Predicting Functional Interactions in the Influenza Hemagglutinin Transmembrane Domain via Simulation

---

A Thesis

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

---

in partial fulfillment

of the requirements for the degree

Master of Science

by

Matthew J. Eckler

May

2015

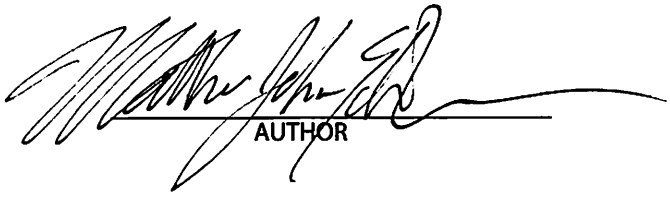
APPROVAL SHEET

The thesis

is submitted in partial fulfillment of the requirements

for the degree of

Master of Science



AUTHOR

The thesis has been read and approved by the examining committee:

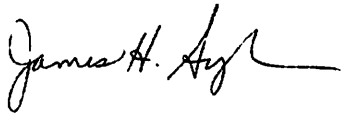
Peter M. Kasson, M.D. Ph.D.

Advisor

Jeffrey J. Saucerman, Ph.D.

Michael R. Shirts, Ph.D.

Accepted for the School of Engineering and Applied Science:



Dean, School of Engineering and Applied Science

May

2015

Abstract

The precise mechanism of cell entry by influenza remains poorly understood, despite many years of research. Entry into the cytoplasm is preceded by a membrane fusion event between the virion and host endosomal membrane, mediated by the fusion protein hemagglutinin. Hemagglutinin resides in the viral membrane with one transmembrane helix. Its ectodomain is trimeric, and this enforces close proximity between the three helices in the viral membrane. Previous research has demonstrated that the transmembrane domain of hemagglutinin is crucial to its pathogenicity, and that while some mutations are allowed, productive viral fusion to the target membrane is blocked by drastic deletions or replacement with a GPI anchor. We have conducted molecular dynamics simulations of the hemagglutinin transmembrane domain to understand any interactions that may occur between the domains and postulate how they can impact hemagglutinin function.

Our simulations indicate that the helices can associate in a membrane without the ectodomain present. We have used a multi-scale simulation approach to examine the stability of encounter complexes in order to identify key interactions and predict changes that would disrupt them. First, coarse-grained simulations were used to generate a large population of encounter complexes between helices, which were then clustered to create a diverse sample set. The resulting complexes were simulated at atomic resolution to test their stability and identify specific interactions in the membrane.

At atomic resolution, simulations show that transmembrane domain complexes remain stable in the membrane environment and exhibit key hydrophobic interactions between residues in the membrane-inserted region of the peptide. In order to test their

importance, we created two additional sets of all-atom simulations. The F205A mutant simulations were designed to ablate the most probable single contact between monomers, while the L198A V201A F205A W208A mutant simulations aimed to more broadly disrupt common inter-peptide interaction modes. Surprisingly, our simulations show that both sets of mutants exhibit similar stability in the membrane environment despite severely diminished interaction at the mutation sites. In contrast, simulations of a previously studied deletion mutant show a substantial disruption in trimer stability and membrane disruption. In order to further quantify trimeric stability in the membrane, a series of pulling simulations aims to determine the relative free energy of dissociation for the wild type and quadruple mutant trimeric complexes.

Acknowledgements

This work would not be possible without the help and support of my family, colleagues, and friends. First, I would like to thank my parents, Paul and Janet, as well as my brother Dan for all the motivation and support they gave me. As I redefined my career path during my time here, I was never lacking in encouragement.

I would like to especially thank my advisor and P.I., Dr. Peter Kasson, for all his help in my project and thesis. His guidance was paramount in the completion of my masters degree. I would also like to thank my committee members, Dr. Jeff Saucerman and Dr. Michael Shirts, for their contributions. I was lucky to have excellent input on my thesis.

Special thanks go to the current and former Kasson lab members, Dr. Per Larsson, Jeff Xing, Małgorzata Latałło, Katarzyna Zawada, Becky Dunning, Jennifer Hays, George Cortina, Dr. Rich Salaway, Dr. Bob Rawle, and Dr. Marta Domańska, for providing an excellent research environment during my time here. Similarly, thanks to all the students, scientists, and professors that made Fontaine such a great place to work and learn. Thanks also to my professors and the BME department.

Lastly, I would like to thank the other colleagues and friends I have made here. Without my peer network, my time here would not have been nearly as enjoyable.

Table of Contents

List of Figures

Figure 6. Coarse grained (CG) beads are transformed into groups of atoms during the resolution switch to all-atom. (a) Individual coarse grain peptides (left) are made of 96 beads. There are 671 atoms in each peptide, whose placement is determined by the "backward" tool described in Methods. An all-atom TMD peptide retains the structure from its CG structure (middle), shown in bonds representation (middle) and as a cartoon (right). (b) Each POPC molecule in the CG bilayer (left) is transformed from 13 beads into 134 atoms in place, yielding the bilayer on the right.

Figure 7. The number of contacts between peptides is stable after an initial decrease. The number of contacts in each snapshot is plotted over time for each simulation set. After an initial decrease, all simulations ((a) wild type, (b) F205A mutant, and (c) Quadruple mutant) relax at a new number of contacts. Averages are shown in red, with surrounding blue 95% confidence intervals. Each trace is plotted over the entire length of simulations.

Figure 8. Trimeric complexes remain tightly associated over the course of all wild type and alanine mutant simulations. The center of mass distance between each peptide is averaged over both peptides and simulations for the (a) wild type simulations, (b) single mutant F205A simulations, and (c) quadruple mutant L198A V201A F205A W208A simulations. Averages in red are surrounded by a blue 95% confidence interval. The original CG starting distance before complexation is shown in the dotted green line for comparison. All simulations display stability for their duration.

Figure 9. The probability of an inter-peptide contact highlights several frequent modes of contact. The probability of each contact is mapped over residues for wild type simulations after removing data accrued in the first 25 ns of each simulation. Areas in dark blue have no contacts in simulations, while areas in yellow have a high likelihood of contact. The probability follows a diagonal pattern due to the peptides residing in the membrane with similar depths. Several single contacts display probabilities of over 0.5, shown in bright yellow.

Figure 10. A discrete set of prevalent contacts defines the most favorable specific interaction modes between peptides. The most likely contact pairs are listed for each simulation set, along with their corresponding probabilities. Probabilities are accompanied by bounds defining a 95% confidence interval. Below, each top contact is displayed in a trimer, colored by its ranking as indicated in column 1 of the table.

Figure 11. A large decrease in probability results at the mutation site. The difference in probability for all residue pairs between the single mutant F205A simulations and the wild type simulations is shown, with data collected during the first 25 ns of each simulation removed. Areas of similar probability are colored tan. Slight increases in probability are in yellow and slight decreases in green, while a large decrease is in blue.

Figure 12. Mutation sites constitute areas of extended decreases in probability. The difference in probability for all residue pairs between the quadruple mutant L198A V201A F205A W208A simulations and the wild type simulations is shown, with data

collected during the first 25 ns of each simulation removed. Areas of similar probability are colored tan. Slight increases in probability are in yellow and slight decreases in green, while a large decrease is in blue.

Figure 13. Stable behavior is exhibited in protonated del12 mutants, but not unprotonated del12 mutant simulations. (a) The inter-peptide center of mass distance is shown with mean values in red surrounded by a 95% confidence interval in blue. (b) The number of contacts over time is shown in red surrounded by a 95% confidence interval in blue. (c) The difference in contact probability of residues 177-199 from their probability in the wild type simulations. Areas of similar probability are colored tan. Slight increases in probability are in yellow and slight decreases in green, while a large decrease is in blue.

Figure 14. Areas of statistically significant changes in contact probability follow mutation sites most directly but show scattered additional differences. The results of a paired t-test on each residue-residue contact dataset were mapped, where a blue box indicates a statistically significant increase in the mutant probability and a red a decrease in the mutant probability. (a) displays results from changes between the wild type and single mutant F205A simulations and (b) for the wild type and quadruple mutant.

List of Abbreviations

HA	hemagglutinin
HA1	polypeptide subunit 1 of hemagglutinin
HA2	polypeptide subunit 2 of hemagglutinin
TMD	transmembrane domain
GPI	Glycophosphatidylinositol
GPI-HA	Glycophosphatidylinositol/hemagglutinin construct
PIV5	Parainfluenza virus 5
POPC	lipid type; 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
X-31	X-31 strain of influenza A
H3	H3 subtype of hemagglutinin
K	Kelvin
NaCl	sodium chloride
vdW	van der Walls
Cl	chlorine
μs	microsecond
nm	nanometer
ns	nanosecond
F205A	mutation replacing phenylalanine at position 205 with alanine
cm/sec	centimeters per second
kJ mol^{-1} nm^{-2}	kilojoules*mol^{-1}*$nanometers^{-2}$
L198A	mutation replacing leucine at position 198 with alanine
V201A	mutation replacing valine at position 201 with alanine
W208A	mutation replacing tryptophan at position 208 with alanine
CG	coarse grain
PCA	principal components analysis
2D	two dimensional
Quad	quadruple (mutant)
Phe	phenylalanine
Leu	leucine
Trp	tryptophan
Val	valine
Ser	serine
Ala	alanine
Del12	mutant containing residues 177 to 199

1. Introduction

Hemagglutinin (HA) is a viral spike glycoprotein crucial to influenza virus entry into the host cell. Hemagglutinin exists as a homotrimer on the viral surface and is comprised of the two linked polypeptides HA1 and HA2. While HA1 plays a role in the viral binding to a host cell receptor, the HA2 polypeptide both anchors the protein in the viral membrane and mediates a later membrane fusion event. Before entering a host cell, the HA1 polypeptides are organized to interact with the environment outside of the virion while HA2 polypeptides are less accessible [1]. A single helical [2] transmembrane domain near the C terminal end of HA2 acts as the viral anchor in each monomer, but is not included in the prefusion structure. The N terminal region of HA2, termed the fusion peptide, remains sequestered as the virus enters the cell. After HA1 binds a receptor on the host cell surface containing a specific sialic acid, it becomes enveloped into an endosome within the cell. The pH in the endosome drops during maturation, triggering a conformational change in the viral hemagglutinin [3-8]. This effectively repurposes the hemagglutinin in order to allow the virus to release its genetic mutation into the host cell cytosol.

Hemagglutinin undergoes a major conformational change, releasing its fusion peptide in order to catalyze membrane fusion. The pH drop during endosomal maturation induces a reorganization that reveals the previously sequestered fusion peptide [3,6-8]. The fusion peptide inserts into the target membrane stably [3,9]. The conformations and effects of fusion peptide insertion in a membrane environment have been studied extensively but a precise structural role remains elusive [10-17]. The fusion peptide includes several highly conserved positions, and many single mutations

will abolish or arrest fusion at the hemifusion intermediate [18-20]. Hemifusion is a metastable intermediate of the membrane fusion pathway [21,22] characterized by a continuity in only the proximal leaflets of two opposed membranes. The fusion peptide has been shown necessary for both achieving hemifusion and continuing to full fusion [18-20]. A productive fusion pore will exhibit lipid mixing between both leaflets and contents mixing.

The HA2 transmembrane domain (TMD) is necessary for productive membrane fusion, but only at the later step leading out of hemifusion. The ten-residue cytoplasmic tail region of HA2 that is located inside the viral compartment and connects to the TMD has been shown unnecessary for fusion [24]. Chimeric constructs of HA that have anchored the protein to the viral membrane with different transmembrane domains have proven the necessity of the TMD [23-28,56]. A lipid linked construct, achieved with a GPI anchor, will arrest fusion at the hemifusion intermediate [25,56]. Additional work revealed that transient non-productive fusion pores could be formed in this construct [26]. A chimera replacing the TMD with the corresponding domain of another viral fusion protein from Sendai virus did not inhibit fusion [28]. The TMD remains largely robust to sequence mutation, but more perturbative effects were discovered through deletion mutants. Successively larger deletions from the TMD show that the receiving 12 residues from the TMD will arrest fusion at the hemifusion intermediate, while larger deletions will not allow the protein to express in the model at all [23]. The inhibitory 12 residue deletion mutant may lose its transmembrane orientation, and instead reside in only a single leaflet. In agreement with the GPI-HA construct, these studies reveal a

possible length requirement for the TMD, defined by the number of residues needed to span both leaflets.

Related viral fusion proteins from Hendra, parainfluenza virus (PIV5), and human metapneumovirus have been shown to have trimeric transmembrane domains in the absence of their ectodomains [29]. While this interaction may be important to fusion protein stability [30], a trimeric ectodomain would almost certainly have importance during the formation and expansion of a fusion pore. Instead of a length requirement [23], the TMD may require some specific interaction or functional structural mode in the bilayer, one that is disturbed by severe deletions and does not occur between GPI anchors. Viral fusion proteins often share similar characteristics and are thought to take advantage of a similar pathway in order to merge opposed bilayers [38]. Recent evidence suggests that the TMD and fusion peptide of PIV5 may interact in the postfusion state and it has been suggested that this could potentially help drive fusion pore formation [31]. Additionally, a PIV5 transmembrane domain mutant scanning study identified two residues crucial to fusion [32]. Together, these results provide a motivation for further study of the influenza TMD structure and interactions in the membrane.

While much effort has been put into characterizing transmembrane domain association with liquid-ordered domains [33,34], some recent work involving the influenza hemagglutinin transmembrane domain has provided evidence of a possible transmembrane interaction. When digested with bromelain and other related enzymes, the cleavage site has been shown to be near the TMD [68]. The different cleavage sites of certain strains of influenza A HA may result from a tight assembly of the transmembrane domains in the membrane environment that promote cleavage in the

TMD [35]. More directly, a fluorescence and gel electrophoresis study by Chang et al. has suggested evidence of a tight transmembrane domain association and a weak interaction between transmembrane and fusion peptides in a model membrane environment [36]. A molecular dynamics simulation study confirmed a perpendicular helical arrangement of TMD in the membrane but failed to suggest inter-peptide interactions in the transmembrane region [37].

Despite previous knowledge dating back over 20 years that the TMD is required for a productive membrane fusion event, its mechanistic role remains unresolved. Structures of HA monomers and trimers lack information on the TMD, but other experimental efforts have characterized it as a predominately alpha-helical domain that may oligomerize in the membrane environment. While the role of the fusion peptide is easily disturbed with mutation, the TMD does not display strict sequence requirements but seems to follow a general length constraint. The present work attempts to characterize TMD interactions in the bilayer environment in order to connect its structure in the bilayer to a mechanistic role during fusion. We utilized a multi-scale molecular dynamics approach to first sample the diffusion process and generate possible encounter complexes and then characterize inter-peptide interactions at the residue level. Coarse grain molecular dynamics simulations of TMD monomers in a bilayer without the HA ectodomain provided a set of trimeric encounter complexes formed from unique initial arrangements of monomers. All-atom simulations on a representative selection of complexes were used to determine interaction tendencies between TMD peptides that contribute to their trimeric state. Our simulations show that the TMD forms stable hydrophobic interactions along a discrete set of residues. Although individual

interactions can be abrogated via mutation, the overall stability of trimeric complexes remains unaffected.

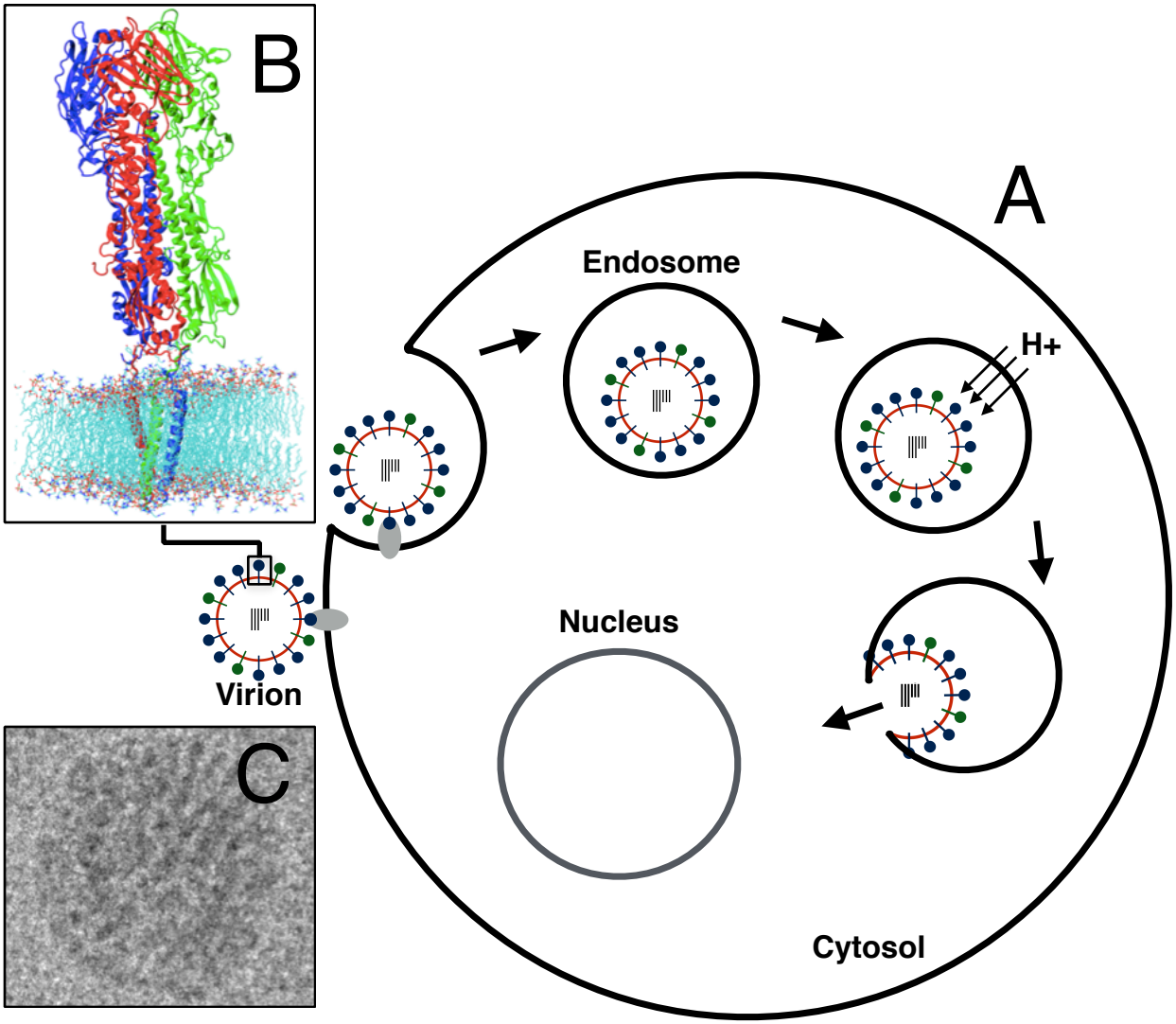


Figure 1. Influenza enters the host cell via endocytosis. (a) The virion enters the cell in an endosome after its spike protein hemagglutinin HA binds a sialic acid-terminated receptor, and upon the pH drop during endosomal maturation the virion fuses with the endosomal membrane to release its contents into the cytoplasm for infection. Hemagglutinin catalyzes the membrane fusion event. (b) Trimeric HA is anchored to the membrane via a transmembrane domain (TMD). The ectodomain crystal structure (pdb code 1HGE [67]) is colored by monomer and attached to our model alpha helical TMD embedded in a POPC membrane. (c) A cryo-EM image of a single X-31 influenza virion (image courtesy PMK).

2. Methods

2.1 Coarse Grain Simulations

A structure of the (H3) X-31 strain was constructed from residue 177 to 216 to initially model an ideal helix. A 256-POPC bilayer was equilibrated from a starting structure made with the CHARMM-GUI online application [39]. TMD peptides were inserted into the equilibrated bilayer in a triangular arrangement, with spacing between each peptide. The insertion procedure, adapted from a previous fusion peptide simulation study [57], first removed lipids with phosphorous atoms located within 5 Å of the intended peptide, and then water molecules within 2.5 Å of the protein. The bilayer and protein system was equilibrated for more than 100 nanoseconds. A post-equilibration snapshot was then used to create a coarse grained system.

Coarse grain simulations have been used to study protein assemblies in the membrane environment [59]. Previous approaches have sampled longer timescale phenomena while sampling conformational relaxation at higher resolution by interconverting between coarse grain and all-atom states [60-61]. Lipids were transformed from CHARMM36 forcefield [40] parameters to MARTINI forcefield [41] parameters and beads with the use of the backward.py script [42], but in the ‘forward’ direction. The protein coordinates were transformed with the martinize.py method [43], enforcing helical structure with the application of backbone constraints designed to hold the helix together. Rotational systems were created at 90 degree increments for every unique assembly of TMD (Figure 2C). The systems were solvated with MARTINI water beads [44], and a NaCl concentration of ~150 mM was applied by replacing water

beads with ion/water MARTINI beads. Each simulation box measured 9.25 by 9.25 by 11.5 nanometers.

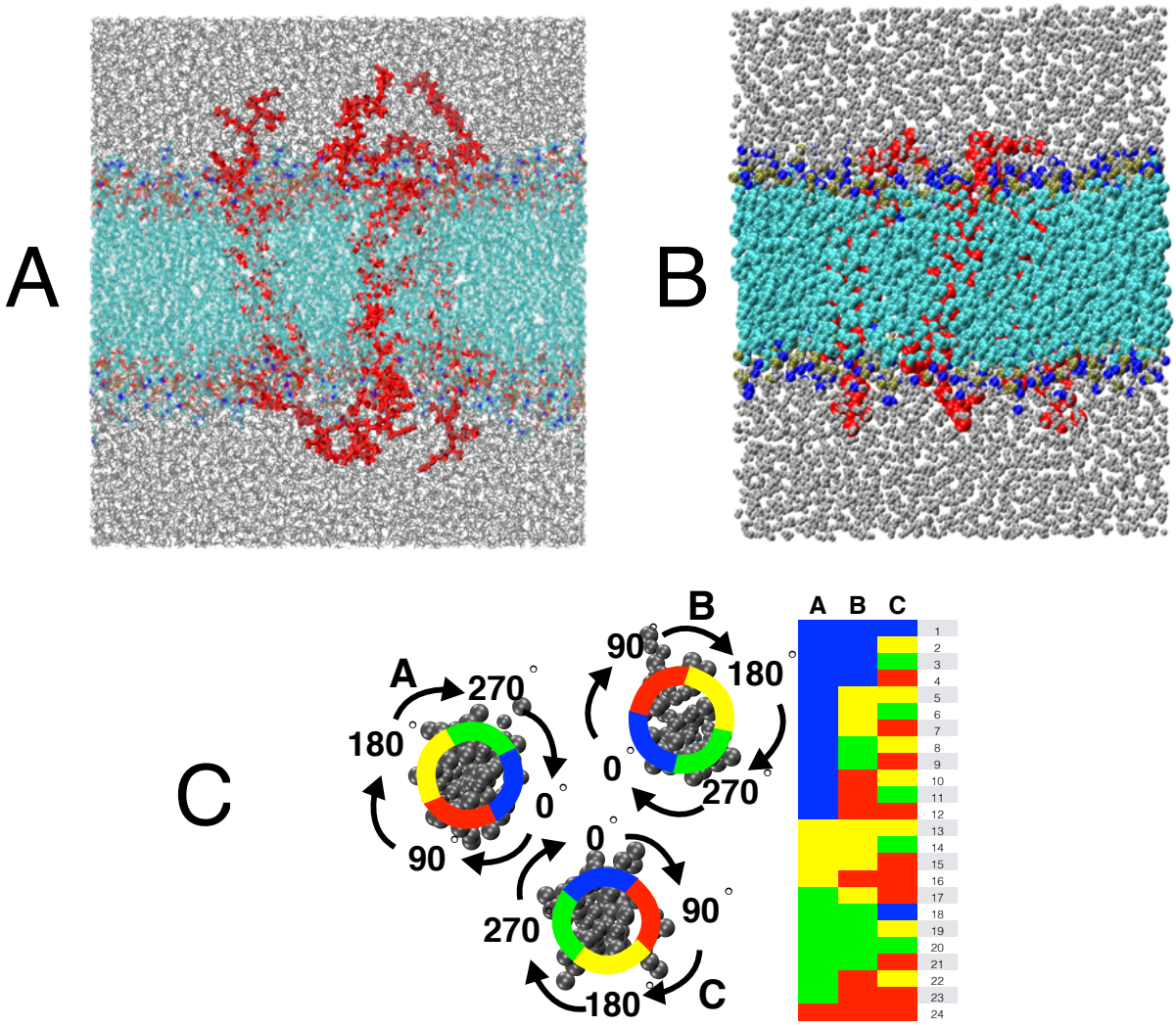


Figure 2. Coarse grain stimulations are created from an equilibrated all-atom system of 3 TMD peptides embedded in a POPC bilayer in order to sample possible peptide encounter complexes. (a) A side view of an all-atom POPC bilayer (colored by atom type) with embedded TMD peptides in red and surrounding water in silver. (b) A side view of the corresponding coarse grained bilayer with TMD peptides in red, POPC in cyan, blue, and gold, and water in silver. (c) 24 simulations are created where each individual monomer is rotated in the membrane plane at 90 degree intervals to sample a diverse set of encounter complexes. Looking down on the membrane (not shown), each gray monomer can be rotated to 0 (blue), 90 (red), 180 (yellow) or 270 (green) degrees. A table of all 24 unique arrangements is shown to the right.

Coarse grain simulations were run for more than 10 microseconds each in Gromacs 4.6.6 [47], totaling over 240 μs aggregate production simulation data. As is standard for the MARTINI coarse grain forcefield, a large time step of 30 femtoseconds was used. Coordinates and velocities were output every 250,000 steps. Production simulations were held at constant temperature and pressure with the use of the velocity rescaling thermostat [45] at 320 K and Parrinello-Rahman pressure coupling [46] at 1 bar, respectively. Electrostatics and vdW potentials were applied over the range of 0 to 0.9 nm and decayed smoothly to zero at 1.2 nm.

2.2 All-Atom Simulations

Coarse grain snapshots selected via clustering (Methods 2.3) were transformed to all-atom structures via the backward.py method [42], with the CHARMM36 forcefield. The bilayer and each peptide had to be transformed separately to properly account for peptide termini settings, and once in all-atom representation the coordinate files were merged. The box size was set to 9.25 by 9.25 by 11.5 nanometers, and the system was solvated with TIP3P [48]. A 150mM NaCl concentration was achieved through water replacement, with additional Cl beads added to neutralize the system.

Production wild type simulations were run for over 190 nanoseconds each, for an aggregate of over 9.5 μs of all-atom production simulation. To confirm results, a second group of simulations was started with the same initial coordinates but randomized velocities. A time step of 3 femtoseconds was used, with all covalent bonds constrained by the LINCS algorithm [49]. Coordinates and velocities were output every .75 nanoseconds (250,000 steps). Simulations were held at constant temperature [45] and

pressure [46] at 300 K. Long range electrostatics were calculated using the Particle Mesh Ewald method [50]. A Lennard-Jones potential was applied between 0 and 1.1 nm and decayed smoothly to zero at 1.2 nm. These settings were repeated for all other all-atom simulation sets.

Mutation snapshots were created from starting configurations used in the wild type simulations by removing and renaming atoms on the specified residues. New topologies were constructed with the Gromacs tool pdb2gmx. Single mutant F205A simulations were run for over 130 nanoseconds each, after initial equilibrations of 50 picoseconds each. The quadruple mutants were run for more than 140 nanoseconds each, totaling over 7 μs of simulation, after the same equilibration steps. Deletion mutants were created from ten random starting trimer complex configurations by cutting 17 residues from the C terminus, and updating the protein topology with pdb2gmx. Three arbitrarily selected starting conformations were left unprotonated at the C terminus, while the rest were protonated to test the effect of each protonation state. Each trimeric complex was inserted into the top leaflet of a previously equilibrated bilayer. After an initial equilibration of 150 picoseconds, production simulations were run for 80 nanoseconds each, totaling 0.8 μs of simulation data.

2.3 Analysis

The gromacs utility g_dist was used in all simulations to measure the center of mass distance between each pair of peptides. Contact maps were made with the tool g_mdmat, set to a cutoff of 0.7 nanometers. Output contact map data was transformed into numerical arrays via a script written by T Wassenaar [58]. All data was analyzed

and plotted in MATLAB [53]. Confidence intervals were calculated from the mean and standard deviation of N parallel simulations on different starting conditions. Paired t-tests were carried out for each contact for both F205A mutant and quadruple mutant simulations in order to determine which changes in probability from the wild type case displayed statistical significance, using a rejection p-value of 0.05.

We used contact maps and K-means clustering to describe and sample our TMD trimeric states. Using a distance metric defined as the sum-squared difference between two contact maps, we implemented a transparent K-means algorithm that allowed us to monitor center locations and groupings each iteration. In MATLAB, over 31,081 unique snapshots were clustered to 200 centers with our K-means implementation. We expect the 200 centers to overpopulate any natural groupings in the data, removing sensitivity due to initial random center assignments. The snapshots closest to the centers, the medoid snapshots, were sampled randomly to form a set of 50 final snapshot targets for an increase in resolution. To visualize our clustering, we plotted the projection of these targets, the medoids, and all snapshots on the first two principal components of the contact map data.

Visualization and rendering were performed using VMD [52].

2.4 Pulling Simulations

Free energy perturbation methods have been used to calculate protein binding free energies from simulation [66]. A snapshot from all-atom simulations in the wild type and quadruple mutant set was taken after 100 nanoseconds to perform a non-equilibrium pulling simulation. Each monomer was assigned a force vector pointing from

the center of the mass of all protein atoms to its own center of mass. Only x and y coordinates were used, as the pulling was designed to be parallel to the membrane surface. The alpha carbons of residues 185 to 210 were defined as pulling groups for each monomer. A harmonic potential was applied at each time step between each of the 78 pulling groups and a moving reference point. The reference moved along the defined pulling vector at a constant velocity of 5 cm/sec, starting from the peptide's center of mass. Forces on each atom were calculated separately using a force constant of 1000 kJ mol^{-1} nm^{-2}. Simulations were continued until either the distance between centers of mass of all protein and individual monomer or the distance between the reference point in space and monomer exceeded half the box size in the x direction. Simulations were run between 20 and 45 nanoseconds, and all contacts between residues inserted in the membrane were separated. Data is currently being analyzed to obtain a work measurement from each simulation, which will provide sufficient information to calculate the free energy of dissociation for both the wild type and quadruple mutant peptides [64,65].

3. Results

We have performed a multi-scale molecular dynamics study showing that influenza hemagglutinin transmembrane domain (TMD) peptides can form stable trimers in the membrane environment in the absence of their ectodomains. By placing sets of 3 TMD monomers in small bilayers in a diverse array of starting orientations, we have studied the tendency of monomers to form higher-ordered, stable assemblies in a POPC membrane on the microsecond timescale. Sampling of this process in coarse

grain representation created an ensemble of trimeric TMD complexes which we explored at all-atom resolution. Trimeric complexes of TMD remained stable for 190 nanoseconds of all-atom simulation, and we described their interactions via contact maps. We mutated a set of common hydrophobic contacts in the membrane-inserted section of the TMD with two mutant simulation sets: F205A and quadruple mutant L198A/V201A/F205A/W208A. Though interactions made by mutated residues decreased, the overall stability of the trimeric complex remained unaffected during simulations. This result is consistent with previous experimental evidence that the transmembrane domain of influenza hemagglutinin is robust to mutation.

3.1 Coarse Grain Molecular Dynamics Simulations

We performed coarse grain molecular dynamics simulations on 24 systems comprising of three influenza hemagglutinin transmembrane domain monomers embedded in a POPC bilayer. Each system was compromised of a unique triplet of TMD starting orientations, sampled at 90 degree intervals (Figure 2C). Sampling multiple starting orientations removed any bias conferred by the initial arrangement of monomers by allowing a diverse array of encounter scenarios in the membrane environment. Peptides were initially placed so that their centers of mass were roughly 5 nanometers apart, far enough that none were initially contacting but close enough that minimal diffusion would be necessary for an encounter. Each system was simulated for over 10 μs, totaling 240 μs of aggregate simulation.

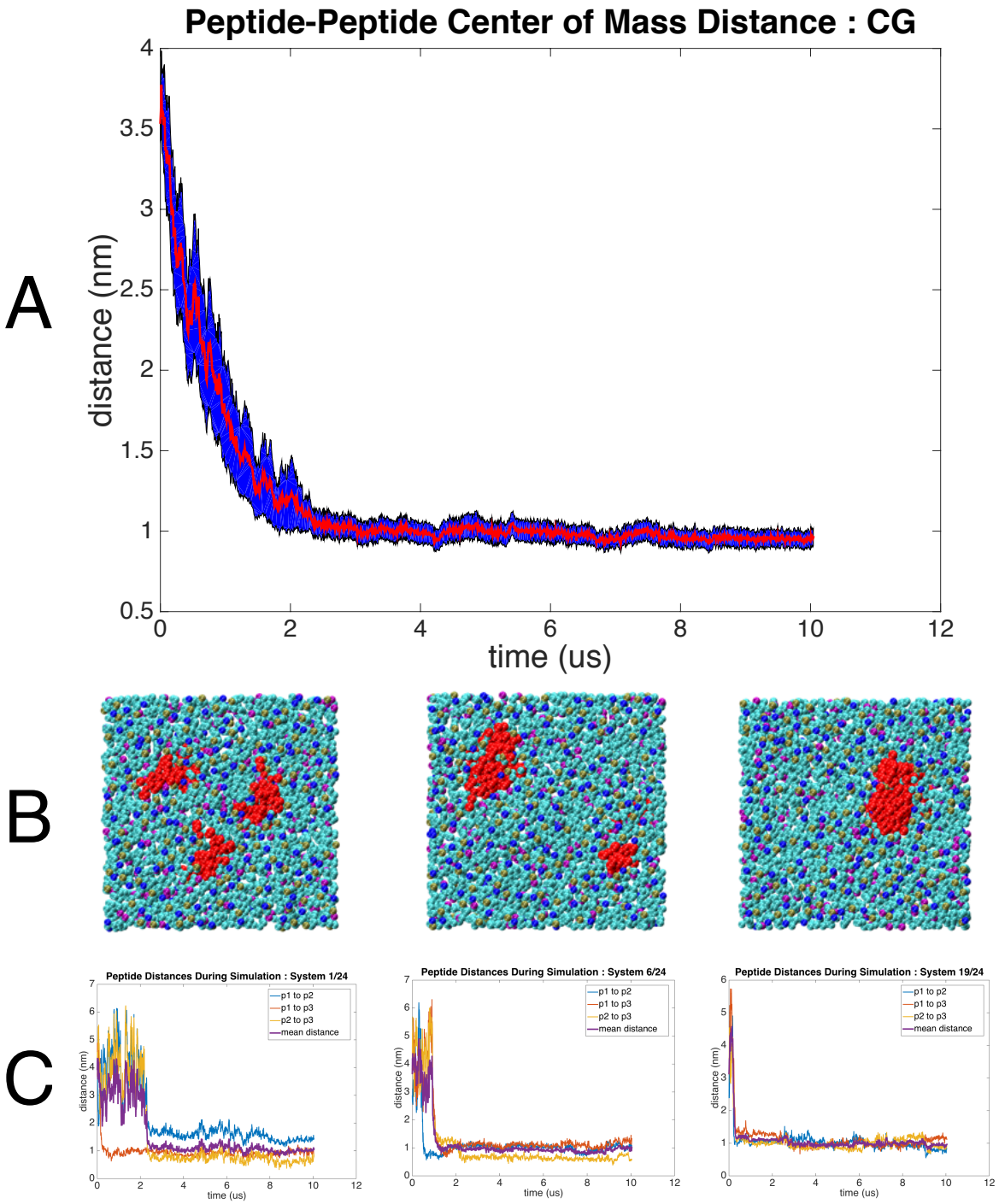


Figure 3. TMD monomers assemble into trimers on the microsecond timescale. (a) Center of mass distance between monomers is averaged over peptides and simulations and plotted as a function of time. Mean values are plotted in red and a 95% confidence interval in blue. (b) A frequent pathway of discrete events leading to trimerization starts with 2 of the 3 monomers (left) joining to form a dimer (middle), which then combines with the final monomer in a trimer (right). (c) Sample time traces of individual simulation center of mass distances display heterogeneity in timescale and pathway. Red, yellow and blue lines are individual monomer pairs while the purple shows the average of the three. The left panel details a quick dimerization and slow trimer formation, the middle has a lag between each phase, and the right shows an immediate trimerization.

All coarse grain simulations resulted in the formation of stable transmembrane domain trimeric complexes. We measured the average center of mass distance between monomers in each simulation snapshot (Figure 3). Sample traces of individual systems displayed heterogenous behavior during the first several microseconds (Figure 3C), but each trace leveled off at a close peptide to peptide distance of around 1 nanometer. At this distance, the monomers had joined into a trimeric complex. While some simulations trimerized in less than half a μs (Figure 3C, far right), after 4 μs all simulations resulted in a compact trimer. In many simulations, a dimer formed prior to the trimer as two monomers diffused together and stuck. This can be seen most clearly in the left panel of Figure 3C, where the dimer formed near the beginning of simulation time but the trimer after 2 microseconds. Snapshots from this pathway are shown in Figure 3B. Some systems will undergo additional trimeric reassembly in a compact state, as can be seen when the individual traces of Figure 3C (far right) cross; one set of peptides moved closer together. Despite reassemblies, all trimer complexes remained together throughout the simulation after formation.

3.2 All-Atom Molecular Dynamics Simulations

In order to further study key interactions and features of transmembrane domain trimers, we returned to all-atom resolution. All coarse grain snapshots were measured by the contacts made between each pair of peptides. A residue-residue contact is defined as having a bead in one residue within 0.7 nanometers of another residue. Each pair of peptides formed one binary array describing the contacts between them (Figure 4). The three arrays in each snapshot of simulation were combined into a single array

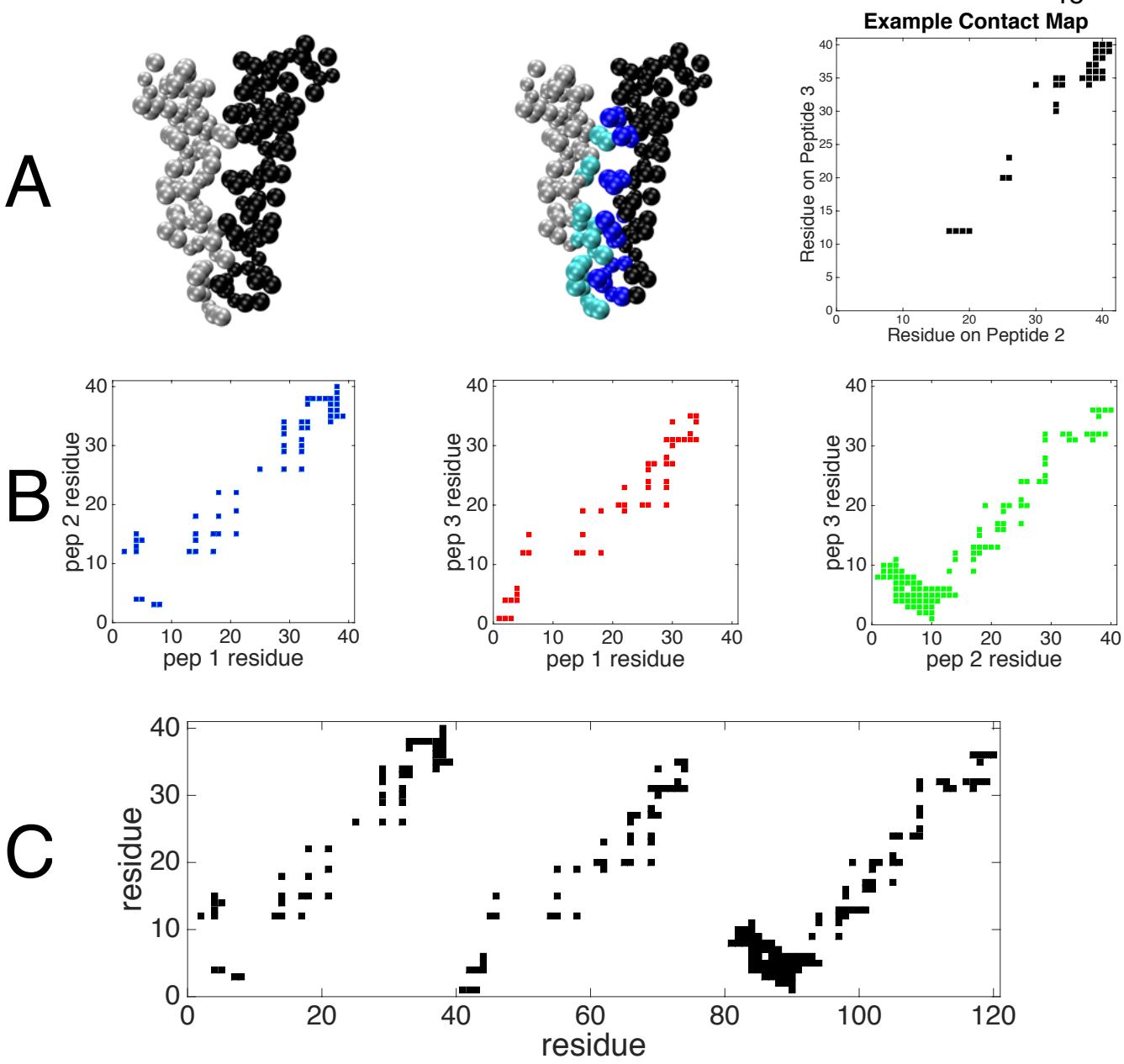


Figure 4. Simulation snapshots are quantified and clustered using contact maps. (a) Between each pair of peptides (left), contacts are defined as residues with nearest atom-atom distances of <0.7nm (middle, light and royal blue). This information is stored in an array of binary variables, a contact map (right). (b) The 3 pairs of peptides in each snapshot produce 3 separate maps, and for clustering the three are combined into a single 120x40 binary variable map (c).

for all snapshots. Snapshots that displayed trimerization, defined in coarse grained simulations as the presence of 5 or more contacts between 2 or more pairs, were considered for clustering. 31,081 resulting snapshots were clustered via K-means. We

used a large number of centers (200) to avoid issues that can arise from under-fitting high dimensional data with too few centers. The medoid snapshot of each cluster is plotted in Figure 5, and the population is spread across the PCA space. We chose 50 of the 200 K-means medoids randomly to simulate at all-atom resolution. The final 50 targets are also plotted in Figure 5 and form a diverse set of trimer assemblies.

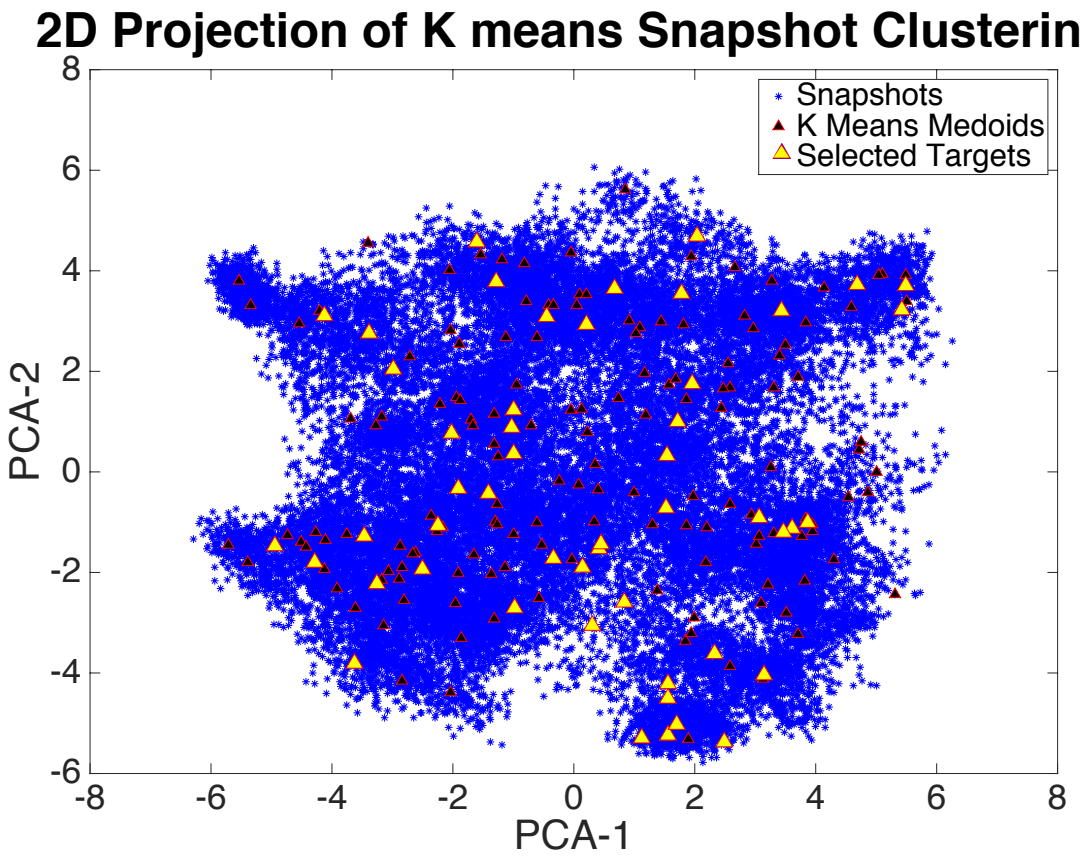


Figure 5. K-means medoids and final target snapshots sample the snapshot population. Plotted on a 2D projection onto the first two principal component axes, the K-means medoids (black triangles) spread over the whole population. A random sample of 50 medoids chosen out of the 200 medoids yields a representative sample of the population space (yellow triangles).

The final selection of 50 snapshots, taken from 20 of the 24 coarse grain simulations, were simulated at all-atom resolution. Once the lipid and protein coarse grain beads were converted into atomic coordinates (Figure 6), the membrane and

protein environment was re-solvated and ions added back in full atomic representation. After several relaxation and equilibration steps, the trimers were simulated for more than 190 nanoseconds.

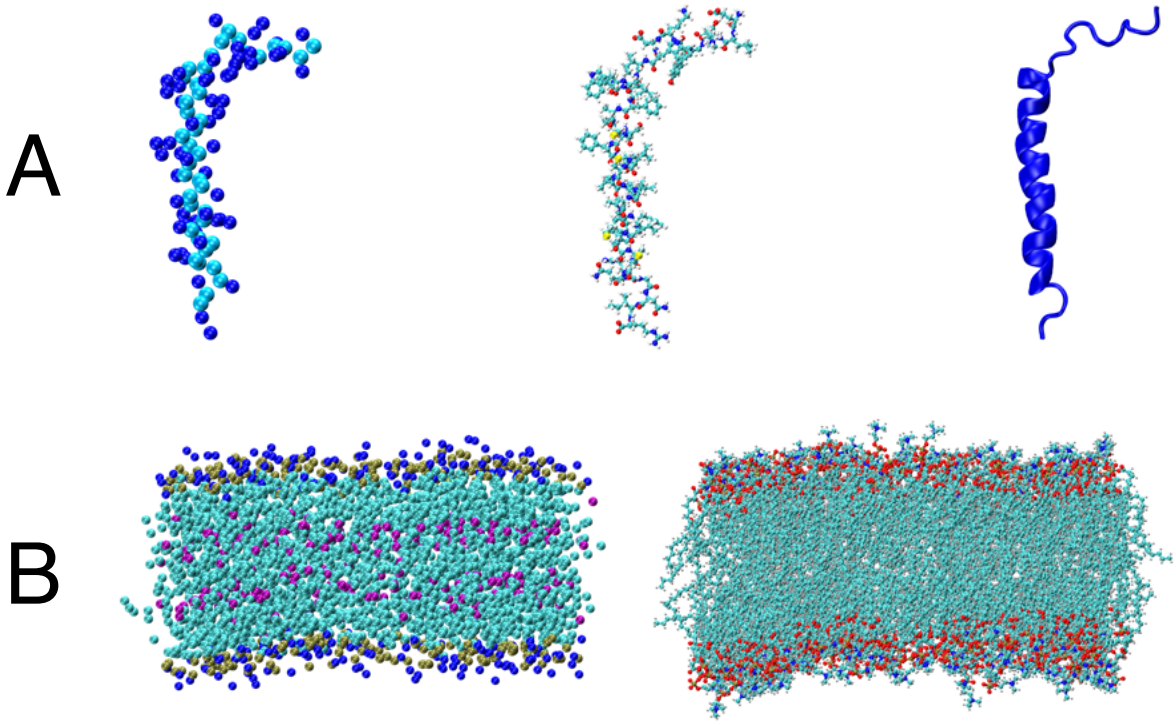


Figure 6. Coarse grained (CG) beads are transformed into groups of atoms during the resolution switch to all-atom. (a) Individual coarse grain peptides (left) are made of 96 beads. There are 671 atoms in each peptide, whose placement is determined by the "backward" tool described in Methods. An all-atom TMD peptide retains the structure from its CG structure (middle), shown in bonds representation (middle) and as a cartoon (right). (b) Each POPC molecule in the CG bilayer (left) is transformed from 13 beads into 134 atoms in place, yielding the bilayer on the right.

We first probed if the trimers remained stable throughout our simulations. No trajectories exhibited gross separation of peptides. The center of mass distance between peptide pairs is plotted for the length of simulations in Figure 8A. With the relaxation of helical restraints and the increased degrees of freedom allowed in our all-atom simulations, we expected that the specific modes of interaction explored in the coarse grained simulations would change upon the resolution increase.

Contact maps were used to gain residue-level interaction information, as with the coarse grained simulations. The number of contacts was averaged over the 50 simulations and plotted in Figure 7A. While the trimers remain stably compact, the number of contacts in simulation snapshots fell as the simulation begins, possibly due to the immediate relaxation of the atomic resolution monomers. The number of contacts evened out at around 380 contacts per frame. Due to our diverse starting population, we hypothesized that contacts appearing most in these simulations would be crucial to trimer stability. We measured the probability of finding each contact over all snapshots. A complete mapping of contact probability is shown in Figure 9. Contact probability greater than zero is mostly exhibited along the diagonal line of equality due to the monomers residing in register in the membrane. Sharp peaks in probability are located in the probability pattern between membrane-inserted residues 185 to 210. Broader probability is seen for contacts outside of the membrane environment, possibly due to a decrease in alpha-helical structure. The peaks in probability along residues 190-210 are spaced along one face of the helical peptide, at every 3 or 4 residues. The data show that inter-peptide contacts in the membrane acted along one face of the helical monomer.

The probabilities of contacts in our simulations are a measure of their relative contributions in the stability of trimeric complexes. A ranking of the top 5 contacts is shown in Figure 10. All are hydrophobic residues in the membrane embedded region, and all are closer to the C terminal end of the TMD. The C terminal end is closest to the interior of the virus, and thus farthest from the hemagglutinin ectodomain. The most common contact was between neighboring phenylalanine residues at residue 205,

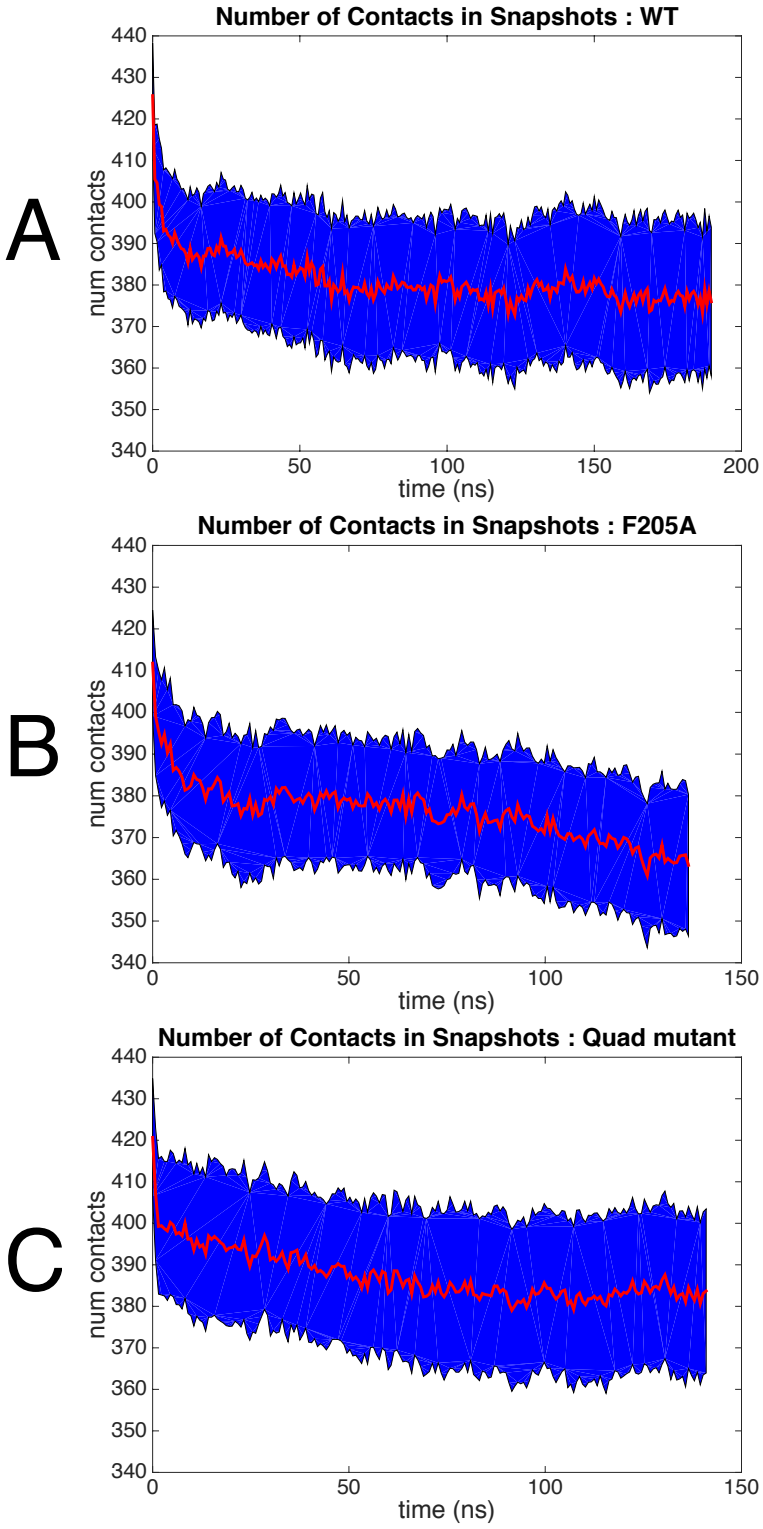
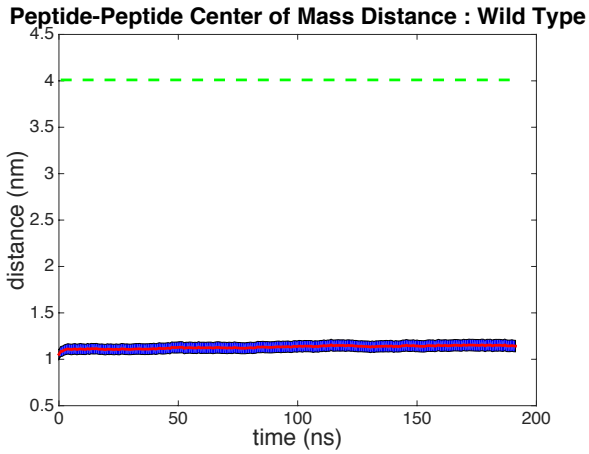
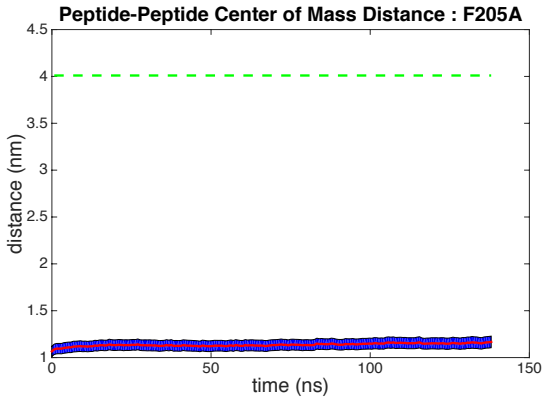


Figure 7. The number of contacts between peptides is stable after an initial decrease. The number of contacts in each snapshot is plotted over time for each simulation set. After an initial decrease, all simulations ((a) wild type, (b) F205A mutant, and (c) Quadruple mutant) relax at a new number of contacts. Averages are shown in red, with surrounding blue 95% confidence intervals. Each trace is plotted over the entire length of simulations.

A



B



C

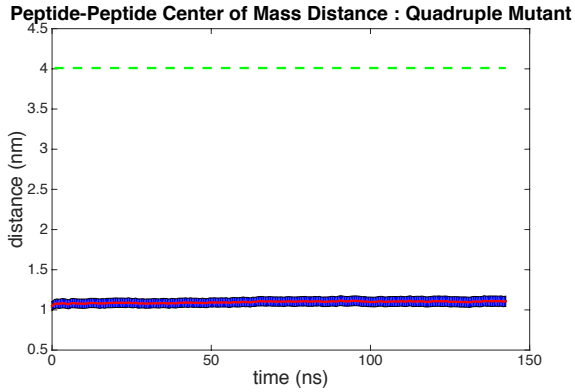


Figure 8. Trimeric complexes remain tightly associated over the course of all wild type and alanine mutant simulations. The center of mass distance between each peptide is averaged over both peptides and simulations for the (a) wild type simulations, (b) single mutant F205A simulations, and (c) quadruple mutant L198A V201A F205A W208A simulations. Averages in red are surrounded by a blue 95% confidence interval. The original CG starting distance before complexation is shown in the dotted green line for comparison. All simulations display stability for their duration.

occurring in 70% of all snapshots. L198, W208, V201, and L202 are the next most common. These hydrophobic residues interacted with others on the neighboring peptide, despite the relatively hydrophobic environment of the membrane interior. No single interaction involving a polar residue, which are at positions 190, 194, 195, and 199, appeared in more than 50% of snapshots. Nonetheless, a serine at position 194 contributes a smaller peak in the probability map. It lies on the same face of the helix as the top contacts.

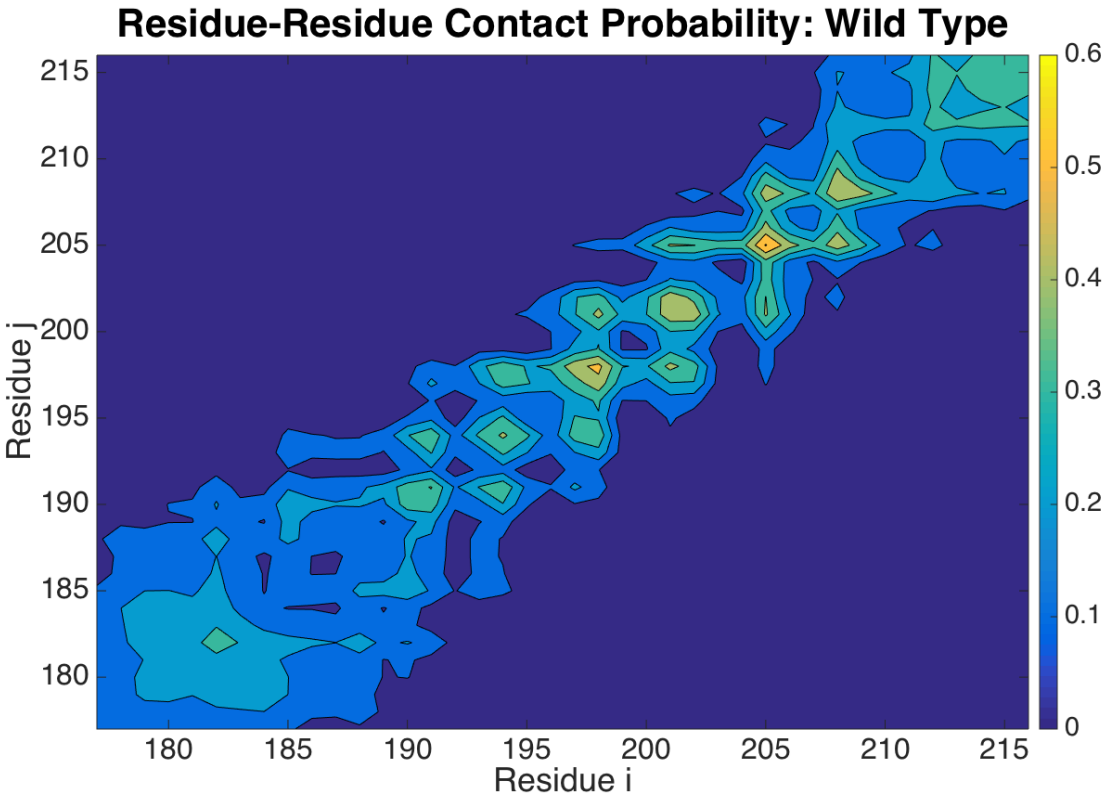


Figure 9. The probability of an inter-peptide contact highlights several frequent modes of contact. The probability of each contact is mapped over residues for wild type simulations after removing data accrued in the first 25 ns of each simulation. Areas in dark blue have no contacts in simulations, while areas in yellow have a high likelihood of contact. The probability follows a diagonal pattern due to the peptides residing in the membrane with similar depths. Several single contacts display probabilities of over 0.5, shown in bright yellow.

Rank	Wild Type			F205A			Quadruple Mutant		
	i	j	P(contact)	i	j	P(contact)	i	j	P(contact)
1	205 Phe	205 Phe	0.70±.06	198 Leu	198 Leu	0.61±.08	194 Ser	194 Ser	0.52±.08
2	198 Leu	198 Leu	0.60±.07	201 Val	201 Val	0.58±.07	197 Leu	197 Leu	0.52±.08
3	208 Trp	208 Trp	0.56±.07	208 Trp	208 Trp	0.57±.07	202 Leu	201 Ala	0.51±.06
4	201 Val	201 Val	0.55±.08	202 Leu	201 Val	0.53±.06	202 Leu	202 Leu	0.51±.08
5	202 Leu	201 Val	0.54±.05	198 Leu	197 Leu	0.53±.05	198 Ala	198 Ala	0.50±.07

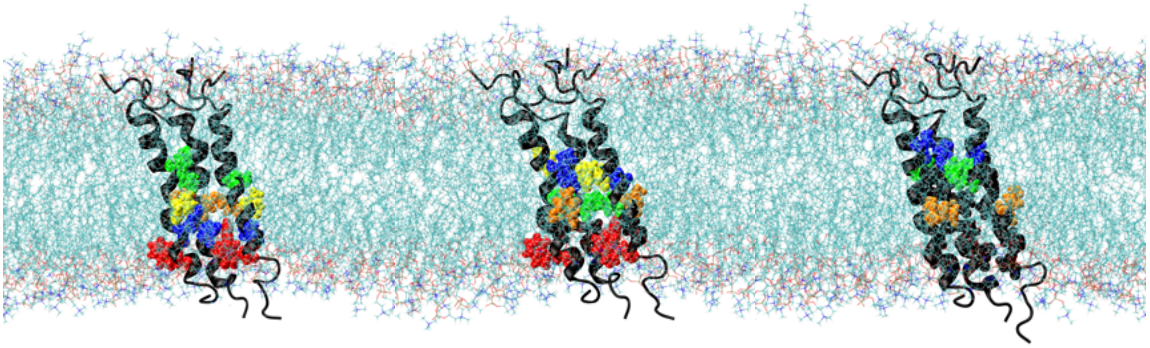


Figure 10. A discrete set of prevalent contacts defines the most favorable specific interaction modes between peptides. The most likely contact pairs are listed for each simulation set, along with their corresponding probabilities. Probabilities are accompanied by bounds defining a 95% confidence interval. Below, each top contact is displayed in a trimer, colored by its ranking as indicated in column 1 of the table.

3.3 All-Atom Molecular Dynamics Simulations: Mutant TMD

In order to test the influence of the most probable contact on trimer stability, we created a second set of all-atom simulations from the same starting snapshots, but mutated each phenylalanine at position 205 to an alanine. While alanine is still hydrophobic, this mutation effectively removed an aromatic ring from the transmembrane domain. The effect of this mutation on the average center of mass distance between peptides is plotted in Figure 8B. This remained stable in both the wild type and F205A mutant simulations. The number of contacts during the F205A simulations underwent the same initial decrease seen in the wild type simulations (Figure 7B). After leveling off at a similar value of around 380 contacts per snapshot, the

F205A trace appears to decrease again during the second half of the simulations, but well within error.

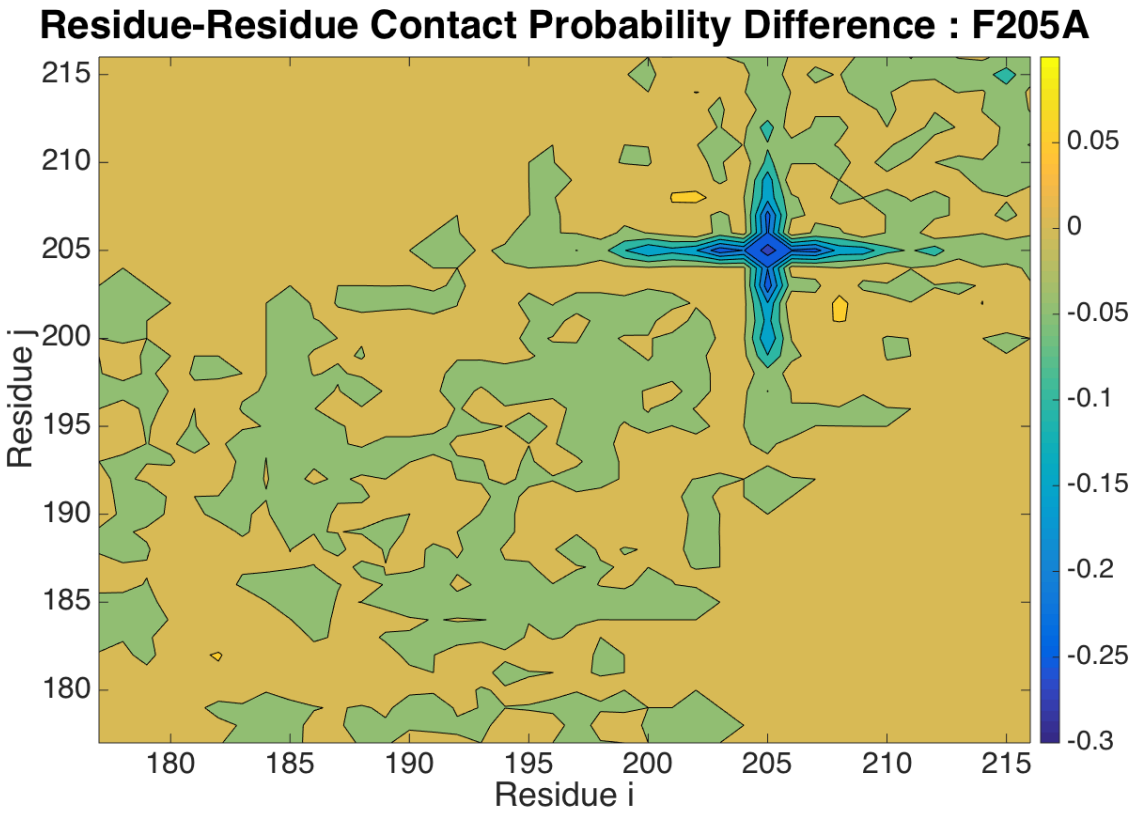


Figure 11. A large decrease in probability results at the mutation site. The difference in probability for all residue pairs between the single mutant F205A simulations and the wild type simulations is shown, with data collected during the first 25 ns of each simulation removed. Areas of similar probability are colored tan. Slight increases in probability are in yellow and slight decreases in green, while a large decrease is in blue.

While the mutation ablated contacts at residue 205, the effects were local to the mutation site. The change in probability of each contact in the F205A simulations from the wild type is mapped in Figure 11. Minor changes in probability span across the map, but a large decrease occurred at residue 205. However, this did not result in an overall decrease in trimer stability. Individual contacts that underwent a significant increase or decrease in probability are shown in Figure 14A. Significant decreases at the mutation

site are larger in magnitude than the compensatory increases surrounding it. All statistically significant increases and decreases are located within 15 residues of the mutation site. The most probable contacts in the F205A simulations are listed in Figure 10. Other top contacts from the wild type simulations do not change significantly.

Since a single mutation seemed only to affect contacts at the mutation site, a more perturbative mutation set of simulations was created. The top four contacting residues from the wild type simulations were all mutated to alanine to determine if the additive effect of multiple single mutations would remain local to mutation sites or perturb trimeric structure more generally. Residues Phe205, Trp208, Val201, and Leu198 were all mutated to alanine in the 50 start states shared by the other all-atom simulations. The peptide-peptide center of mass distance again shows no destabilization during the 140 nanoseconds of simulation each (Figure 8C). The number of contacts per frame also did not show a large change from the wild type (Figure 7C). The probability map for the quadruple mutant system (Figure 12) exhibits large decreases in interactions contributed by Phe205 and Trp208, visible as plus patterns. A smaller decrease appears for Leu198, and the smallest decrease for Val201. Since leucine and valine are smaller hydrophobic residues, without the large interacting rings that phenylalanine and tryptophan have, the mutation to alanine was not as effective at abrogating the contacts they formed. Nonetheless, mutating each residue to alanine had a large local effect on individual contacts, but not overall trimer stability.

While all decreases in probability at mutation sites were significant (Figure 14), many nearby residues display an increase in contact probability. Increases are smaller in magnitude than the mutation decreases (Figure 12), but many nearby residues

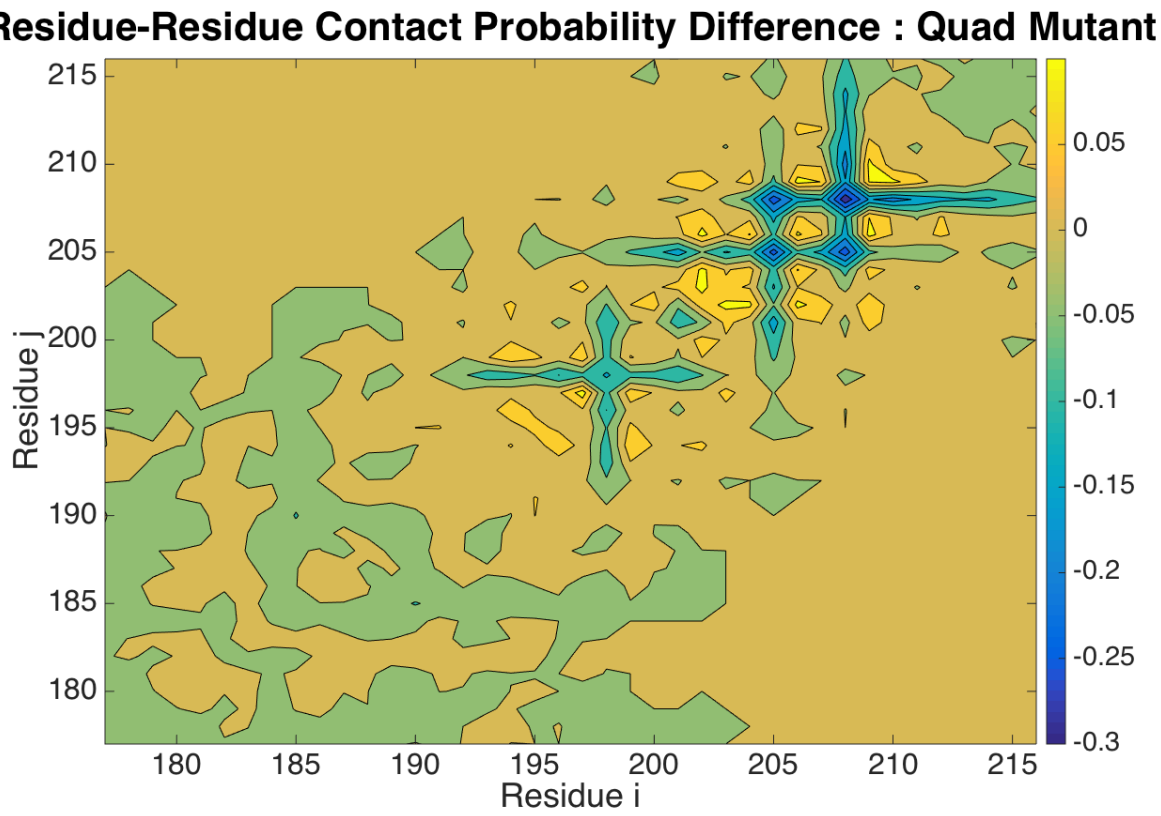


Figure 12. Mutation sites constitute areas of extended decreases in probability. The difference in probability for all residue pairs between the quadruple mutant L198A V201A F205A W208A simulations and the wild type simulations is shown, with data collected during the first 25 ns of each simulation removed. Areas of similar probability are colored tan. Slight increases in probability are in yellow and slight decreases in green, while a large decrease is in blue.

compensated for the loss of contact at mutation sites. This characteristic appears to be the reason behind trimeric stability in the case of TMD mutants. The top contacts formed in the quadruple mutant simulations still included residues 198 and 201, but at significantly reduced probability (Figure 14). A serine-serine interaction became the most prevalent contact after the previous top contacts were ablated. Located 4 residues toward the N terminal end of the peptide from the nearest mutation, this serine continued to follow the pattern of prevalent contacts being spaced by roughly one

helical turn. This interaction appeared in all cases, and did not significantly increase in the quadruple mutant simulations (Figure 14).

An additional set of ten simulations were created to investigate the effect of our all-atom simulations on a mutation known to block productive fusion. The trimers from ten all-atom starting snapshots were cut at residue 199, leaving only the N terminal 23 residues of the original construct, 15 of which were membrane-inserted in one leaflet. Three of the ten systems had charged carboxylates at their unprotonated C termini located in the membrane. Each was simulated for more than 80 nanoseconds.

Trimeric complexes made up of the unprotonated C termini TMD peptides became destabilized, while complexes in the protonated state were able to retain their structure (Figure 13). Unprotonated TMD peptides rapidly adjusted their orientation in the membrane in order for the charged C termini to lie in the membrane head groups, either rotating to become parallel with the membrane surface or penetrating the membrane more deeply. This reorientation caused a rise in inter-peptide distance and a large reduction in the number of contacts in each snapshot, especially at the charged terminus (Figure 13C). The membrane thinned significantly as the monomers restructured and a water-accessible pore was formed. The trimeric complexes of 6 out of the 7 protonated state simulations remained stable in structure, but also thinned the membrane. Interestingly, the shorter TMD monomers were able to span both membrane leaflets. While it seems that trimeric complexes of protonated TMD peptides can retain their structure, it is unclear if our complexes can form in the first place. A shortened TMD monomer may not lie perpendicular to the membrane surface, blocking the formation of our trimeric assemblies that lie perpendicular to the membrane surface. This rotation

may occur during or immediately following the insertion of the newly created HA into a membrane, preceding trimerization. Further reductions in TMD length may lead to expression or structural problems in the protein by further membrane destabilization or enhanced insertion errors.

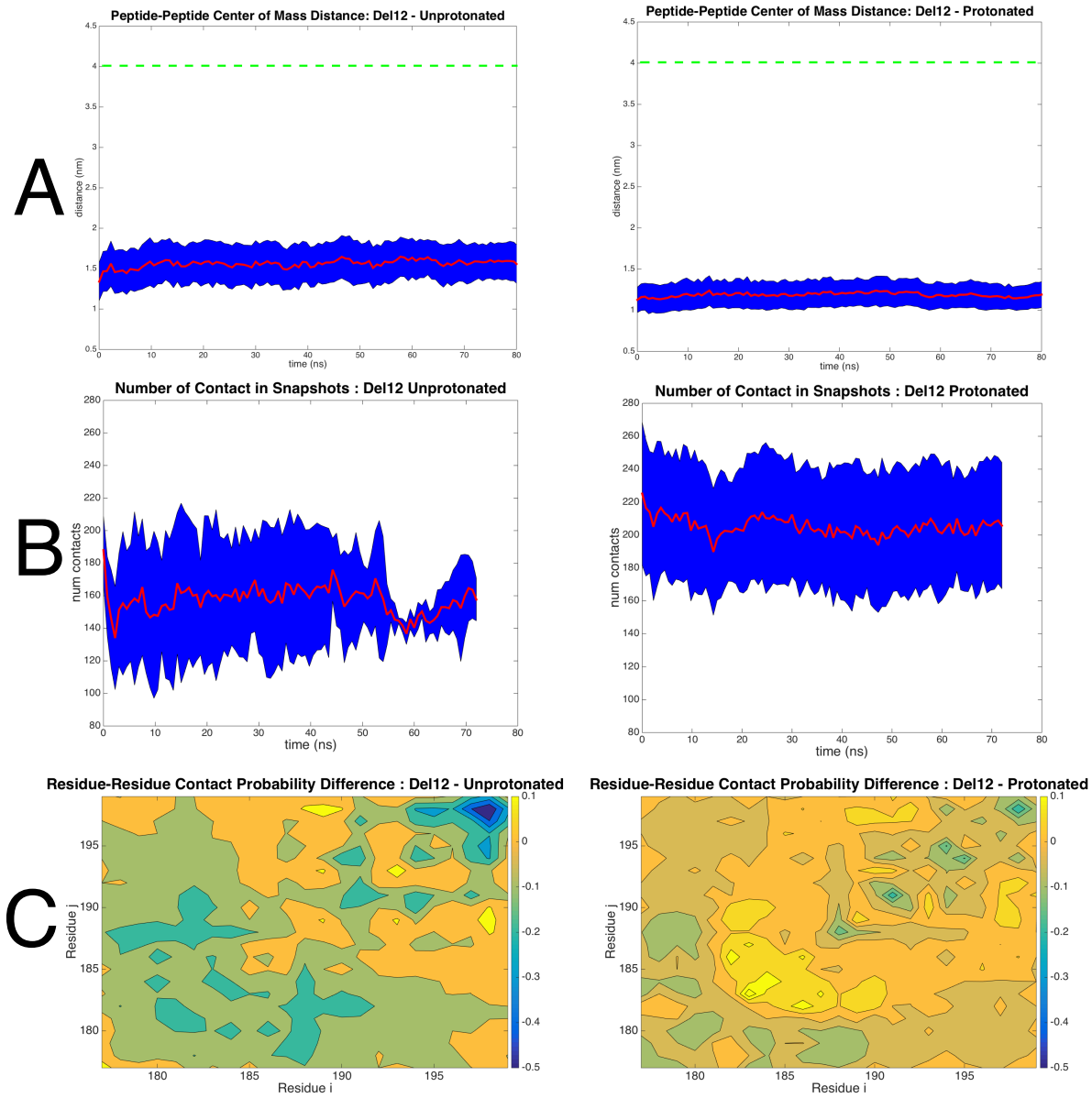
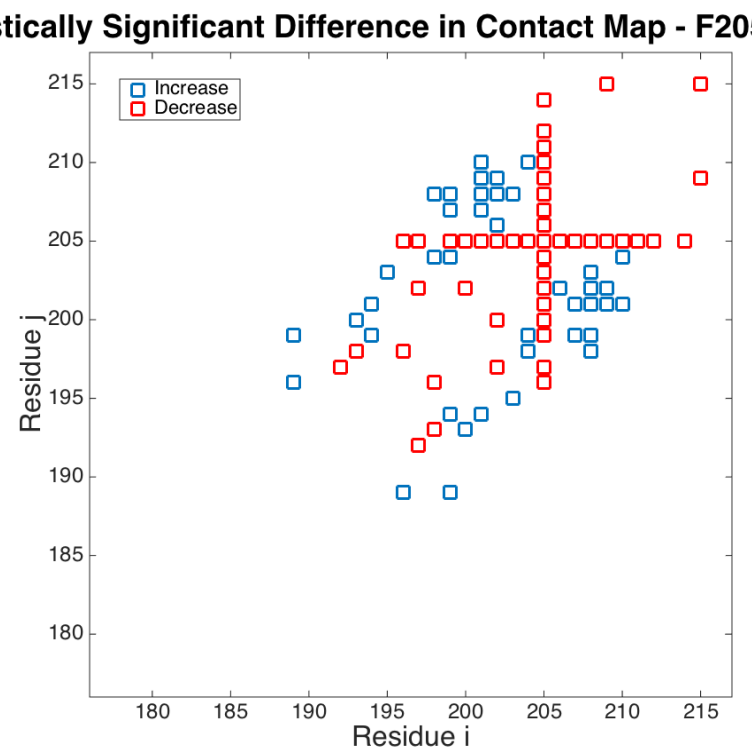


Figure 13. Stable behavior is exhibited in protonated del12 mutants, but not unprotonated del12 mutant simulations. (a) The inter-peptide center of mass distance is shown with mean values in red surrounded by a 95% confidence interval in blue. (b) The number of contacts over time is shown in red surrounded by a 95% confidence interval in blue. (c) The difference in contact probability of residues 177-199 from their probability in the wild type simulations. Areas of similar probability are colored tan. Slight increases in probability are in yellow and slight decreases in green, while a large decrease is in blue.

A



B

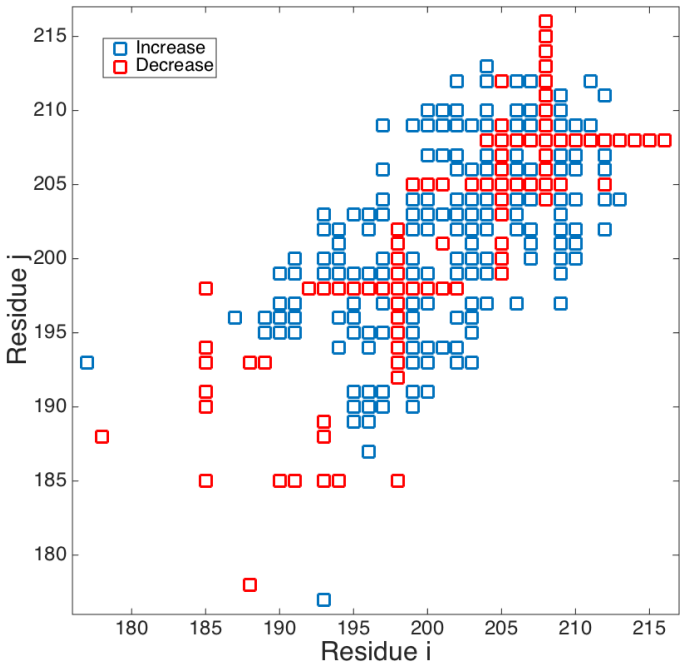


Figure 14. Areas of statistically significant changes in contact probability follow mutation sites most directly but show scattered additional differences. The results of a paired t-test on each residue-residue contact dataset were mapped, where a blue box indicates a statistically significant increase in the mutant probability and a red a decrease in the mutant probability. (a) displays results from changes between the wild type and single mutant F205A simulations and (b) for the wild type and quadruple mutant.

4. Discussion

Many studies on influenza hemagglutinin-mediated membrane fusion focus on the fusion peptide's effects, without taking into account the important role of the transmembrane domain. Here, we used a multi-scale simulation approach to investigate the interactions between neighboring transmembrane domains in a POPC bilayer. Trimeric complexes generated in high-throughput coarse grain simulations were examined for their stability and key features with all-atom simulations. Prevalent hydrophobic contacts provided a discrete set of key interaction features of the trimeric complex. Though the individual contacts formed by these discrete residues were removed via mutation, the overall trimeric structure remained stable, in agreement with the previously documented mutational plasticity of the region.

The apparent stability of TMD trimers is an important determinant in our understanding of viral fusion protein-mediated fusion. Breaking the hemifusion diaphragm requires both the fusion peptide and the transmembrane domain, but it remains unclear if their effect is concerted or additive. Since the TMD must anchor HA in the viral membrane, it is unlikely to be playing a membrane destabilization role as fusion peptides are thought to. Previous work has shed some light on possible oligomeric activity of the TMD, but here we demonstrate stable trimeric complexes that are resistant to mutations, even at common inter-monomer contacts. Recent work has suggested three to four neighboring HA trimers are required for fusion [54,62], in agreement with older kinetic studies [55]. Therefore, we can expect multiple copies of the HA trimer surrounding a fusion site. Having each copy structured in a trimeric complex would prevent random interaction between single monomers of neighboring HA

proteins, but would allow for trimer to trimer interactions at the hemifusion site. Future work stemming from this study will include probing possible interactions between multiple TMD trimers in the membrane.

While Chang et al. showed a loose propensity for fusion peptides to interact with TMD peptides, this needs to be investigated further. Multi-scale simulations could provide computational predictions of any functional interactions between fusion peptides and TMD trimers. Studies of related viruses have predicted these interactions to be important in the fusion pathway. The fusion peptides, which are not thought to form trimeric bundles, may interact with a TMD trimer during hemifusion or in an expanding fusion pore. An interaction between loose fusion peptides and structured TMDs may localize the fusion peptides leading into fusion. This could concentrate any perturbative effect the fusion peptide has on the membrane, possibly creating a fusion pore. Common contacts were made along what appears to be a single face of the TMD in our simulations. The proposed specific interaction between hydrophobic contacts at this site may present interaction sites on the other side of the TMD for fusion peptides as fusion proceeds. Without a structured TMD trimer the fusion peptide may not be able to interact, or any interaction between loose fusion peptides and loose TMD peptides may not lead to a productive fusion pore. Future work adding fusion peptides into membranes with multiple copies of TMD trimers has been started.

We have also began further investigations into measuring the relative change in stability between the wild type and quadruple mutant simulations, after our simulations did not show gross destabilization. By comparing the work needed to pull trimers apart, we hope to quantify the difference in free energy of dissociation between the two

constructs. If we have not destabilized the TMD trimer with the quadruple mutant, this metric will be useful in measuring further mutations to the domain to find a TMD that will not trimerize as favorably. This mutant can then be used to confirm the functional importance of a TMD trimer experimentally in membrane fusion assays. If a less favorably-interacting TMD inhibits fusion kinetics then we can expect the trimeric structure to play a mechanistic role during fusion. Such an effect has been shown for the other major influenza spike protein, Neuraminidase, where a less stably interacting transmembrane mutant can interfere with the function of the entire protein assembly [69].

Interestingly, a 12-residue deletion mutant known to inhibit fusion at the hemifusion intermediate still showed the ability to stabilize trimers in our simulations when in the protonated state. However, the immediate membrane environment was severely thinned and allowed for nearly continuous water penetration at the complex site. These destabilizing effects may interfere with proper protein expression and structure in further deletions as is seen in previous work. Additional simulations can be run to determine if the shortened TMD can form trimers before destabilization, since our study allowed the mutants to start as trimeric complexes. In our simulations, having a terminal end of a TMD peptide is not favorable in the hydrophobic core of the membrane, as they quickly rotated to become parallel to the membrane surface at the hydrophilic head groups or penetrated more deeply to interact with the bottom leaflet head groups. Monomers creating encounter complexes in these states may not form stable trimeric complexes like the full-length TMD.

In conclusion, our simulations have demonstrated that TMD trimers can form stable trimeric complexes in the absence of the HA ectodomain. Mutations, even to residues participating in the most common contacts, did not grossly perturb this property in our simulations. As a necessary part of the influenza membrane fusion machinery, this result provides one piece of clarity in the unresolved mechanism of protein mediated membrane fusion, and also provides a basis for future work simulating multiple copies of TMD trimers and fusion peptides in membrane environments.

References

[1] Wilson, I. A., Skehel, J. J., & Wiley, D. C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature, 289(5796), 366-373.

[2] Tatulian, S. A., & Tamm, L. K. (2000). Secondary structure, orientation, oligomerization, and lipid interactions of the transmembrane domain of influenza hemagglutinin. Biochemistry, 39(3), 496-507.

[3] Gruenke, J. A., Armstrong, R. T., Newcomb, W. W., Brown, J. C., & White, J. M. (2002). New insights into the spring-loaded conformational change of influenza virus hemagglutinin. Journal of virology, 76(9), 4456-4466.

[4] Tamm, L. K. (2003). Hypothesis: spring-loaded boomerang mechanism of influenza hemagglutinin-mediated membrane fusion. Biochimica et Biophysica Acta (BBA)-Biomembranes, 1614(1), 14-23.

[5] Huang, Q., Opitz, R., Knapp, E. W., & Herrmann, A. (2002). Protonation and stability of the globular domain of influenza virus hemagglutinin. Biophysical journal, 82(2), 1050-1058.

[6] Doms, R. W., Helenius, A., & White, J. (1985). Membrane fusion activity of the influenza virus hemagglutinin. The low pH-induced conformational change. Journal of Biological Chemistry, 260(5), 2973-2981.

[7] Carr, C. M., Chaudhry, C., & Kim, P. S. (1997). Influenza hemagglutinin is spring-loaded by a metastable native conformation. Proceedings of the National Academy of Sciences, 94(26), 14306-14313.

[8] Chen, J., Skehel, J. J., & Wiley, D. C. (1999). 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. Proceedings of the National Academy of Sciences, 96(16), 8967-8972.

[9] Harter, C., James, P., Bächi, T., Semenza, G., & Brunner, J. (1989). Hydrophobic binding of the ectodomain of influenza hemagglutinin to membranes occurs through the" fusion peptide". Journal of Biological Chemistry, 264(11), 6459-6464.

[10] Han, X., Bushweller, J. H., Cafiso, D. S., & Tamm, L. K. (2001). Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. Nature Structural & Molecular Biology, 8(8), 715-720.

[11] Sun, Y., & Weliky, D. P. (2009). 13C– 13C Correlation Spectroscopy of Membrane-Associated Influenza Virus Fusion Peptide Strongly Supports a Helix-Turn-Helix Motif and Two Turn Conformations. Journal of the American Chemical Society, 131(37), 13228-13229.

[12] Lorieau, J. L., Louis, J. M., & Bax, A. (2010). The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid: water interface. Proceedings of the National Academy of Sciences, 107(25), 11341-11346.

[13] Lorieau, J. L., Louis, J. M., & Bax, A. (2011). Helical Hairpin Structure of Influenza Hemagglutinin Fusion Peptide Stabilized by Charge– Dipole Interactions between the N-Terminal Amino Group and the Second Helix. Journal of the American Chemical Society, 133(9), 2824-2827.

[14] Légaré, S., & Lagüe, P. (2012). The Influenza fusion peptide adopts a flexible flat V conformation in membranes. Biophysical journal, 102(10), 2270-2278.

[15] Huang, Q., Chen, C. L., & Herrmann, A. (2004). Bilayer conformation of fusion peptide of influenza virus hemagglutinin: a molecular dynamics simulation study. Biophysical journal, 87(1), 14-22.

[16] Larsson, P., & Kasson, P. M. (2013). Lipid tail protrusion in simulations predicts fusogenic activity of influenza fusion peptide mutants and conformational models. PLoS computational biology, 9(3), e1002950.

[17] Vaccaro, L., Cross, K. J., Kleinjung, J., Straus, S. K., Thomas, D. J., Wharton, S. A., ... & Fraternali, F. (2005). Plasticity of influenza haemagglutinin fusion peptides and their interaction with lipid bilayers. Biophysical journal, 88(1), 25-36.

[18] Han, X., Steinhauer, D. A., Wharton, S. A., & Tamm, L. K. (1999). 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, 38(45), 15052-15059.

[19] Qiao, H., Armstrong, R. T., Melikyan, G. B., Cohen, F. S., & White, J. M. (1999). A specific point mutant at position 1 of the influenza hemagglutinin fusion peptide displays a hemifusion phenotype. Molecular biology of the cell, 10(8), 2759-2769.

[20] Li, Y., Han, X., Lai, A. L., Bushweller, J. H., Cafiso, D. S., & Tamm, L. K. (2005). 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. Journal of virology, 79(18), 12065-12076.

[21] Chernomordik, L. V., & Kozlov, M. M. (2008). Mechanics of membrane fusion. Nature structural & molecular biology, 15(7), 675-683.

[22] Cohen, F. S., & Melikyan, G. B. (2004). The energetics of membrane fusion from binding, through hemifusion, pore formation, and pore enlargement. The Journal of membrane biology, 199(1), 1-14.

[23] Armstrong, R. T., Kushnir, A. S., & White, J. M. (2000). The transmembrane domain of influenza hemagglutinin exhibits a stringent length requirement to support the hemifusion to fusion transition. The Journal of cell biology, 151(2), 425-438.

[24] Melikyan, G. B., Lin, S., Roth, M. G., & Cohen, F. S. (1999). Amino acid sequence requirements of the transmembrane and cytoplasmic domains of influenza virus hemagglutinin for viable membrane fusion. Molecular biology of the cell, 10(6), 1821-1836.

[25] Melikyan, G. B., White, J. M., & Cohen, F. S. (1995). GPI-anchored influenza hemagglutinin induces hemifusion to both red blood cell and planar bilayer membranes. The Journal of cell biology, 131(3), 679-691.

[26] Markosyan, R. M., Cohen, F. S., & Melikyan, G. B. (2000). The lipid-anchored ectodomain of influenza virus hemagglutinin (GPI-HA) is capable of inducing nonenlarging fusion pores. Molecular biology of the cell, 11(4), 1143-1152.

[27] Melikyan, G. B., Markosyan, R. M., Roth, M. G., & Cohen, F. S. (2000). A point mutation in the transmembrane domain of the hemagglutinin of influenza virus stabilizes a hemifusion intermediate that can transit to fusion. Molecular biology of the cell, 11(11), 3765-3775.

[28] Schroth-Diez, B., Ponimaskin, E., Reverey, H., Schmidt, M. F., & Herrmann, A. (1998). Fusion activity of transmembrane and cytoplasmic domain chimeras of the influenza virus glycoprotein hemagglutinin. Journal of virology, 72(1), 133-141.

[29] Smith, E. C., Smith, S. E., Carter, J. R., Webb, S. R., Gibson, K. M., Hellman, L. M., ... & Dutch, R. E. (2013). Trimeric Transmembrane Domain Interactions in

Paramyxovirus Fusion Proteins ROLES IN PROTEIN FOLDING, STABILITY, AND FUNCTION. Journal of Biological Chemistry, 288(50), 35726-35735.

[30] Smith, E. C., Culler, M. R., Hellman, L. M., Fried, M. G., Creamer, T. P., & Dutch, R. E. (2012). Beyond anchoring: The expanding role of the hendra virus fusion protein transmembrane domain in protein folding, stability, and function. Journal of virology, 86(6), 3003-3013.

[31] Donald, J. E., Zhang, Y., Fiorin, G., Carnevale, V., Slochower, D. R., Gai, F., ... & DeGrado, W. F. (2011). Transmembrane orientation and possible role of the fusogenic peptide from parainfluenza virus 5 (PIV5) in promoting fusion. Proceedings of the National Academy of Sciences, 108(10), 3958-3963.

[32] Bissonnette, M. L. Z., Donald, J. E., DeGrado, W. F., Jardetzky, T. S., & Lamb, R. A. (2009). Functional analysis of the transmembrane domain in paramyxovirus F protein-mediated membrane fusion. Journal of molecular biology, 386(1), 14-36.

[33] Parton, D. L., Tek, A., Baaden, M., & Sansom, M. S. (2013). Formation of raft-like assemblies within clusters of influenza hemagglutinin observed by MD simulations. PLoS computational biology, 9(4), e1003034.

[34] Scolari, S., Engel, S., Krebs, N., Plazzo, A. P., De Almeida, R. F., Prieto, M., ... & Herrmann, A. (2009). Lateral distribution of the transmembrane domain of influenza virus hemagglutinin revealed by time-resolved fluorescence imaging. Journal of Biological Chemistry, 284(23), 15708-15716.

[35] Kordyukova, L. V., Serebryakova, M. V., Polyansky, A. A., Kropotkina, E. A., Alexeevski, A. V., Veit, M., ... & Baratova, L. A. (2011). Linker and/or transmembrane regions of influenza A/Group-1, A/Group-2, and type B virus hemagglutinins are packed differently within trimers. Biochimica et Biophysica Acta (BBA)-Biomembranes, 1808(7), 1843-1854.

[36] Chang, D. K., Cheng, S. F., Kantchev, E. A., Lin, C. H., & Liu, Y. T. (2008). Membrane interaction and structure of the transmembrane domain of influenza hemagglutinin and its fusion peptide complex. BMC biology, 6(1), 2.

[37] Victor, B. L., Baptista, A. M., & Soares, C. M. (2012). Structural Determinants for the Membrane Insertion of the Transmembrane Peptide of Hemagglutinin from Influenza Virus. Journal of chemical information and modeling, 52(11), 3001-3012.

[38] Kasson, P. M., & Pande, V. S. (2011). A bundling of viral fusion mechanisms. Proceedings of the National Academy of Sciences, 108(10), 3827-3828.

[39] Jo, S., Kim, T., Iyer, V. G., & Im, W. (2008). CHARMM-GUI: a web-based graphical user interface for CHARMM. Journal of computational chemistry, 29(11), 1859-1865.

[40] Klauda, J. B., Venable, R. M., Freites, J. A., O'Connor, J. W., Tobias, D. J., Mondragon-Ramirez, C., ... & Pastor, R. W. (2010). Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. The journal of physical chemistry B, 114(23), 7830-7843.

[41] Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P., & De Vries, A. H. (2007). The MARTINI force field: coarse grained model for biomolecular simulations. The Journal of Physical Chemistry B, 111(27), 7812-7824.

[42] Wassenaar, T. A., Pluhackova, K., Böckmann, R. A., Marrink, S. J., & Tieleman, D. P. (2014). Going backward: a flexible geometric approach to reverse transformation from coarse grained to atomistic models. Journal of Chemical Theory and Computation, 10(2), 676-690.

[43] de Jong, D. H., Singh, G., Bennett, W. D., Arnarez, C., Wassenaar, T. A., Schafer, L. V., ... & Marrink, S. J. (2012). Improved parameters for the martini coarse-grained protein force field. Journal of Chemical Theory and Computation, 9(1), 687-697.

[44] Yesylevskyy, S. O., Schäfer, L. V., Sengupta, D., & Marrink, S. J. (2010). Polarizable water model for the coarse-grained MARTINI force field. PLoS computational biology, 6(6), e1000810.

[45] Bussi, G., Donadio, D., & Parrinello, M. (2007). Canonical sampling through velocity rescaling. The Journal of chemical physics, 126(1), 014101.

[46] Parrinello, M., & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. Journal of Applied physics, 52(12), 7182-7190.

[47] Hess, B., Kutzner, C., Van Der Spoel, D., & Lindahl, E. (2008). GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. Journal of chemical theory and computation, 4(3), 435-447.

[48] Bjelkmar, P., Larsson, P., Cuendet, M. A., Hess, B., & Lindahl, E. (2010). Implementation of the CHARMM force field in GROMACS: Analysis of protein stability effects from correction maps, virtual interaction sites, and water models. Journal of Chemical Theory and Computation, 6(2), 459-466.

[49] Hess, B. (2008). P-LINCS: A parallel linear constraint solver for molecular simulation. Journal of Chemical Theory and Computation, 4(1), 116-122.

[50] Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. The Journal of chemical physics, 98(12), 10089-10092.

[51] Kellogg, E. H., Lange, O. F., & Baker, D. (2012). Evaluation and optimization of discrete state models of protein folding. The journal of physical chemistry B, 116(37), 11405-11413.

[52] Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. Journal of molecular graphics, 14(1), 33-38.

[53] MATLAB 2014b, The MathWorks, Inc., Natick, Massachusetts, United States.

[54] Ivanovic, T., Choi, J. L., Whelan, S. P., van Oijen, A. M., & Harrison, S. C. (2013). Influenza-virus membrane fusion by cooperative fold-back of stochastically induced hemagglutinin intermediates. Elife, 2, e00333.

[55] Danieli, T., Pelletier, S. L., Henis, Y. I., & White, J. M. (1996). Membrane fusion mediated by the influenza virus hemagglutinin requires the concerted action of at least three hemagglutinin trimers. The Journal of cell biology, 133(3), 559-569.

[56] Kemble, G. W., Danieli, T., & White, J. M. (1994). Lipid-anchored influenza hemagglutinin promotes hemifusion, not complete fusion. Cell, 76(2), 383-391.

[57] Kasson, P. M., & Pande, V. S. (2007, January). Predicting structure and dynamics of loosely-ordered protein complexes: influenza hemagglutinin fusion peptide. In Pacific Symposium on Biocomputing (Vol. 12, pp. 40-50).

[58] Wassenaar, T. (2010, October). *numerical matrix from xpm file* [Mailing list post]. Retrieved from https://groups.google.com/forum/#!topic/archive-gmx-users/FzgrP2tgM4g5

[59] Bond, P. J., & Sansom, M. S. (2006). Insertion and assembly of membrane proteins via simulation. Journal of the American Chemical Society, 128(8), 2697-2704.

[60] Crowet, J. M., Parton, D. L., Hall, B. A., Steinhauer, S., Brasseur, R., Lins, L., & Sansom, M. S. (2012). Multi-scale simulation of the simian immunodeficiency virus fusion peptide. The Journal of Physical Chemistry B, 116(46), 13713-13721.

[61] Kasson, P. M., & Pande, V. S. (2010). Cross-Graining: Efficient Multi-Scale Simulation via Markov State Models. In Pacific Symposium on Biocomputing (Vol. 15, pp. 260-268).

[62] Floyd, D. L., Ragains, J. R., Skehel, J. J., Harrison, S. C., & van Oijen, A. M. (2008). Single-particle kinetics of influenza virus membrane fusion. Proceedings of the National Academy of Sciences, 105(40), 15382-15387.

[63] Kasson, P. M., Lindahl, E., & Pande, V. S. (2010). Atomic-resolution simulations predict a transition state for vesicle fusion defined by contact of a few lipid tails. PLoS computational biology, 6(6), e1000829.

Chicago

[64] Hummer, G., & Szabo, A. (2001). Free energy reconstruction from nonequilibrium single-molecule pulling experiments. Proceedings of the National Academy of Sciences, 98(7), 3658-3661.

[65] Liphardt, J., Dumont, S., Smith, S. B., Tinoco, I., & Bustamante, C. (2002). Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. Science, 296(5574), 1832-1835.

[66] Woo, H. J., & Roux, B. (2005). Calculation of absolute protein–ligand binding free energy from computer simulations. Proceedings of the national academy of sciences of the united states of america, 102(19), 6825-6830.b

[67] Sauter, N. K., Hanson, J. E., Glick, G. D., Brown, J. H., Crowther, R. L., Park, S. J., ... & Wiley, D. C. (1992). Binding of influenza virus hemagglutinin to analogs of its cell-surface receptor, sialic acid: analysis by proton nuclear magnetic resonance spectroscopy and X-ray crystallography. Biochemistry, 31(40), 9609-9621.

[68] Dopheide, T. A., & Ward, C. W. (1981). The location of the bromelain cleavage site in a Hong Kong influenza virus Haemagglutinin. The Journal of general virology, 52(Pt 2), 367-370.

[69] da Silva, D. V., Nordholm, J., Madjo, U., Pfeiffer, A., & Daniels, R. (2013). Assembly of subtype 1 influenza neuraminidase is driven by both the transmembrane and head domains. Journal of Biological Chemistry, 288(1), 644-653.