Pacific Symposium on Biocomputing 12:40-50(2007)

PREDICTING STRUCTURE AND DYNAMICS OF LOOSELY-ORDERED PROTEIN COMPLEXES: INFLUENZA HEMAGGLUTININ FUSION PEPTIDE

PETER M. KASSON
Medical Scientist Training Program
Stanford University, Stanford CA 94305 USA

VIJAY S. PANDE
Department of Chemistry
Stanford University, Stanford CA 94305 USA

Transient and low-affinity protein complexes pose a challenge to existing experimental methods and traditional computational techniques for structural determination. One example of such a disordered complex is that formed by trimers of influenza virus fusion peptide inserted into a host cell membrane. This fusion peptide is responsible for mediating viral infection, and spectroscopic data suggest that the peptide forms loose multimeric associations that are important for viral infectivity. We have developed an ensemble simulation technique that harnesses >1000 molecular dynamics trajectories to build a structural model for the arrangement of fusion peptide trimers. We predict a trimer structure in which the fusion peptides are packed into proximity while maintaining their monomeric structure. Our model helps to explain the effects of several mutations to the fusion peptide that destroy viral infectivity but do not measurably alter peptide monomer structure. This approach also serves as a general model for addressing the challenging problem of higher-order protein organization in cell membranes.

1. Introduction

Seasonal influenza infection is responsible for an estimated 41,000 deaths each year in the United States [2], and an avian influenza pandemic is projected to cause illness in 90 million Americans, with up to 1.9 million deaths [3]. Viral infection of host cells is mediated by a trimeric hemagglutinin protein, which inserts an approximately 20-residue N-terminal fusion peptide into the target membrane. Many mutations to the fusion peptide destroy viral infectivity, demonstrating the importance of understanding the fusogenic activity of hemagglutinin and enabling the design of novel antiviral drugs. The hemagglutinin ectodomain is a tightly structured trimer [5], which enforces a high local concentration of fusion peptide. This trimeric structure of hemagglutinin and additional fluorescence quenching experiments [6] support peptide complex formation, but infrared spectroscopy data suggest the fusion peptide does not strongly self-associate in lipid bilayers [7, 8]. In combination,

these data suggest a model of a loose complex of fusion peptide trimers where complex formation is not associated with a detectable change in peptide conformation. While the available spectroscopic data show the individual monomers to be structured, the relationship of the monomers in the complex is not well defined.

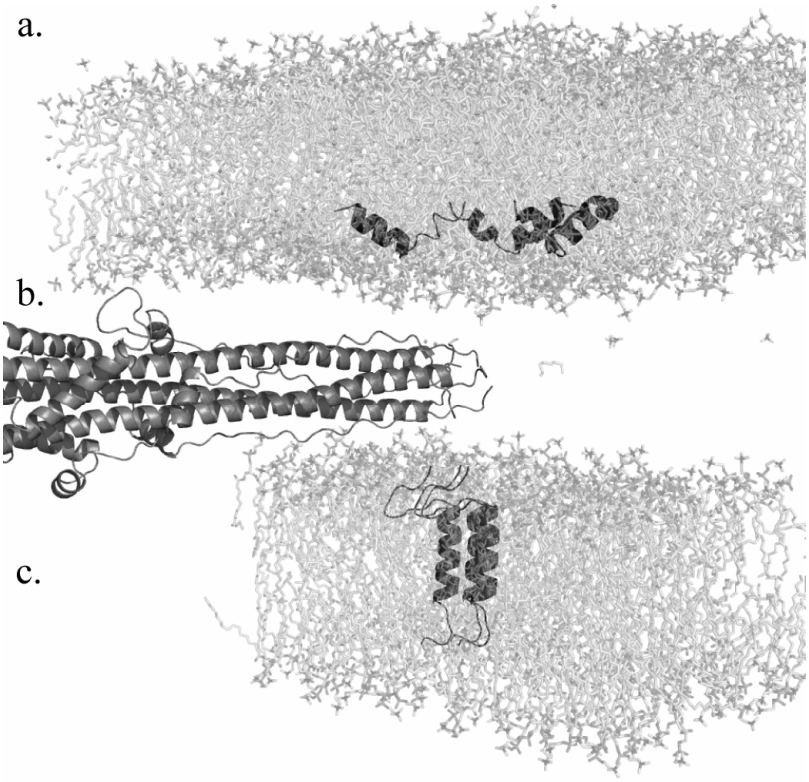

Figure 1. Cut-away view of influenza virus hemagglutinin.
Rendered in (**a**) are three copies of the influenza fusion peptide, which are inserted in the target cell membrane and connected via flexible linkers to the hemagglutinin ectodomain rendered in (**b**). This complex is anchored in the viral membrane by three copies of the C-terminal transmembrane domain rendered in (**c**). This model was constructed using crystal structures of the ectodomain [1] , NMR structures of the fusion peptide [4], and a homology model of the transmembrane domain.

This loosely-ordered nature of influenza fusion peptide complexes is representative of a large class of membrane signaling processes that are mediated by loose or transient interactions. Compared to more traditional multimeric complexes formed in solution, these complexes are much more challenging to probe using traditional methods for experimental structure

determination or computational structure prediction. Molecular dynamics simulation can yield physically-based models of protein behavior, but such approaches have typically been limited to exploring only a few molecular trajectories. Additionally, previous molecular dynamics simulations of influenza fusion peptides [9-11] have considered only monomers rather than trimers. To overcome these obstacles, we have developed ensemble molecular dynamics methodology for membrane protein simulation, a robust method that computes thousands of separate simulations to yield statistically accurate prediction of transient complexes or slow-timescale processes. To achieve calculations on this scale, we have employed worldwide distributed computing via the Folding@Home project [12]. In this report, we predict the structure and dynamic movement of influenza fusion peptide trimers via ensemble molecular dynamics simulation, thus demonstrating a powerful new approach for examining protein complexes and protein signaling in lipid membranes.

2. Methods

Molecular dynamics calculations were performed using GROMACS [13] running under the Folding@Home distributed-computing architecture [12]. Fusion peptide coordinates were taken from the NMR structure [4], and POPC bilayer coordinates were used as previously reported [14]. Peptide conformations were generated as in Fig. 2 with the central Glu11 residues placed at a radial distance of 19A. Peptides were inserted in the bilayer by deleting any lipids with phosphate head groups within 7A of the peptide backbone and then equilibrating the bilayer with the peptide conformation fixed before beginning simulation. Simulations were performed in a periodic box under NPT conditions with semi-isotropic pressure coupling and in explicit TIP3P water with 150 mM sodium chloride. The GROMOS87 united-atom force field was used with modifications for lipid parameters from Berger [15]. Each nanosecond of a single simulation trajectory requires approximately 4 days of compute time on a 2.8 GHz Pentium 4.

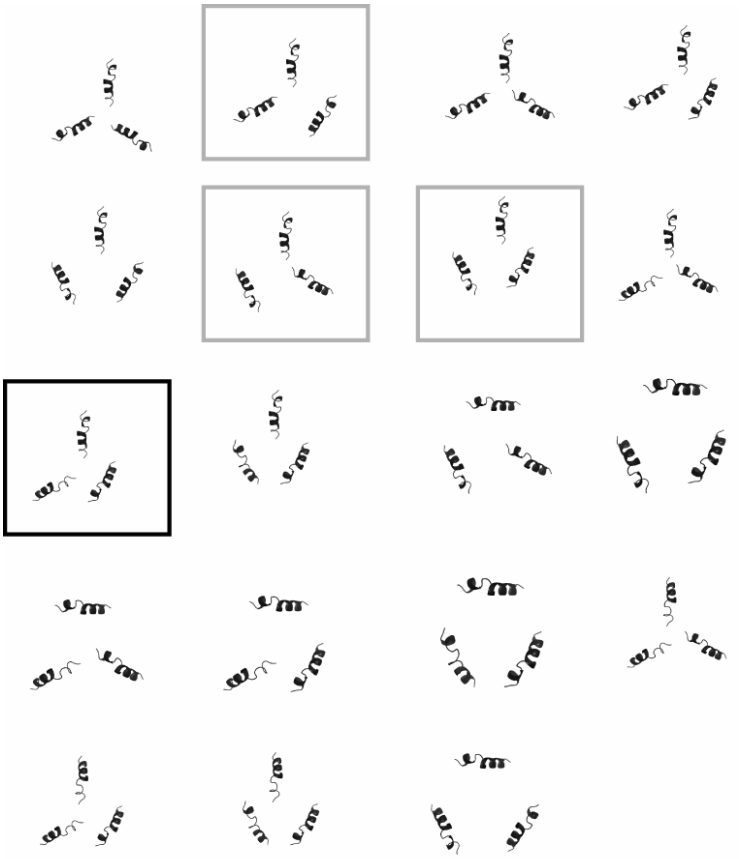

Figure 2. Starting conformations for fusion peptide trimers
The 19 starting conformations were taken by placing the peptides in a radially symmetric orientation and then generating all possible combinations of 90-degree rotations around a bilayer normal centered at Glu11 of each peptide, eliminating degenerate conformations resulting from symmetry of identical monomers. The trimers were then placed in a POPC bilayer and simulated with explicit water and ions. Boxes denote the four most stable conformations, with the most stable in black and the three following in gray. (See Fig. 4 for analysis.)

3. Results

Using ensemble molecular dynamics, we have developed a structural and dynamic model for the interaction of hemagglutinin fusion peptide trimers. We incorporate known structural information regarding the peptide monomer conformations with a broad sampling approach to create a large ensemble of simulations. This ensemble allows us to predict the conformational dynamics of the relatively disordered trimeric fusion peptide complex. We use

experimentally-measured bilayer insertion depths for the peptide [4] and the NMR structure of the peptide in micelles [4] to specify the initial conformation of each monomer. All possible rotational arrangements of the peptides within the bilayer were sampled at 90-degree increments (Fig. 2). This broad array of starting conformations allows us to test the relative stability of different trimer conformations in an efficient manner. We report the results of >1000 separate molecular dynamics simulations of up to 25 ns in length and an additional 100 simulations of fusion peptide monomers for comparison. Use of this ensemble technique allows us to predict the conformational dynamics and relative stability of fusion peptide trimeric complexes.

3.1. *Conformational change of peptide monomers*

We assess peptide conformational change by measuring root mean squared deviation (RMSD) of the protein backbone from the starting configuration at nanosecond intervals. The resulting data (Fig. 3a) show that the conformational fluctuation of peptides within the trimeric complex matches that of monomeric peptides to within error, suggesting that trimer formation does not involve a significant ordering of the internal conformation of the fusion peptide. RMSD values for the individual starting conformations are identical to within statistical error (data not shown). These findings are in good agreement with experimental infrared spectroscopic measurements [7] that show no conformational change from the monomeric NMR structure of the fusion peptide in micelles to the membrane-inserted form. Experimental measurements of short synthetic peptides containing fusion peptide sequences also suggests that the conformation of membrane-inserted peptides is primarily alpha-helical but that non-inserted beta-sheet aggregates may form at the lipid-solvent interface [7]. In our simulations, peptides remain membrane-inserted and no relevant beta sheet conformation is detected.

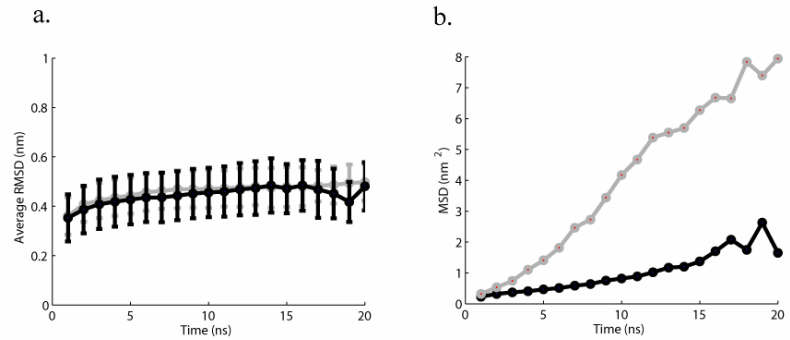

Figure 3. Conformational flexibility and translational motion of peptides. Conformational stability (**a**) is assessed for each peptide in the trimer by performing a rigid-body alignment to the starting conformation at each time point and then measuring RMSD between the backbone α-carbons. Average RMSD values are plotted here for 1230 separate molecular dynamics trajectories of the trimers (black) and overlaid with analogous values for monomeric fusion peptides (gray). Error bars represent one standard deviation of the mean. Translational motion (**b**) is assessed via mean squared deviation of the center of mass of each peptide from its starting position. Plotted values are the mean over each monomer in 1230 separate molecular dynamics trajectories of peptide trimers (black) and analogous values calculated for 100 trajectories of peptide monomers (gray).

3.2. *Translational movement*

In addition to conformational changes, peptides may undergo both translational and rotational movements. We measure translational movements via mean squared displacement (MSD) of the peptide center of mass as a function of time. Average MSD values are plotted in Fig. 3b. Since peptides in a complex undergo constrained rather than Brownian diffusion, standard methods for estimating diffusion coefficients do not apply. However, a linear fit of MSD versus time provides a good means to compare peptide movements in trimeric complexes to movements of peptide monomers. In our simulations, individual peptides in trimers move approximately 4-fold more slowly than monomeric peptides in lipid bilayers (0.1 nm^2/s versus 0.4 nm^2/s), consistent with the diffusional constraint posed by other members of the trimer.

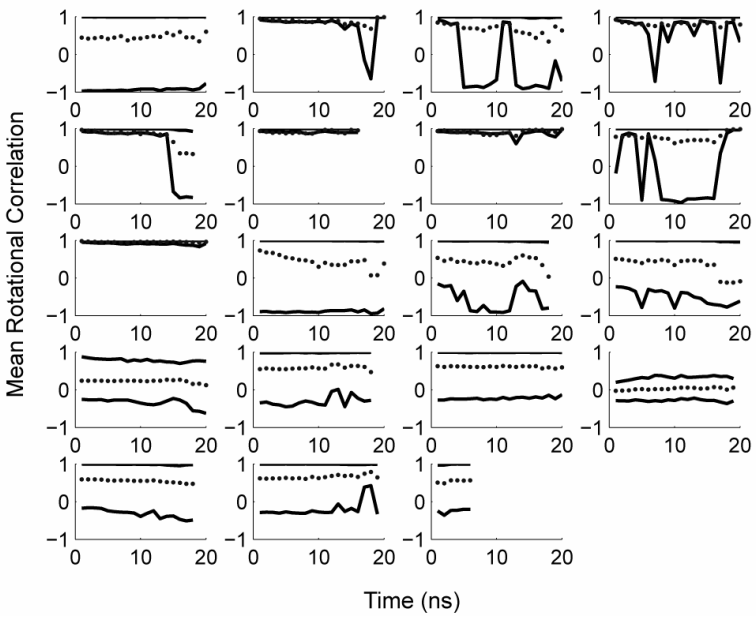

Figure 4. Rotational autocorrelation functions for fusion peptide trimers. Plotted are mean rotational autocorrelation functions for all monomers, with each subplot corresponding to the analogous trimer starting configuration in Fig. 2. Dotted lines represent mean values, while solid lines represent 80% confidence intervals. Correlation functions are calculated as $\langle \mathbf{r}(0), \mathbf{r}(t) \rangle$, where $\mathbf{r}(t)$ is the unit vector corresponding to the first principal axis of the protein molecule at time t and are averaged over each of ~70 separate simulations per starting configuration.

3.3. *Orientation of the trimeric complex*

To assess the rotational dynamics of peptide monomers within the fusion peptide complex, angular correlation functions were measured for each monomer at each starting configuration. Angular (or rotational) correlation functions measure the average "decay" of a population of peptide trimers from a single given starting configuration into other configurations. They therefore provide a good metric for relative conformational stability, and indeed the results of these measurements (Fig. 4) show significant differences in stability between starting conformations (p-values for rotational correlation functions of the four most stable configurations after 15 ns of simulation are 0.004, 0.001, 0.007, and 0.009 calculated using the Wilcoxon rank sum test with Bonferroni multiple-hypothesis correction). Notably, the most stable rotational arrangements are not radially symmetric; instead, they feature two monomers facing end-in (with a slight but not statistically significant preference for "nose-

to-tail" over "nose-to-nose") and a third aligned sideways. This protein packing arrangement is unusual in that homotrimeric soluble proteins typically form radially symmetric complexes.

We hypothesize that our predicted trimer conformation may be driven by lipid energetics: favored rotational states allow trimer packing such that the protein-lipid interface is reduced and fewer lipids are "confined" by protein helices on multiple sides. Local ordering of lipid bilayers by protein helices has been reported in the experimental literature [16], and the most favored packing arrangements we observe are likely to have increased lipid entropy. According to this model, membrane-embedded peptides exert a local ordering effect on adjacent lipid molecules. We postulate that a more compact packing arrangement that has a smaller peptide-lipid interfacial area would cause less ordering of the bilayer lipids and thus have a smaller entropic penalty.

The asymmetric packing arrangements we simulate (Fig. 5) also help to explain functional mutation data on influenza hemagglutinin. Mutants to the N-terminus of the fusion peptide G1S and G1V prevent membrane fusion in cell-based assays [17, 18] but have near-identical conformations as determined by NMR in micelles [19]. Each stable trimer conformation that we predict has at least one N-terminus packed towards the center of the complex, such that substitution of the terminal glycine for a bulkier amino acid could disrupt trimer packing. We hypothesize that such a disruption could interfere with the fusogenic effects of the peptide complex on lipid membranes and potentially inhibit fusion via that mechanism. This would explain how point mutations that do not alter monomeric peptide structure can destroy viral infectivity.

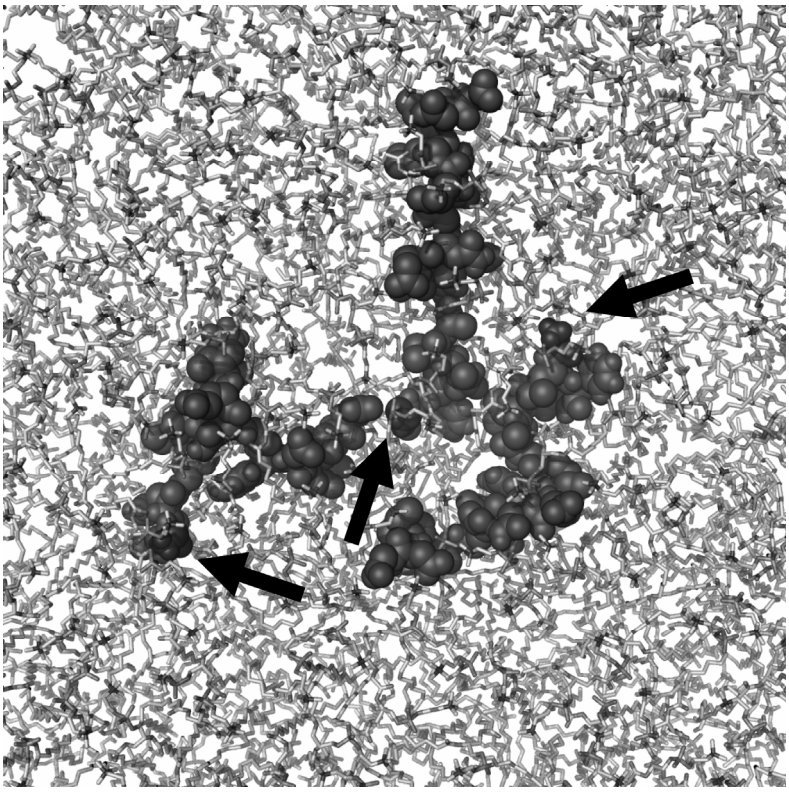

Figure 5: Representative structure of fusion peptide trimer after 20 ns simulation. The structure shown is from the most stable trimer orientation. The peptides are rendered in space-filling form, and the surrounding lipids (light gray) are rendered in stick form, showing the packing of protein and lipid. The N-termini of the peptides are marked with arrows.

4. Conclusions

Many protein interactions critical for disease processes occur within loosely ordered membrane complexes. Because of their conformational diversity and the complicating presence of the lipid bilayer, these complexes pose a challenge for experimental structure determination and computational prediction. In this report, we introduce a general means of studying loosely ordered complexes via ensemble molecular dynamics and apply it to influenza virus fusion peptides. Using distributed-computing techniques, we are able to calculate conformational stability for a broadly-sampled ensemble of candidate conformations, generating a robust structural model for protein-protein and protein-lipid interactions in fusion peptide trimers. Experimental mutants of influenza fusion peptides have

been identified that have unchanged monomeric structure but lack fusogenic capability, thus destroying viral infectivity. The structural understanding generated by our computational approach aids the interpretation of such mutational data, suggesting a means by which such mutants may cause altered protein-protein packing among fusion peptide trimers. In addition, we suggest a novel paradigm for protein complex structure, in which lipid entropy plays a important role in determining the structure of loosely ordered complexes within lipid membranes.

Acknowledgments

The authors would like to thank K. Branson for helpful discussions and Folding@Home volunteers worldwide.

References

1.Bullough, P.A., F.M. Hughson, J.J. Skehel, and D.C. Wiley, *Structure of influenza haemagglutinin at the pH of membrane fusion.* Nature, 1994. **371**(6492): p. 37-43.
2.Dushoff, J., J.B. Plotkin, C. Viboud, D.J. Earn, and L. Simonsen, *Mortality due to influenza in the United States--an annualized regression approach using multiple-cause mortality data.* Am J Epidemiol, 2006. **163**(2): p. 181-7.
3.*Pandemic Influenza Plan*, U.S.D.o.H.H. Services, Editor. 2005.
4.Han, X., J.H. Bushweller, D.S. Cafiso, and L.K. Tamm, *Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin.* Nat Struct Biol, 2001. **8**(8): p. 715-20.
5.Wilson, I.A., J.J. Skehel, and D.C. Wiley, *Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution.* Nature, 1981. **289**(5796): p. 366-73.
6.Cheng, S.F., A.B. Kantchev, and D.K. Chang, *Fluorescence evidence for a loose self-assembly of the fusion peptide of influenza virus HA2 in the lipid bilayer.* Mol Membr Biol, 2003. **20**(4): p. 345-51.
7.Han, X. and L.K. Tamm, *pH-dependent self-association of influenza hemagglutinin fusion peptides in lipid bilayers.* J Mol Biol, 2000. **304**(5): p. 953-65.
8.Haque, M.E., V. Koppaka, P.H. Axelsen, and B.R. Lentz, *Properties and structures of the influenza and HIV fusion peptides on lipid membranes: implications for a role in fusion.* Biophys J, 2005. **89**(5): p. 3183-94.
9.Vaccaro, L., K.J. Cross, J. Kleinjung, S.K. Straus, D.J. Thomas, S.A. Wharton, J.J. Skehel, and F. Fraternali, *Plasticity of influenza haemagglutinin fusion peptides and their interaction with lipid bilayers.* Biophys J, 2005. **88**(1): p. 25-36.

10.Huang, Q., C.L. Chen, and A. Herrmann, *Bilayer conformation of fusion peptide of influenza virus hemagglutinin: a molecular dynamics simulation study.* Biophys J, 2004. **87**(1): p. 14-22.
11.Lague, P., B. Roux, and R.W. Pastor, *Molecular dynamics simulations of the influenza hemagglutinin fusion peptide in micelles and bilayers: conformational analysis of peptide and lipids.* J Mol Biol, 2005. **354**(5): p. 1129-41.
12.Shirts, M. and V.S. Pande, *Computing - Screen savers of the world unite!* Science, 2000. **290**(5498): p. 1903-1904.
13.Van der Spoel, D., E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, and H.J.C. Berendsen, *GROMACS: Fast, flexible, and free.* Journal of Computational Chemistry, 2005. **26**(16): p. 1701-1718.
14.Tieleman, D.P., L.R. Forrest, M.S. Sansom, and H.J. Berendsen, *Lipid properties and the orientation of aromatic residues in OmpF, influenza M2, and alamethicin systems: molecular dynamics simulations.* Biochemistry, 1998. **37**(50): p. 17554-61.
15.Berger, O., O. Edholm, and F. Jahnig, *Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature.* Biophys J, 1997. **72**(5): p. 2002-13.
16.Killian, J.A., *Hydrophobic mismatch between proteins and lipids in membranes.* Biochim Biophys Acta, 1998. **1376**(3): p. 401-15.
17.Gray, C., S.A. Tatulian, S.A. Wharton, and L.K. Tamm, *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**(5): p. 2275-86.
18.Qiao, H., R.T. Armstrong, G.B. Melikyan, F.S. Cohen, and J.M. White, *A specific point mutant at position 1 of the influenza hemagglutinin fusion peptide displays a hemifusion phenotype.* Mol Biol Cell, 1999. **10**(8): p. 2759-69.
19.Li, Y., X. Han, A.L. Lai, J.H. Bushweller, D.S. Cafiso, and L.K. Tamm, *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**(18): p. 12065-76.