane fusion, fusion peptide, peptide-lipid interactions

Received: 30 May, 2014; revised: 14 August, 2014; accepted: 18 August, 2014; available on-line: 08 September, 2014

INTRODUCTION

Composition and function of the influenza fusion peptide were thoroughly reviewed in 2009 by the Stein-hauer group (Cross et al., 2009), but since that time new structural aspects have emerged. This short review updates the current knowledge on the structural biology of the influenza fusion peptide, including a summary of the structures deposited in the Protein Data Bank (Table 1a). For more general topics the reader is directed to the literature related to fusion proteins (Tamm et al., 2009; Apellániz et al., 2014), membrane fusion (Chernomordik et al., 2006; Chernomordik & Kozlov, 2008; Harrison 2008; Wickner & Schekman, 2008; Schick 2010), or to the influenza hemagglutinin (Skelch & Wiley, 2000; Cross et al., 2009; Luo 2012).

Definition of the fusion peptide

The influenza fusion peptide is defined as the N-terminal part of hemagglutinin (HA) HA2 subunit, which protrudes into the endosomal membrane during virus entry (Fig. 1B). The free HA2 N-terminal part results from cleavage of hemagglutinin precursor, a glycoprotein HA0, by cellular proteases, with both cleaved subunits (HA1 and HA2) linked by a single disulfide bond. The C-terminus of the HA2 subunit is embedded in the membrane of the influenza envelope with a single transmembrane domain (TMD) (Fig. 1B). Both HA0 and cleaved subunits function as homotrimers.

The length of the fusion peptide is not clearly defined (Epand, 2003). This shortcoming comes from the fact that, the fusion peptides are usually studied as short synthetic fragments of peptides, often chosen arbitrarily. Obviously, the behavior of a complete protein may be different, due to inter-chain interactions. In the case of the influenza virus, a vast majority of biophysical studies were carried out on a 20 amino acid monomeric fragment of HA2 (abbreviated here as HAfp1-20). However, recent structures were obtained for a 23 amino acid fragment (HAfp1-23). The length issue should be kept in mind when comparing ‘influenza fusion peptides’ between various research papers.

The sequence of the fusion peptide is very well conserved between the 16 recognized serotypes of the influenza A type, especially in the region comprising the first N-terminal 11 amino acids (Cross et al., 2009). The variability of sequences between serotypes is shown in Fig. 2A. Similarly to the sequences of other viral fusion peptides, influenza fusion peptide contain GxxG or GxxXG motifs, known to be responsible for interhelical interactions of transmembrane segments (Russ & Engelman, 2000; Kleiger et al., 2002). However, as these motifs were not shown to be involved in interactions with membranes, it can be speculated that these fragments can be potentially involved in inter fp or fp:TMD interactions instead (Chang et al., 2008). The fusion peptide 1-23 contains many hydrophobic amino acids with some of acidic residues at the conserved position 11 and, with a larger level of diversity, at position 15 and 19 (Fig. 2A). Since the fusion peptide net charge vary between serotypes, the role of electrostatic interactions does not seem to be a crucial driving force for HAfp attachment to the membrane. This is in agreement with the presence of only negatively charged residues in the HAfp and the negative charge of naturally occurring lipids. What is more, the HAfp comes into contact with the inner membrane of the endosome resulting from invagination of the cell surface, where the charged lipids are not present, but the composition of which is enriched with phosphatidylcho-
line, sphingolipids and cholesterol. Early endosomes are similar in composition to the plasma membrane. However, their composition is altered during maturation, mainly by a decrease in sterols and phosphatidylserine and an increase in bis(monoacylglycero)phosphate (van Meer et al., 2008). Additionally, a dedicated system of kinases and phosphatases produces and hydrolyses specific phosphoinositides (Di Paolo & De Camilli, 2006). Nevertheless, their role in membrane fusion upon the virus entry is poorly characterized.

Role in the membrane fusion

The fusion peptide has to be inserted into target cellular membrane to release RNA from the virus trapped in the endosome during entry (Fig. 1A). The mechanics of membrane fusion is usually thought to involve a hemifusion intermediate, in which lipid mixing with no aqueous content mixing occurs (i.e. lipid mixing occurs in contacting membrane leaflets but not in distal leaflets (Fig. 1C, D). This stage is transient and is transformed eventually to a fusion pore (Fig. 1E). Depending on the geometry of hemifusion contact, stalks (Fig. 1C) or hemifusion diaphragms (Fig. 1D) are identified (Colman & Lawrence, 2003; Chernomordik & Kozlov, 2008; Harrison, 2008; Wickner & Schekman, 2008). In most of the models involving the elastic theory of lipid monolayers and the self-consistent mean field theory of lipid bilayer interior, the hemifusion states are symmetric (Chernomordik & Kozlov, 2008). However, electron cryo-tomography observations of the influenza-liposome complexes showed a π-shaped hemi diaphragm instead (Lee, 2010), also observed in computer simulations (Risselada et al., 2012). Further studies are needed to elucidate transient geometries of membrane fragments upon protein-mediated fusion.

A drop of pH in the late endosome leads to a large conformational change of HA2, which opens from a ‘jackknife-like’ or ‘spring-loaded’ conformation at pH 7.4 (Cross et al., 2009; Luo, 2012). This opening leads to fp exposition, which enables its insertion into the cellular membrane. By this step, two ends of the same molecule (HA2) become embedded in two different membranes, viral and cellular, which eventually fuse (Fig. 1E). During a further conformational change, the protein folds

<table>
<thead>
<tr>
<th>PDB ID</th>
<th>Magnetic nuclei</th>
<th>pH</th>
<th>HAp sequence</th>
<th>Serotype, mutations</th>
<th>Solubility tag</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1IBN</td>
<td>$^1$H</td>
<td>5.0</td>
<td>GLFGAIAFGI</td>
<td>ENGWEGMIDG</td>
<td>H3</td>
<td>GCGKKKK</td>
</tr>
<tr>
<td>1IBO</td>
<td>$^1$H</td>
<td>7.4</td>
<td>GLFGAIAFGI</td>
<td>ENGWEGMIDG</td>
<td>H3</td>
<td>GCGKKKK</td>
</tr>
<tr>
<td>1XO0</td>
<td>$^1$H</td>
<td>5.0</td>
<td>GLFGAIAFGI</td>
<td>ENGWEGMIDG</td>
<td>H3, G15</td>
<td>GCGKKKK</td>
</tr>
<tr>
<td>1XOP</td>
<td>$^1$H</td>
<td>5.0</td>
<td>VLFGAIAFGI</td>
<td>ENGWEGMIDG</td>
<td>H3, G1V</td>
<td>GCGKKKK</td>
</tr>
<tr>
<td>2DCI</td>
<td>$^1$H</td>
<td>5.0</td>
<td>GLFGAIAFGI</td>
<td>ENGAEGMIDG</td>
<td>H3, W14A</td>
<td>GCGKKKK</td>
</tr>
<tr>
<td>2DKA</td>
<td>$^1$H, $^{13}$C, $^{15}$N</td>
<td>7.4</td>
<td>GLFGAIAFGI</td>
<td>EGGWTGMIDG WYG</td>
<td>H1</td>
<td>SGKKKKK</td>
</tr>
<tr>
<td>2L4G</td>
<td>$^1$H</td>
<td>5.0</td>
<td>GLFGAIAFGI</td>
<td>ENAWEGMIDG</td>
<td>H1, G13A</td>
<td>GCGKKKK</td>
</tr>
<tr>
<td>2LWA</td>
<td>$^1$H, $^{13}$C, $^{15}$N</td>
<td>7.3</td>
<td>GLFGAIAFGI</td>
<td>EGGWTGMIDG WYG</td>
<td>H1, G8A</td>
<td>SGKKKKK</td>
</tr>
</tbody>
</table>

Table 1a. Summary of the influenza fusion peptide structures in DPC micelles solved by liquid NMR and deposited in the Protein Data Bank.

Figure 1. Schematic representation of the influenza virus endocytosis-mediated entry into a host cell. (A) Fusion of the viral and cellular membranes (red and blue, respectively) leads to membrane fusion and eventually to viral content release. A dashed gray rectangle shows details of membrane fusion depicted in steps (B)-(E). (B) Hemagglutinin HA2 subunit is embedded in the viral envelope by its TMD and the fusion peptide anchors in the inner endosomal membrane after a pH drop in late endosomes. (C) and (D) show fusion intermediates called fusion stalk and hemifusion diaphragm, respectively. The process ends by formation of a fusion pore (E).
into a hairpin-like shape with TMD and fp at the same side of a rigid structure, leading to a close contact of the two membranes. As mentioned above, a direct fp:TMD interaction was shown (Chang et al., 2008). However, fusion of liposomes was also shown to be driven by the HA2 ectodomain itself (Kim et al., 2011). Nevertheless, anchoring of the fp in the inner endosomal leaflet is a key step in the fusion process, since all deletions along the length of the fp led to inhibition of fusion (Langley et al., 2009).

STRUCTURES OF THE INFLUENZA FUSION PEPTIDES

The presentation of structures in chronological order ends with a summary of conformational dynamics. New structure types receive more attention, while previous studies on HAfp1-20 (Cross et al., 2009; Skehel & Wiley, 2000) are only shortly summarized to highlight the main features.

A ‘boomerang’-like structure of HAfp1-20

The first structure of HAfp1-20 in membrane mimicking environment (DPC micelles) was solved by Tamm and colleagues (Han et al., 2001). In DPC micelles, HAfp1-20 adopted a conformation of an inverted V shaped structure, similar to a boomerang, with a high helical content. The insertion in the membrane was thought to orient the apex at the aqueous interface. In the N-terminal part, regardless of pH, there was a number of $i+4\rightarrow i$ hydrogen bonds stabilizing the helix (residues L2 to F9). Yet the authors pointed out a difference between the structures in acidic (fusogenic) and neutral conditions, i.e. the C-terminal part became more helical at pH 5.0 (Fig. 2C, D). This fact could be noticed by an increased number of hydrogen bonds in Fig. 3C. Additionally, at pH 5.0 the kink of the boomerang became sharper, followed by a turn of acidic residues E15 and D19. A rotation of acidic residues was also accompanied by formation of a short $3_{10}$ helix in the C-terminal part (Fig. 2D).

A tight helical hairpin structure of HAfp1-23

The most recent HAfp structure was resolved by the group of Bax (Lorieau et al., 2010). A major difference from the previous structures was the presence of the three conserved residues W21-Y22-G23 Their presence played a key role in stabilizing a highly ordered structure.

Figure 2. A variety of structures adopted by the influenza fusion peptide. Polar residues marked in gray. (A) A helical hairpin formed by H1 HAfp1-23. Residues in rectangles show point mutations found in other serotypes. (B) G8A HAfp1-23 mutant. A ‘boomerang’ structure formed by HAfp1-20 at pH 7.4 and (D) pH 5.0. (E) G1V HAfp1-20 mutant. Structures drawn with Chimera (www.cgl.ucsf.edu/chimera) (corresponding PDB codes are given in Table 1a).
of a helical hairpin in which N- and C-terminal helices pack together (Fig. 2A). In these studies the multi-dimensional isotope-based NMR techniques were applied, with the use of \(^1\)H, \(^{13}\)C and \(^{15}\)N nuclei (Table 1a), in contrast to the previous, mainly \(^1\)H-based, experiments. The helical hairpin shows a dense network of HN–Hα and Hα–Hα hydrogen bonds, including the \(i+4\rightarrow i\), stabilizing α-helical structures, as well as several inter-helical hydrogen bonds, four of them of Hα to C=O type. The latter type is also exemplified by a double hydrogen bond between A5 and M17 (Fig. 3A).

Beside hydrogen bonds, the same authors pointed out another stabilizing factor, a charge-dipol interaction between the protonated N-terminus and a dipol of C-terminal helix arm (Lorieau et al., 2011). A strongly elevated pK of 8.8 was observed for the N-terminal glycine and for two acidic side-chains E11 and D19: 5.31 and 4.35, respectively. Using the reported pK values, a drop of pH from 7.4 to 5.0 changes the total HAfp1-23 charge from \(-1.03\) to \(-0.15\), resulting in a decreased electrostatic barrier for peptide-membrane interactions. Strikingly, there were no major structural rearrangements of the helical hairpin observed between pH 7.4 and 4.0, in contrast to the previously determined structures of HAfp1-20 (see below and Fig. 2C, D). However, at low pH the native peptide transiently changes its conformation to a totally open conformation (\(\sim 20\%\) of population), with a lower bent angle in the middle, which is similar to one seen in a G8A mutant (Fig. 2B) (Lorieau et al., 2012).

Serotypes and fusion peptide mutants

Charge variations between the fp serotypes may raise a question about its influence on the fp structure. For instance, the hairpin and ‘boomerang’ structures were originally solved for H1 and H3 serotypes, respectively, whose net charges differ, since T15 in H1 is replaced by E15 in H3 (Fig. 2A). To check the influence of such substitution, H3 HAfp1-23 structure was solved recently by liquid NMR in DPC micelles and it appeared to share a tight helical hairpin fold with H1 HAfp1-23 (Du et al., 2014). Serotype — structure relationships can be explained thanks to the bioinformatics tools, such as the recently described hydrophobic moment map pattern in relationship with the peptide structure (Worch, 2013). The structure of the influenza helical hairpin could be assigned thanks to the characteristic pattern on the map, in contrast to a distinct pattern obtained for a straight helix of HIV gp41 fusion peptide (Jaroniec et al., 2005).

It was noticed that the pattern on a hydrophobic moment map is similar for all influenza serotypes (Worch, 2013), suggesting that all of them may adopt a helical hairpin fold.

The unique structures of naturally occurring fusion peptides were a starting point of the structural studies of point mutations. The main focus was directed at the N-terminal glycine (G1) and the kink region. The structures of G1S and G1V mutants (Table 1b) were discussed in the light of V-shaped, ‘boomerang’ structure (Li et al., 2005). In principle, G1S was described as ‘a V with a disrupted ‘glycine edge’ on its N-terminal arm’ and G1V as adopting ‘a slightly tilted linear helical structure in membranes’ (Li et al., 2005). Compared to wild type, both mutations showed altered behavior during membrane fusion; G1S caused hemifusion only and G1V abolished fusion completely. It is worth noting that the kink was absent in the G1V structure (Fig. 2E), and the mutant had an increased propensity to form β-structures on the membrane surface, as observed by solid-state NMR with cholesterol-containing membranes for wild type fp (Wasniewski et al., 2004). The importance of G1 residue was also shown in biochemical studies, in which

![Figure 3. A network of hydrogen bonds in the fusion peptide structures drawn in a form of a helical wheel with the assumption of the helical hairpin structure. Polar residues are marked in gray and a dotted line depicts a hydrogen bond between Hα and C=O. (A) HAfp1-23 at pH 7.4, (B) HAfp1-20 at pH 7.4, (C) HAfp1-20 at pH 5.0.](image-url)
all N-terminal deletion mutants were shown to abolish fusion (Langley et al., 2009).

The kink region was further studied in the experiments with W14A and F9A mutants. While the F9A mutation did not change fusion properties and almost did not change thermodynamic binding parameters compared to the wild type, W14A mutant was not able to mediate fusion. Its impaired function was described as related to the structure of a flexible kink that points out of the membrane, in contrast to the more structured kink region of the wild type peptide (Lai et al., 2006). The F9A mutant was consistently shown to have a similar structure as the wild type peptide. Another mutant, G13A, was described as having a shallower kink with an angle of ~150° and perturbed the membrane to a lesser degree than wild type (with a kink angle of ~105°), but to a larger degree than non-fusogenic mutants (Lai & Tamm, 2010). Coarse-grained molecular simulations explained the mutant’s behavior by the formation of a peptide bundle formed by four or more fusion-mediating peptides, which were arranged in a transmembrane conformation, that was abolished for G14V and W14A (Ris selada et al., 2012). All these studies indicate that the degree of flexibility in the kink region is crucial for peptide rearrangements in the membrane.

Membrane-mimicking properties

Taking into consideration the fact that fusion peptides fuse lipid vesicles but fail to fuse surfactant micelles, an issue of membrane-mimicking properties of DPC micelles can be discussed. For this purpose, comparing the structures obtained by liquid- and solid-state-NMR is of great significance, since in the latter technique the membranes are prepared as bilayers of naturally occurring lipids. The structures of HAfp1-20 at pH 7.4 and 5.0 were solved in POPC/POPG and DTPC/DTPG bilayers by 13C–15N correlation spectroscopy (proton-driven spin diffusion; PDSD) (Sun & Welkly, 2009) and recently by 13C–2H NMR and double resonance (REDOR) (Ghosh et al., 2013) (Table 1b). The H3 HAfp1-20 peptide at pH 5.0 appeared similar to the open ‘boomerang’ in detergents, whereas a structure at pH 7.4 with an extended C-terminal terminus, previously solved by liquid-state NMR (Fig. 2B), was not observed. In addition, more closed conformers for HAfp1-20 fragments were observed that coexisted with an open (‘boomerang’) structures and did not appear to be exchanging on the NMR timescale. The authors ascribe the contradiction to previous incorrect G13 dihedral angles analysis of experimentally correct 13C chemical shifts, however they do not exclude a real difference between peptide-detergent and peptide-membrane interactions (Ghosh et al., 2013). Hydrophobic residues occur in the interior of the open structure, in contrast to their side face location in the closed structure and discerning between the conformers implies a different mode of membrane binding.

Conformational changes

Describing the structures of the influenza fusion peptides as having one particular shape bears certain simplifications. The NMR technique, besides giving structural insight, may also serve as a tool for studying peptide dynamics. Fusion peptides varied by length showed different internal motions, traced by N–H vector motions on the ps-ns time-scale, using conventional Lipari-Szabo model-free analysis (Lipari & Szabo, 1982). The overall correlation times, calculated from 15N and 1H NOE data sets, were 7.1 ± 0.1 ns for HAfp1-20 and 8.4 ± 0.1 ns for both HAfp1-23 and HAfp1-28 (Lorieau et al., 2013). These values corresponded well to the structures of a ‘boomerang’ and the helical hairpin, respectively. However, a closed hairpin was also observed as a transient conformer of HAfp1-20 with the population estimated as 11 ± 4% by correlation of chemical shifts and for ~10% from NOE measurements (Lorieau et al., 2013). Transitions between a kinked helix, a straight helix and a helical hairpin, were also observed recently in the molecular dynamics simulations (Larsson & Kasson, 2013).

It has to be remembered that in most of the studies, the peptides were used in a tagged form to mitigate their rather poor solubility (see Table 1). Although the additional five residues (Y24-HHQ-N28) were shown to have near random coil chemical shifts and tended not to form stable hydrogen bonds with the hairpin, nor altered the hairpin structure (Lorieau et al., 2013), the charge of +3 or +4 on the solubility tag can be repelling to the positively charged N-terminus. Thus, the presence of extra, non-native residues may alter conformer population and/or transition rates. In fact molecular simulations pointed out that the presence of charged C-terminal tag significantly alters the conformational sampling and results in more diverse conformational ensembles (Panahi & Feig, 2010). Conformations of the fusion peptide are inevitably related to the fusion mechanism and formation of the fusion pores. Nevertheless, further studies are needed to resolve all the details and explain somewhat contradictory results.

CONCLUSIONS

The high sequence conservation of the first 23 amino acids of hemagglutinin HA2 subunit is related to its structure, which, depending on the peptide length, was shown to adopt a ‘boomerang’-like shape or a helical hairpin. It is still not clear whether such structures indicate the mode of binding, since it has been proposed that a ‘boomerang’-like peptide binds to the membrane with its apex pointed to the aqueous surroundings, whereas a hairpin is oriented parallel to the membrane surface. A new light can be shed on the fusion processes by application of the model membrane systems (for a review see Marsden et al., 2011), especially those with the application of single particle analysis (Otterstrom & van Oijen, 2013). A major question for the influenza fusion peptide is related to its length: since both HAfp1-20 and HAfp1–23 lead to fusion of vesicles, what is the role of the distinct structures they adopt? It is very likely that these peptides have distinct pore-forming properties, or diverse abilities of plausible interactions with the transmembrane region. Regarding the latter issue, a six-helix bundle of fp:TMD was proposed for the parainfluenza virus (Donald et al., 2011). It is also know that the TM region of HA2 is post-translationally modified (palmitoylated) and, in addition, a groove in the TMD helix could accommodate a fatty acid (Veit, 2012). Therefore, the palmitic acid would also be a part of the bundle assembled in the membrane. However, these speculations, as many others, remain to be elucidated for a complete understanding of the hemagglutinin mediated fusion process, a key step in the influenza virus replication.

Acknowledgements of financial support

The article was written as a part of the Polpharma Scientific Foundation project, partially performed in the NanoFun laboratories at the Institute of Physics, PAS.
REFERENCES


