Microsecond Molecular Dynamics Simulations of Influenza Neuraminidase Suggest a Mechanism for the Increased Virulence of Stalk-Deletion Mutants

Jacob D. Durrant,*,[†] Robin M. Bush,[‡] and Rommie E. Amaro[†]

[†]Department of Chemistry & Biochemistry and the National Biomedical Computation Resource, University of California San Diego, La Jolla, California 92093, United States

[‡]Department of Ecology & Evolutionary Biology, University of California Irvine, Irvine, California 92697, United States

S Supporting Information

ABSTRACT: Deletions in the stalk of the influenza neuraminidase (NA) surface protein are associated with increased virulence, but the mechanisms responsible for this enhanced virulence are unclear. Here we use microsecond molecular dynamics simulations to explore the effect of stalk deletion on enzymatic activity, contrasting NA proteins from the A/swine/Shandong/N1/2009 strain both with and without a stalk deletion. By modeling and simulating neuraminidase apo glycoproteins embedded in complex-mixture lipid bilayers, we show that the geometry and dynamics of the neuraminidase enzymatic pocket may differ depending on stalk length, with possible repercussions on the binding of the endogenous sialylated-oligosaccharide receptors. We also use these simulations to predict previously unrecognized druggable "hotspots" on the neuraminidase surface that may prove useful for future efforts aimed at structure-based drug design.
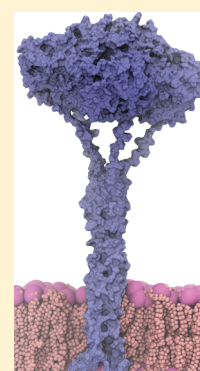
INTRODUCTION

There is an urgent need to better understand the molecular factors that govern influenza A virus (IAV) virulence. Seasonal (human-adapted) IAV is a major cause of morbidity and mortality each year, but the pandemic potential of IAV strains that originate in animal hosts poses an even more serious threat to public health. There have been four influenza pandemics in the past century, in 1918, 1957, 1968, and 2009. The precise origin of the 1918 pandemic strain, which killed roughly 50 million people, is unclear. However, the three subsequent pandemic strains resulted from the process of genetic reassortment, in which gene segments from different viruses mix during host coinfection to produce novel viral strains.[1]

There is much concern that a new pandemic strain might result from reassortment involving the highly pathogenic H5N1 and H7N9 avian influenza viruses currently circulating in Asia. The majority of these viruses undergo a deletion in the stalk of the neuraminidase surface protein, a characteristic that is associated with virulence when causing outbreaks in domestic poultry.[2] It is thus vitally important to understand the effect of stalk deletion on viral function.

Influenza A strains are named according to the alleles coding for their two surface antigens, the hemagglutinin (HA) and neuraminidase (NA) glycoproteins. These alleles are assigned numbers based on the order of their discovery.[3] Hemagglutinin (HA, 18 antigenically distinct alleles identified to date) facilitates viral adhesion to human host cells by binding to cellular sialylated-oligosaccharide receptors, and neuraminidase (NA, 11 antigenically distinct alleles) facilitates viral release by

cleaving sialic acid linkages. A number of factors modulate NA and HA activities. For example, hyperglycosylation and active-site mutation tend to reduce HA/sialic acid binding. NA activity is generally proportional to the length of the highly variable NA stalk,[2,4−9] and recent neuraminidase stalk deletions often produce highly pathogenic avian influenza viruses such as H5N1.[2]

Influenza virulence is determined in large part by the balance between NA and HA activities at the virion surface. Recent viral strains with stalk-deletion NA glycoproteins (NA_{del}) are more virulent because the reduced NA_{del} activity alters this balance. For example, the virulence of both a pandemic 2009 H1N1 virus and a more recent 2013 H7N9 strain was enhanced when deletion mutations containing 20 and 5 amino acids, respectively, were introduced into the corresponding NA stalks.[2,4−13] In these strains, reduced NA activity due to a stalk deletion may enhance viral infectivity by reducing oligosaccharide-receptor cleavage and thus increasing the number of receptors available for HA binding. This strategy is particularly advantageous when HA activity is itself compromised (e.g., due to hyperglycosylation, active-site mutations, etc.), leading many to express concerns that a current H5N1 strain might obtain a long-stalk NA through

reassortment with a human-adapted virus. Optimal viral replication is achieved only when the activities of NA and HA are ideally complementary.[2,4−7,9]

The mechanism by which reduced NA stalk height impacts NA activity remains uncertain. One prevailing theory suggests that NA$_{del}$ glycoproteins have reduced sialidase activity relative to wild type (NA$_{wt}$) because their diminished height hinders access to cellular sialylated-oligosaccharide receptors (i.e., the "limited access" theory).[4,8,9] The implication is that a towering canopy of HA effectively buries the short-stalk NA$_{del}$ such that cell-bound oligosaccharides simply cannot reach the NA enzymatic pockets. While there is merit to the limited-access theory, especially in cases where the stalk length is profoundly reduced, experiment and computational modeling suggest it may not fully explain the reduced NA$_{del}$ activity.

We hypothesize that, in addition to any limited-access effects, NA stalk length also reduces NA enzymatic activity by altering the binding affinity of sialic acid, the native substrate, to the enzymatic pocket. Given that the stalk is roughly 40 Å from the active site, we propose an indirect mechanism by which stalk length impacts binding-pocket dynamics at a distance. Other studies have demonstrated that the NA enzymatic pocket is highly flexible and so samples many conformations. For example, X-ray crystallography and molecular dynamics simulations of the isolated neuraminidase catalytic domain suggest that key loops surrounding the active site (i.e., the 150- and 430-loops) are highly mobile, permitting large cavities to form immediately adjacent to the main pocket.[14−19] Given that the enzyme is known to adopt a variety of conformational states, it is reasonable to hypothesize that stalk length may have a quasi-allosteric impact on ligand binding affinity by shifting the binding-pocket conformational population toward a subtly different distribution.[20−23]

To properly study the impact of stalk length, we built a full model of 2009 H1N1 neuraminidase as well as a model that included a 20 amino acid stalk-deletion mutation similar to one known to enhance virulence in a recent strains.[24] Our approach was inspired by Blumenkrantz et al.,[25] who replaced the long-stalk NA$_{wt}$ of a 2009 H1N1 pandemic virus with a short-stalk H5N1 NA$_{del}$. They attributed the resulting decrease in transmissibility at least in part to a reduction in the ability of the short-stalk NA$_{del}$ to cleave tethered substrates and penetrate the host mucus barrier. Rather than exactly replicating the work of Blumenkrantz et al., who inserted a foreign NA (from H5N1) into a H1N1 virus, we made our deletion directly in the H1N1 NA$_{wt}$ stalk itself. Our models were constructed using data from both X-ray crystallography (used to model the enzymatic head) and also protein−protein docking and modeling (used to model the stalk, transmembrane, and intravirion domains, which have never been crystallized). All NA models were embedded in complex-mixture lipid bilayers (Figure 1).

Subsequent atomistic molecular dynamics simulations of these systems support the theory that the NA stalk length impacts the dynamics of the NA binding pocket and therefore the affinity for the endogenous substrate. Furthermore, these simulations reveal previously uncharacterized druggable "hotspots" that may open up additional avenues for structure-based drug design. Previous simulations have included only the NA head; our models, which include additional NA domains as well as the associated lipid bilayer, allow us to take the next steps toward more realistic atomic simulations of viral components
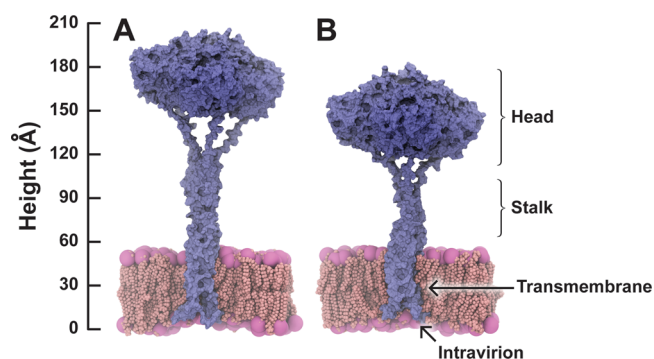
Figure 1. Equilibrated neuraminidase models used in the current study: (A) wild-type (long-stalk) NA$_{wt}$ and (B) stalk-deletion NA$_{del}$. The four domains of the glycoprotein are indicated.

■ MATERIALS AND METHODS

Constructing the Neuraminidase Models. To generate a complete model of long-stalk (wild-type) apo NA$_{wt}$, a monomer of the tetrameric NA head was built by homology using Schrödinger's Prime and the NA sequence of the 2009 H1N1 pandemic reference strain A/California/04/2009 (GenBank accession FJ966084), which codes for a NA protein identical to that of the A/England/195/09 virus used by Blumenkrantz et al.[25] Default Prime parameters were used. The monomeric homology model, based on the 3NSS crystal structure,[26] was missing a single lysine residue on the C terminus, which was added manually. The appropriate tetramer was formed by aligning copies of the homology-model monomer to chains A−D of the 2HU4 crystal structure[27] using MultiSeq,[28] as implemented in VMD.[29] The resulting tetrameric head was then processed with Schrödinger Maestro's Protein Preparation Wizard and subjected to repeated minimization in the Maestro environment. Calcium cations were added using the 3CL0 structure[30] as a guiding template. Coordination was manually inspected and corrected when necessary.

In contrast, the stalk and transmembrane domains were built entirely de novo. Both SSPro 4.5,[31] a neural-network-based algorithm for predicting secondary structure from primary sequence, and ABTMpro,[32] a SVM algorithm for determining whether a given transmembrane segment is α helical or β barrel, suggested that the stalk has an α-helical structure. We therefore used the Amber XLEAP module to generate a single α helix with the appropriate long-stalk sequence.

RosettaDock 3.4[33−37] was used to assemble four identical copies of this helix into a four-helix bundle, one helix at a time. At each step, ∼100 000 dockings were performed. The top 200 best-scoring structures were clustered by root-mean-square distance (RMSD) using cutoffs ranging from 0.75 to 3.0 Å, and representative dockings were examined visually. The top representative docked pose that positioned the newly introduced helix in the expected parallel conformation was retained. When no such pose was identified by clustering, all docked poses were examined sequentially, starting with the pose that had the highest score, until an acceptable pose was identified. The resulting four-helix bundle was then processed with Schrödinger Maestro's Protein Preparation Wizard and subjected to repeated minimization in the Maestro environment.

Each monomer of the NA$_{wt}$ model contained two cysteine residues. We assumed that the cysteine residues located in the

extra-virion region of the stalk likely participate in interchain disulfide bonds,[38] but that the cysteine residues located in the transmembrane region likely do not. The appropriate bonds were built in Maestro. The tetrameric-head and stalk/transmembrane-region models were next positioned near each other using VMD. The combined models were subsequently loaded into Maestro and again subjected to multiple minimizations.

Finally, Maestro was used to build and position the residues of the short intravirion topological domain. Based on the assumption that this domain lies at the interface of the inner membrane leaf and bulk intravirion water, all hydrophobic side chains were manually pointed upward toward the bilayer, and all polar side chains were pointed downward.

To generate a model of apo NA with a virulence-enhancing stalk deletion (hereafter called NA_{del}), the appropriate mutant amino acid sequence was first identified.[24] This sequence suggested that the stalk residues Cys49 to Asn68 should be deleted. These residues were manually removed from the NA_{wt} model, and the flanking NA segments were juxtaposed and fused using VMD and Schrödinger's Maestro, followed by multiple minimizations as needed.

Constructing the Bilayer. Our efforts to construct a complex-mixture lipid-membrane model were guided by a recent paper describing the composition of the influenza viral envelope.[39] We identified and classified all listed membrane components with mol % > 1.5, taking care to keep the composition of the inner and outer leaflets separate. As our ultimate goal was to perform molecular dynamics simulations of this system, we were limited to the lipids of the CHARMM 36 force field.[40,41] Consequently, we had to substitute some experimentally identified lipids with available lipid models. 3-palmitoyl-2-oleoyl-D-glycero-1-phosphatidylethanolamine (POPE) was used for any phosphatidylethanolamine (PE) variant, 2,3-distearoyl-D-glycero-1-phosphatidylserine (DSPS) for phosphatidylserine (PS), 3-palmitoyl-2-oleoyl-D-glycero-1-phosphatidylglycerol (POPG) for the Forssman glycolipid hapten, and 3-palmitoyl-2-oleoyl-D-glycero-1-phosphatidylcholine (POPC) for sphingomyelin (SM). Cholesterol and phosphatidylcholine, also major components of the envelope membrane, are included in the CHARMM 36 force field.[40] Ultimately, an appropriate planar bilayer model was built using CHARMM-GUI.[42,43] The NA models were then positioned within this generated bilayer, and clashing lipid residues were deleted with PyMolecule.[44]

In preparation for molecular dynamics simulations, we resolved multiple steric clashes in both the long-stalk and stalk-deletion systems using serial iterations of minimization and geometry optimization in Schrödinger's Maestro, coupled with manual atomic and molecular manipulation in VMD.[29]

Solvation and Parameterization. Partial charges were assigned to all atoms according to the specifications of the CHARMM22[45,46] all-hydrogen force field for proteins and the CHARMM36[40,41] all-hydrogen force field for lipids. The VMD[29] plug-in *cionize*[47] was used to position sodium and chloride ions as required to bring the systems to electrical neutrality and to simulate a 20 nM solution. Limited manual adjustments to the positions of some ions that had been placed far from any protein or lipid molecule were made.

For both the NA_{wt} and NA_{del} systems, Amber's XLEAP module[48] was used to generate a water box extending 15 Å beyond any atom. Water molecules were subsequently removed if they were not positioned directly above the bilayer, so that

the 15 Å margin was ultimately retained only along the Z axis, perpendicular to the bilayer itself. The NA_{wt} and NA_{del} systems had 366 293 and 351 699 atoms, respectively.

Molecular Dynamics Simulations. PSFGEN was used to fully parametrize each system according to the CHARMM22,[45,46] CHARMM36,[40] and TIP3P[49] force fields for proteins, lipids, and water molecules, respectively. NAMD 2.9[50] running on the Stampede supercomputer at the Texas Advanced Computing Center was used for all MD simulations. Periodic boundary conditions were employed with the particle mesh Ewald method to account for electrostatic effects (smoothing cutoff: 14 Å). Because of observed instabilities when our standard minimization protocols were used, we subjected the initial NA_{wt} system to only 1000 steps of unrestrained conjugate gradient minimization, with the first and max initial steps for the line minimizer reduced to 1.0×10^{-20} and 1.0×10^{-10}, respectively. Minimization of the NA_{del} system was more straightforward. The structure was minimized in four distinct steps: hydrogen atoms were first relaxed for 5 000 steps; hydrogen atoms, water molecules, and ions were next relaxed for 5 000 steps; hydrogen atoms, water molecules, ions, protein side chains, and lipid tails were then relaxed for 10 000 steps; and, finally, all atoms were relaxed for 25 000 steps.

Following minimization, each system was subjected to a constrained equilibration using an *NPT* ensemble at 310 K. The equilibration was realized in four steps, using stepwise harmonic-constraint force constants of 4, 3, 2, and 1 kcal/mol/Å² on the protein backbone. For each force constant (1 fs time step), 250 000 steps of MD simulation were performed. Langevin dynamics were applied to maintain the temperature, and a modified Langevin piston Nosé−Hoover thermostat was used to maintain 1 atm pressure.

Each system was next subjected to an unconstrained equilibration. The NA_{wt} model underwent an additional ∼5 million 2 fs steps of unconstrained simulation, followed by another ∼21 million 1 fs steps. The NA_{del} model underwent an additional ∼5.5 million 2 fs steps of unconstrained simulation, followed by another 36.5 million 1 fs steps.

Satisfied that both systems were sufficiently equilibrated, five distinct 100 ns productive runs were performed for each (∼100 million 1 fs steps) using distinct random seeds in order to sample diverse protein configurations. Snapshots were saved every 50 ps for subsequent analysis. (5 simulations) × (100 ns per simulation) × (2 systems) × (4 monomers per system) = 4 μs of monomeric simulation total. 5000 and 7000 frames were extracted from the last 50% and 70% of the tetrameric NA_{wt} and NA_{del} simulations, respectively, for subsequent analysis.

For comparison purposes, we also analyzed an archived simulation of the isolated apo NA head (hereafter called NA_{head}), described previously.[14] In brief, the NA head (PDB ID: 3NSS[26]) was simulated for 100 ns using the AMBER99SB force field, yielding 400 ns of monomeric trajectory. For the current study, we extracted 5000 frames from this simulation for subsequent analysis.

Root-Mean-Square-Distance Analysis of Binding-Pocket Residues. We used root-mean-square-distance (RMSD) calculations to judge the geometric similarity between the primary NA enzymatic pockets of crystallographic holo and simulated apo models. We first identified active-site residues as those that came within 5.0 Å of a crystallographic oseltamivir ligand (PDB ID: 2HU4:A, H5N1). These residues are Arg118, Glu119, Asp151, Arg152, Arg156, Trp178, Ser179, Ile222, Arg224, Glu227, Ser246, Glu276, Glu277, Arg292, Asn294,

Tyr347, Gly348, Arg371, and Tyr406. For the purposes of our RMSD analysis, we discarded Tyr347 because that tyrosine is an asparagine in H1N1 NA. Hydrogen atoms were also removed prior to performing all RMSD calculations; RMSD alignment was performed only on the remaining heavy atoms.

Principal Component Analysis. We performed principal component analysis (PCA) on these same active-site heavy atoms. After RMSD aligning the NA_{wt} trajectories by these atoms using VMD, we used the python modules scikit-learn[51] and numpy[52−54] to calculate their first two principal components and to project each trajectory frame onto those components. The frames of the NA_{del} trajectory were similarly projected onto the same two NA_{wt} principal components to facilitate comparison.

Volumetric Analysis. The POVME algorithm[55] was used to measure the volume of the primary sialic-acid-binding pocket over the course of the NA_{wt}, NA_{del}, and NA_{head} simulations. To calculate the volume of the binding pocket for a given trajectory frame, we monitored the space within a pocket-centered sphere of radius 16.0 Å. In all cases, the pocket volume was calculated from the portion of the sphere that was unoccupied by protein atoms. Only volumetric regions contiguous with the pocket were included. Grid spacing and padding parameters were set to 1.0 and 1.09 Å, respectively.

Root-Mean-Square Fluctuations. All-atom, residue root-mean-square fluctuations (RMSF)[56] were calculated for each monomer of each simulation. For the NA_{wt} and NA_{del} simulations, there were 20 associated monomeric simulations (5 runs × 4 monomers). For the NA_{head} simulation, there were 4 associated monomeric simulations (1 run × 4 monomers). In all cases, the monomeric trajectories were first aligned by the α carbons of residues Ser90 to Asp469 (i.e., the residues of the NA head). The residue center-of-mass RMSF values were calculated for each monomeric simulation separately.

Identifying Druggable Hotspots. To identify druggable hotspots, we first identified representative protein conformations from the MD simulations. The heads of the aligned monomeric trajectories were subjected to RMSD clustering,[57] as implemented in the GROMACS computer package.[58] We adjusted the RMSD cutoff until the trajectory frames were separated into five clusters. The required cutoffs were 1.72, 1.74, and 1.74 Å for the NA_{wt}, NA_{del}, and NA_{head} simulations, respectively. The centroid members of each of these clusters were considered representative of the corresponding simulated model.

The 15 representative NA structures (3 models × 5 cluster centroids) were then submitted to the FTMAP server[59,60] to identify druggable hotspots. To simplify the analysis, all the docked molecular probes associated with each simulation (NA_{wt}, NA_{del}, and NA_{head}) were considered collectively. To identify druggable hotspots that were specific to a given model, the associated docked probes were superimposed, and any probe atom associated with one model that came within 5 Å of any probe atom associated with the other was deleted. The remaining probe atoms congregated at hotspots that were unique to their respective models.

■ RESULTS AND DISCUSSION

The "Limited-Access" Theory. As described in the Introduction, the limited-access theory holds that stalk-deletion neuraminidase glyocproteins (NA_{del}) have reduced sialidase activity relative to NA_{wt} because their diminished height hinders access to cellular sialylated-oligosaccharide receptors. The fact that NA enzymatic activity varies depending on ligand bulkiness is perhaps the most compelling evidence in favor of this theory. Studies have shown that short-stalk NA_{del} glycoproteins have significantly reduced activity when a bulky substrate (e.g., fetuin) is used,[61] presumably because the large substrate cannot physically reach the NA enzymatic sites. In contrast, NA activity is largely stalk-independent when cleaving the small molecule MUNANA (essentially a single sialic acid with a 4-methylcoumarin fluorophore attached to the 2-carbon),[62,63] apparently because this smaller substrate is not subject to the same large-scale steric hindrance.

However, several lines of evidence suggest that the limited-access theory may not fully explain the reduced NA_{del} activity. First, MUNANA is very different than the endogenous sialylated-oligosaccharide receptor. Binding to these receptors, which are extended, large, and cell bound, is likely diffusion limited; in contrast, MUNANA is a relatively small molecule with an affinity that is likely little dependent on k_{on}. While MUNANA may be a fine substrate for measuring relative NA enzymatic activity in the context of a small-molecule screen, it is arguably an inadequate surrogate for the actual receptor.

Second, fetuin (α-2-HS-glycoprotein) is also very different than the endogenous receptor, and the theory that its bulk alone prevents it from accessing the NA_{del} enzymatic sites is hardly certain. We note that all available monoclonal antibodies known to target NA antigenic sites bind equally well to virions with NA_{del} and NA_{wt} glycoproteins,[61] despite the fact that antibodies are much larger than fetuin (∼160 and ∼1000 kDa for IgG and IgM, respectively, vs ∼40 kDa for fetuin).[61] This evidence suggests that access to NA_{del} is not terribly hindered. Furthermore, recent cryo-electron microscopy studies of whole-virion particles demonstrate that the NA glycoproteins are not evenly interspersed among the more prevalent HA's as the limited-access theory suggests; rather, the NA particles tend to cluster together on the viral surface.[64] Even when stalk mutations substantially reduce the NA height relative to HA, this clustering effect alone should in theory preserve much access to the NA enzymatic head.
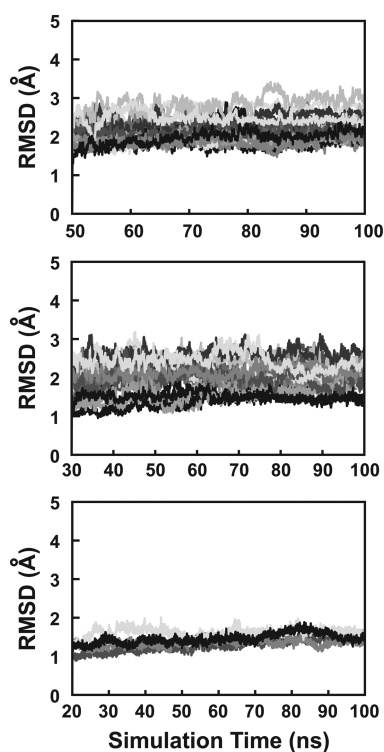
Computational evidence likewise suggests that simple steric considerations are not solely responsible for reduced NA_{del} activity. Our models of 2009 H1N1 glycoproteins demonstrate that each stalk amino acid contributes only ∼1.2 Å to the total NA height (data not shown). The 2013 H7N9 five-amino-acid stalk deletion with enhanced virulence mentioned in the Introduction therefore reduced the NA height by only ∼6 Å (roughly 4%).[13] More substantial stalk deletions do not necessarily force the NA enzymatic head to retreat entirely behind a towering canopy of HA glycoproteins either. In our models of H1N1 HA and NA, we found the height of NA_{wt} was nearly equal to that of HA (149 vs 154 Å, respectively). When a virulence-enhancing 20 amino acid NA stalk deletion[24] was modeled (NA_{del}), the height of the NA fell to only 125 Å, still 81% of the HA height. Given that the NA head is itself a boxlike structure 80 Å square, the resulting 24 Å height reduction seems unlikely to substantially interfere with access.

We suggest that, beyond limited-access effects, actual differences in the affinities of NA_{wt} and NA_{del} for the endogenous sialylated-oligosaccharide receptors contribute to the differences in enzymatic activity. As support for our reduced-affinity hypothesis, we analyze atomistic molecular dynamics simulations of three NA models: five 100 ns simulations of 2009 H1N1 NA_{wt}; five 100 ns simulations of 2009 H1N1 NA with a 20-amino-acid deletion in the stalk

(NA$_{del}$[24]); and one 100 ns simulation of the isolated H1N1 NA head, NA$_{head}$ (Figure 1). These simulations indicate that stalk length does in fact impact the dynamics of the enzymatic sialic acid binding pocket, suggesting a potential mechanism for the proposed stalk-length-dependent differences in affinity.

Whole-Glycoprotein Root-Mean-Square-Distance Analysis. To verify that the simulations of the three NA systems (i.e., NA$_{wt}$, NA$_{del}$, and NA$_{head}$) had properly equilibrated, we RMSD aligned each tetrameric trajectory by its C$_\alpha$ atoms using VMD.[29] The RMSD was then calculated between each trajectory frame and a corresponding reference structure (Figure S1). Despite extensive pre-equilibration, the RMSD values of the whole-glycoprotein NA$_{wt}$ and NA$_{del}$ simulations continued to drift slightly throughout the productive runs.

To further investigate the source of this drift, we calculated separate C$_\alpha$ RMSD values for the residues of the NA head (Figure 2), the lower (membrane-embedded) stalk, and the

of this region equilibrated. Given the instability of the upper stalk, however, we chose to discard the initial 50% and 30% of the NA$_{wt}$ and NA$_{del}$ trajectories, respectively, to ensure that subsequent analysis focused on the most equilibrated portion of the simulations. In contrast, NA$_{head}$, based on the 3NSS crystal structure,[26] equilibrated far faster. We therefore discarded only the first 20% of the NA$_{head}$ simulation (Figures 2 and S1).

RMSD Analysis: Judging Binding-Pocket Dynamics. We next sought to characterize the effects of the stalk on binding-pocket dynamics by comparing the geometries of our simulated apo binding sites to that of holo (ligand-bound) NA. We calculated the RMSD values between the simulated binding-pocket conformations and that of a reference crystallographic structure, 2HU4[27] (H5N1), which shares in common with H1N1 a total of 18 out of 19 binding-pocket residues. In calculating these RMSD values, the single differing residue was ignored.

A histogram of the RMSD values suggests that the long-stalk NA$_{wt}$ pocket exists in two conformational states (Figure 3a), with RMSDs to the apo (reference) conformation of roughly 1.8 and 2.7 Å, respectively. In contrast, the pocket conformations sampled during the NA$_{del}$ and NA$_{head}$ simulations were not binary, suggesting the pocket is less flexible in these systems.

Figure 2. RMSDs of all trajectory monomers to their respective reference structures. In all cases, only the C$_\alpha$ of the head domain (i.e., the C$_\alpha$ of all residues homologous to the 388 residues present in the 3NSS crystal structure) were considered. (A) NA$_{wt}$ (5 tetrameric trajectories × 4 monomers per tetramer = 20 runs); (B) NA$_{del}$ (5 tetrameric trajectories × 4 monomers per tetramer = 20 runs); and (C) NA$_{head}$ (1 tetrameric trajectory × 4 monomers per tetramer = 4 runs).

upper stalk. The dynamics of the head and lower stalk were in fact stable in both the NA$_{wt}$ and NA$_{del}$ simulations. It was the upper stalk that was dynamically unstable. Given that the upper stalk serves as a linker connecting two fairly ridged bodies (the lower stalk and the head), it is reasonable to suppose that this region is legitimately hypermobile.
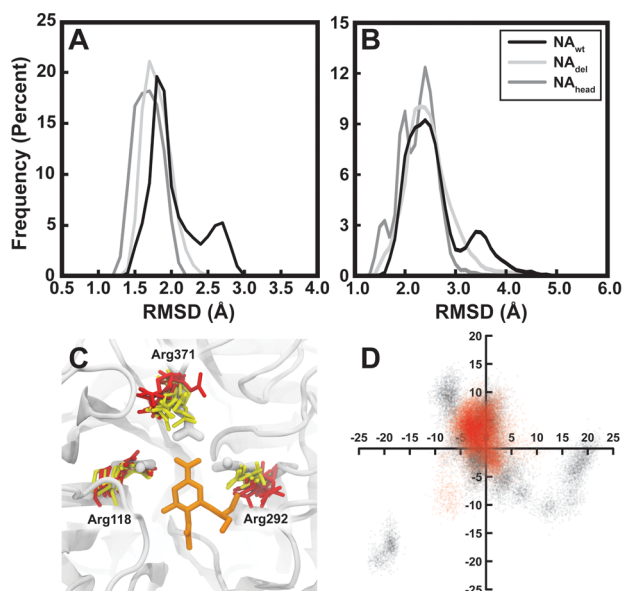
Subsequent analyses (described below) focused primarily on the residues of the NA binding pocket, located on the enzymatic head. It is therefore encouraging that the dynamics

Figure 3. Dynamics of the sialic-acid-binding pocket. (A) Simulated apo binding-pocket conformations are compared to the crystallographic (2HU4) holo conformation. Geometric similarity was judged by calculating the RMSDs between simulated and crystallographic pocket residues. The NA$_{wt}$ pocket assumes two distinct conformational states. (B) Dynamics of the arginine triad. Geometric similarity was judged by calculating the RMSDs between simulated and crystallographic triad conformations. Note that the NA$_{wt}$ simulation again assumes two conformational states. (C) Representative NA$_{wt}$ arginine-triad conformations sampled from the 2.4 Å state (in yellow) and the 3.4 Å state (in red). The crystallographic position of the triad residues is shown in thick, white licorice. An oseltamivir molecule has been positioned within the binding pocket for reference (in orange), though the simulations did not include any ligand. (D) The simulated conformations of NA$_{wt}$ and NA$_{del}$ active-site heavy atoms, in black and red, respectively, projected onto the first two principal components of the NA$_{wt}$ simulation. The NA$_{wt}$ site assumes a greater variety of conformations, in harmony with the RMSD analysis.

An arginine triad formed by Arg118, Arg292, and Arg371 is a key feature of the enzymatic pocket. This triad forms a positively charged arginine nest that surrounds the negatively charged carboxylate groups of the endogenous substrate as well as all FDA-approved NA inhibitors. As with the whole pocket, the dynamics of this key triad are similarly dependent on stalk length. The average RMSD between the atom positions of the simulated apo triad configurations and the holo 2HU4[27] conformation were 2.5 and 2.4 Å for the NA_{wt} and NA_{del} simulations, respectively. A t-test led us to reject the null hypothesis that these two averages are statistically equal ($p = 0.0$).

A histogram of the arginine-triad RMSD values suggests that, unlike the NA_{del} triad, the NA_{wt} triad again assumes two distinct conformations (Figure 3b). To further characterize these two NA_{wt} states, we visually inspected representative conformations from each (Figure 3c). In the state with RMSD = ∼ 2.4 Å, Arg371 is near its crystallographic conformation, and Arg292 is displaced only slightly in the direction opposite of Arg371. In contrast, in the state with RMSD = ∼3.4 Å, Arg371 is generally displaced in the direction of Arg118, and Arg292 is even further from Arg371. Arg118 possesses roughly the same conformation in both states.

To further verify that the NA_{wt} pocket assumes a greater variety of conformational states than the NA_{del} pocket, we next performed a principal component analysis (PCA) of the NA_{wt} active-site heavy atoms (Figure 3d). Pocket conformations from both the NA_{wt} and NA_{del} simulations were projected onto the first two principal components. The NA_{wt} pocket (in black) sampled the same region of PCA space as the NA_{wt} simulation (in red), but the NA_{wt} pocket additionally sampled other conformations not seen in the NA_{del} simulation, confirming that the NA_{wt} pockets adopted a greater number of conformational states.

Stalk-Length Effects on the Movements of Specific Binding-Pocket Residues. Having studied the impact of stalk length on binding-pocket dynamics generally, we next considered the dynamics of specific binding-pocket residues. We calculated the associated all-atom, residue root-mean-square fluctuations[56] for each of the 20 monomeric trajectories associated with the NA_{wt} and NA_{del} simulations. For each residue, we compared the ensembles of NA_{wt} and NA_{del} RMSF values using a t-test. A number of residues exhibited statistically significant average RMSF differences when a liberal significance level was used (t-test, $\alpha = 0.10$). Most interestingly, the RMSF values of two of the three arginine residues of the key active-site triad (Arg371 and Arg118) differed significantly (Table 1). This analysis again implicates the arginine triad and 371 loop as contributing to differential, stalk-length-dependent binding-pocket dynamics.

To better visualize differences in active-site dynamics, we colored the residues of a NA_{wt} conformation by $\Delta RMSF = RMSF_{wt} - RMSF_{del}$ (Figure 4). The 371, 406, and 430 loops were generally more flexible in the NA_{wt} simulation (shown in red). However, a portion of the 150 loop was more flexible in the NA_{del} simulation (shown in blue).

Stalk Length May Subtly Alter Active-Site Druggability. Finally, we also used our molecular dynamics simulations to study NA druggability. To get an initial impression of the full extent of pocket dynamics, we first performed a volumetric analysis of the primary sialic-acid-binding pocket using the POVME algorithm.[55] The trajectory frames with both the largest and smallest volumes from the

Table 1. Binding-Pocket Residues with RMSF Values That Differed Significantly between the NA_{wt} and NA_{del} Simulations ($p < 0.10$)[a]

|	NA_{wt}	NA_{del}	p value	NA_{head}
Arg118 (triad)	0.65 ± 0.09	0.75 ± 0.21	0.06	0.95 ± 0.43
Glu119	0.67 ± 0.17	0.74 ± 0.10	0.10	1.02 ± 0.23
Ile222	0.82 ± 0.05	0.85 ± 0.06	0.09	0.84 ± 0.02
Ser246	1.24 ± 0.28	1.40 ± 0.23	0.06	1.30 ± 0.24
Glu277	0.58 ± 0.20	0.67 ± 0.13	0.07	0.68 ± 0.18
Asn347	1.04 ± 0.11	1.14 ± 0.19	0.06	0.83 ± 0.07
Arg371 (371 loop, triad)	1.24 ± 0.37	1.44 ± 0.32	0.09	0.98 ± 0.21

[a]Data from the corresponding NA_{head} (stalk-absent) simulations are also included. The average RMSF value is shown plus or minus the standard deviation. The residue-specific populations of all NA_{wt} and NA_{del} RMSF values were compared using t-tests.
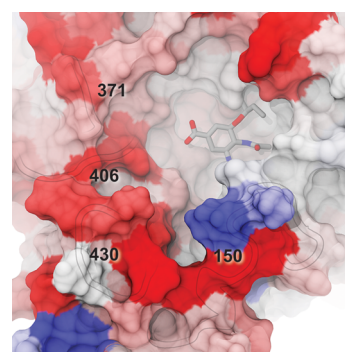
Figure 4. NA active-site residues colored by $\Delta RMSF = RMSF_{wt} - RMSF_{del}$, from −0.2 to 0.2. Residues in red and blue were more flexible in the NA_{wt} and NA_{del} simulations, respectively. The backbones of key binding-pocket loops are subtly outlined. A crystallographic oseltamivir molecule is shown for reference, though no ligand was included in the simulations.

NA_{wt}, NA_{del}, and NA_{head} simulations were compared separately (Figure 5). In all cases, the pocket fluctuated between a closed conformation (Figure 5a) and an open conformation (Figure 5b), suggesting a highly dynamic pharmacophore. Interestingly, the NA_{wt} pocket tended to be larger than the NA_{del} pocket (Figure 5c, $p = 0.0$ per a t-test), implying possible differences in the pharmacophore distributions of these two NA variants that may have important implications for drug discovery. To further study variant-dependent differences in NA-pocket druggability. FTMAP seeks to replicate in silico the experimental multiple solvent crystal structures (MSCS) method developed by Mattos et al.[65,66] In brief, the FTMAP algorithm uses computer docking to flood the surface of a protein with small organic molecular probes. These probes are then clustered to identify regions of the protein surface that are most likely to bind small molecules.

We used FTMAP to identify druggable hotspots on the surfaces of five representative NA structures extracted from each of our molecular dynamics simulations. For the most part, the druggable regions of the enzymatic site did not differ substantially between the NA_{wt} and NA_{del} simulations. However, FTMAP did position some molecular probes near a 430-loop-adjacent region of the NA_{wt} pocket that was not deemed "druggable" when the NA_{del} conformations were considered, suggesting subtle differences that may be pharmacologically relevant (Figure 6a).
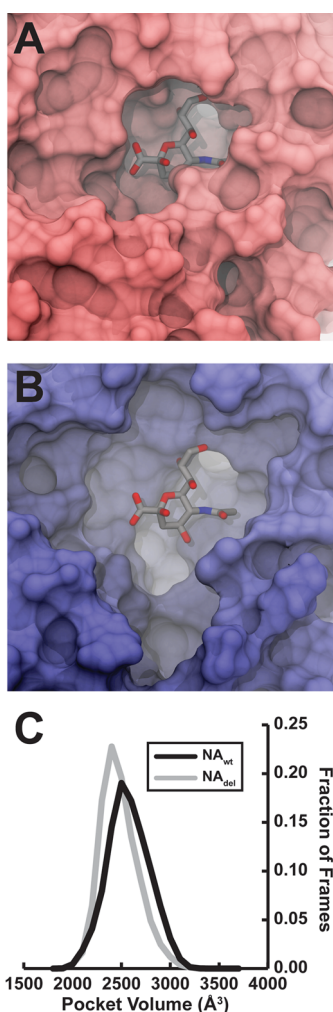
Figure 5. Volumetric analysis. (A) The smallest and (B) largest binding-pocket conformations sampled by the NA_{wt} simulation. The smallest and largest conformations from the NA_{del} and NA_{head} simulations were similar (data not shown). Sialic acid molecules (in licorice) have been placed in each pocket for reference, though the simulations included no ligand. () Histograms of the volumes sampled over the course of the NA_{wt} and NA_{del} simulations, in black and gray, respectively. The NA_{wt} pocket tended to be larger ($p = 0.0$).

The NA_{wt} and NA_{del} hotspots were then merged and similarly compared to the NA_{head} hotspots. Surprisingly, in both stalk-inclusive NA simulations (NA_{wt} and NA_{del}), the 371 loop occasionally flipped upward, opening a novel druggable hotspot that was not apparent when the representative structures of the NA_{head} simulation were considered (Figure 6b). This pocket has been described previously,[19,67] but the NA_{wt} and NA_{del} simulations confirm its importance. To date, no known NA inhibitor exploits this pocket, and it has not been the focus of previous CADD efforts. Finally, another large druggable hotspot located at the base of the NA head was unique to the stalk-inclusive (i.e., NA_{wt} and NA_{del}) simulations (Figure 6c).

◼ CONCLUSION

Factors that alter the sialic acid binding/association mechanism may profoundly alter NA affinity and activity in additional ways not explained by the "theory of limited access." The length of the stalk appears to influence the dynamics of the 150, 430, 371, and 406 binding-pocket loops. The 371 loop and the critical arginine catalytic triad are particularly affected.

As a consequence, the apo NA_{wt} pocket may adopt a greater variety of conformational states than does that of NA_{del}. Specifically, the NA_{del} pocket tends to sample a single conformational state that is generally similar to that of the holo conformation. In contrast, NA_{wt} additionally samples more distant pocket geometries (Figure 3).

While the impact of stalk length on binding-pocket dynamics may well be partially responsible for the reduced NA_{del} affinity, we note also that stalk length likely has a profound influence on the long-range electrostatics of the whole-virion surface, which could also impact k_{on}. Future efforts will focus on elucidating stalk-length-dependent differences in this effect as well.

Finally, the predicted druggable hotspots evident in the stalk-inclusive simulations also suggest novel and urgently needed opportunities for drug discovery. Prior to the 2009 H1N1 pandemic, seasonal H1N1 in the United States was almost entirely oseltamivir resistant.[68] Fortunately, that pandemic strain itself remained largely susceptible to oseltamivir,[69] but given the viral proclivity for rapid mutation and adaptation, multidrug combinatorial therapies may be necessary in the future.[68]

These novel hotspots also provide means for experimentally testing some of the hypotheses generated in the current study. Virtual screening into the NA pocket conformations sampled over the course of the NA_{wt} and NA_{del} simulations could identify compounds that are uniquely suited to the corresponding pocket geometries. If subsequent experimental validation confirms binding, X-ray crystallography could be used to
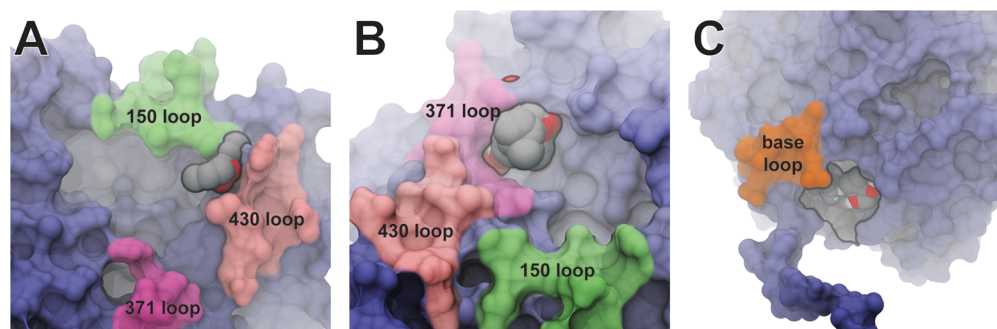
Figure 6. Unique druggable hotspots (outlined). The 371, 430, and 150 loops are shown in mauve, pink, and green, respectively. (A) A druggable hotspot was identified near the 430 loop of the NA_{wt} pocket. (B) A second druggable hotspot was identified beneath the 371 loop of both the NA_{wt} and NA_{del} pockets. (C) A large druggable hotspot was identified near the base of the NA_{wt} and NA_{del} heads.

determine if the novel ligand(s) stabilize one of the predicted NA conformations.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.6b02655.

Graphs that show the tetrameric RMSDs of all trajectory frames to their respective reference structures (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Department of Chemistry & Biochemistry, University of California San Diego, 9500 Gilman Drive, Mail Code 0365, La Jolla, CA 92093-0365. Tel: 858.534.9629. Fax:858.534.9645. E-mail: jdurrant@ucsd.edu.

Notes

The authors declare the following competing financial interest(s): R.E.A. is a co-founder of Actavalon, Inc.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Cox, N. J.; Neumann, G.; Donis, R. O.; Kawaoka, Y. Orthomyxoviruses: Influenza. In *Topley & Wilson's Microbiology and Microbial Infections*; Mahy, B. W. J.; Collier, L., Eds.; Wiley-Blackwell: London, 2009.

(2) Li, J. L.; Dohna, H. Z.; Cardona, C. J.; Miller, J.; Carpenter, T. E. Emergence and Genetic Variation of Neuraminidase Stalk Deletions in Avian Influenza Viruses. *PLoS One* **2011**, *6*, e14722.

(3) Tong, S. X.; Zhu, X. Y.; Li, Y.; Shi, M.; Zhang, J.; Bourgeois, M.; Yang, H.; Chen, X. F.; Recuenco, S.; Gomez, J.; et al. New World Bats Harbor Diverse Influenza a Viruses. *PLoS Pathog.* **2013**, *9*, e1003657.

(4) Baigent, S. J.; McCauley, J. W. Glycosylation of Haemagglutinin and Stalk-Length of Neuraminidase Combine to Regulate the Growth of Avian Influenza Viruses in Tissue Culture. *Virus Res.* **2001**, *79*, 177–185.

(5) Banks, J.; Speidel, E. S.; Moore, E.; Plowright, L.; Piccirillo, A.; Capua, I.; Cordioli, P.; Fioretti, A.; Alexander, D. J. Changes in the Haemagglutinin and the Neuraminidase Genes Prior to the Emergence of Highly Pathogenic H7n1 Avian Influenza Viruses in Italy. *Arch. Virol.* **2001**, *146*, 963–973.

(6) Matrosovich, M.; Zhou, N.; Kawaoka, Y.; Webster, R. The Surface Glycoproteins of H5 Influenza Viruses Isolated from Humans, Chickens, and Wild Aquatic Birds Have Distinguishable Properties. *J. Virol.* **1999**, *73*, 1146–1155.

(7) Mitnaul, L. J.; Matrosovich, M. N.; Castrucci, M. R.; Tuzikov, A. B.; Bovin, N. V.; Kobasa, D.; Kawaoka, Y. Balanced Hemagglutinin and Neuraminidase Activities Are Critical for Efficient Replication of Influenza a Virus. *J. Virol.* **2000**, *74*, 6015–6020.

(8) Baigent, S. J.; McCauley, J. W. Influenza Type a in Humans, Mammals and Birds: Determinants of Virus Virulence, Host-Range and Interspecies Transmission. *BioEssays* **2003**, *25*, 657–671.

(9) Wagner, R.; Matrosovich, M.; Klenk, H. D. Functional Balance between Haemagglutinin and Neuraminidase in Influenza Virus Infections. *Rev. Med. Virol.* **2002**, *12*, 159–166.

(10) Castrucci, M. R.; Kawaoka, Y. Biologic Importance of Neuraminidase Stalk Length in Influenza a Virus. *J. Virol.* **1993**, *67*, 759–764.

(11) Guangxiang, L.; Jeffrey, C.; Palese, P. Alterations of the Stalk of the Influenza Virus Neuraminidase: Deletions and Insertions. *Virus Res.* **1993**, *29*, 141–153.

(12) Matsuoka, Y.; Swayne, D. E.; Thomas, C.; Rameix-Welti, M. A.; Naffakh, N.; Warnes, C.; Altholtz, M.; Donis, R.; Subbarao, K. Neuraminidase Stalk Length and Additional Glycosylation of the Hemagglutinin Influence the Virulence of Influenza H5n1 Viruses for Mice. *J. Virol.* **2009**, *83*, 4704–4708.

(13) Gao, R. B.; Cao, B.; Hu, Y. W.; Feng, Z. J.; Wang, D. Y.; Hu, W. F.; Chen, J.; Jie, Z. J.; Qiu, H. B.; Xu, K.; et al. Human Infection with a Novel Avian-Origin Influenza a (H7n9) Virus. *N. Engl. J. Med.* **2013**, *368*, 1888–1897.

(14) Amaro, R. E.; Swift, R. V.; Votapka, L.; Li, W. W.; Walker, R. C.; Bush, R. M. Mechanism of 150-Cavity Formation in Influenza Neuraminidase. *Nat. Commun.* **2011**, *2*, 388.

(15) Wu, Y.; Qin, G. R.; Gao, F.; Liu, Y.; Vavricka, C. J.; Qi, J. X.; Jiang, H. L.; Yu, K. Q.; Gao, G. F. Induced Opening of Influenza Virus Neuraminidase N2 150-Loop Suggests an Important Role in Inhibitor Binding. *Sci. Rep.* **2013**, *3*, 1551.

(16) Han, N. Y.; Mu, Y. G. Plasticity of 150-Loop in Influenza Neuraminidase Explored by Hamiltonian Replica Exchange Molecular Dynamics Simulations. *PLoS One* **2013**, *8*, e60995.

(17) Zhu, X.; Yang, H.; Guo, Z.; Yu, W.; Carney, P. J.; Li, Y.; Chen, L. M.; Paulson, J. C.; Donis, R. O.; Tong, S.; et al. Crystal Structures of Two Subtype N10 Neuraminidase-Like Proteins from Bat Influenza a Viruses Reveal a Diverged Putative Active Site. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 18903–18908.

(18) Xu, X.; Zhu, X.; Dwek, R. A.; Stevens, J.; Wilson, I. A. Structural Characterization of the 1918 Influenza Virus H1n1 Neuraminidase. *J. Virol.* **2008**, *82*, 10493–10501.

(19) Amaro, R. E.; Minh, D. D.; Cheng, L. S.; Lindstrom, W. M., Jr.; Olson, A. J.; Lin, J. H.; Li, W. W.; McCammon, J. A. Remarkable Loop Flexibility in Avian Influenza N1 and Its Implications for Antiviral Drug Design. *J. Am. Chem. Soc.* **2007**, *129*, 7764–7765.

(20) Ma, B.; Kumar, S.; Tsai, C. J.; Nussinov, R. Folding Funnels and Binding Mechanisms. *Protein Eng., Des. Sel.* **1999**, *12*, 713–720.

(21) Kumar, S.; Ma, B.; Tsai, C. J.; Wolfson, H.; Nussinov, R. Folding Funnels and Conformational Transitions Via Hinge-Bending Motions. *Cell Biochem. Biophys.* **1999**, *31*, 141–164.

(22) Tsai, C. J.; Kumar, S.; Ma, B.; Nussinov, R. Folding Funnels, Binding Funnels, and Protein Function. *Protein Sci.* **1999**, *8*, 1181–1190.

(23) Ma, B.; Shatsky, M.; Wolfson, H. J.; Nussinov, R. Multiple Diverse Ligands Binding at a Single Protein Site: A Matter of Pre-Existing Populations. *Protein Sci.* **2002**, *11*, 184–197.

(24) Zhou, H. B.; Yu, Z. J.; Hu, Y.; Tu, J. G.; Zou, W.; Peng, Y. P.; Zhu, J. P.; Li, Y. T.; Zhang, A. D.; Yu, Z. N.; et al. The Special Neuraminidase Stalk-Motif Responsible for Increased Virulence and Pathogenesis of H5n1 Influenza a Virus. *PLoS One* **2009**, *4*, e6277.

(25) Blumenkrantz, D.; Roberts, K. L.; Shelton, H.; Lycett, S.; Barclay, W. S. The Short Stalk Length of Highly Pathogenic Avian Influenza H5n1 Virus Neuraminidase Limits Transmission of Pandemic H1n1 Virus in Ferrets. *J. Virol.* **2013**, *87*, 10539–10551.

(26) Li, Q.; Qi, J.; Zhang, W.; Vavricka, C. J.; Shi, Y.; Wei, J.; Feng, E.; Shen, J.; Chen, J.; Liu, D.; et al. The 2009 Pandemic H1n1 Neuraminidase N1 Lacks the 150-Cavity in Its Active Site. *Nat. Struct. Mol. Biol.* **2010**, *17*, 1266–1268.

(27) Russell, R. J.; Haire, L. F.; Stevens, D. J.; Collins, P. J.; Lin, Y. P.; Blackburn, G. M.; Hay, A. J.; Gamblin, S. J.; Skehel, J. J. The Structure of H5n1 Avian Influenza Neuraminidase Suggests New Opportunities for Drug Design. *Nature* **2006**, *443*, 45–49.

(28) Roberts, E.; Eargle, J.; Wright, D.; Luthey-Schulten, Z. Multiseq: Unifying Sequence and Structure Data for Evolutionary Analysis. *BMC Bioinf.* **2006**, *7*, 382.

(29) Humphrey, W.; Dalke, A.; Schulten, K. Vmd: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.

(30) Collins, P. J.; Haire, L. F.; Lin, Y. P.; Liu, J. F.; Russell, R. J.; Walker, P. A.; Skehel, J. J.; Martin, S. R.; Hay, A. J.; Gamblin, S. J. Crystal Structures of Oseltamivir-Resistant Influenza Virus Neuraminidase Mutants. *Nature* **2008**, *453*, 1258−U1261.

(31) Pollastri, G.; Przybylski, D.; Rost, B.; Baldi, P. Improving the Prediction of Protein Secondary Structure in Three and Eight Classes Using Recurrent Neural Networks and Profiles. *Proteins: Struct., Funct., Genet.* **2002**, *47*, 228−235.

(32) Cheng, J.; Randall, A. Z.; Sweredoski, M. J.; Baldi, P. Scratch: A Protein Structure and Structural Feature Prediction Server. *Nucleic Acids Res.* **2005**, *33*, W72−W76.

(33) Chaudhury, S.; Gray, J. J. Conformer Selection and Induced Fit in Flexible Backbone Protein-Protein Docking Using Computational and Nmr Ensembles. *J. Mol. Biol.* **2008**, *381*, 1068−1087.

(34) Chaudhury, S.; Sircar, A.; Sivasubramanian, A.; Berrondo, M.; Gray, J. J. Incorporating Biochemical Information and Backbone Flexibility in Rosettadock for Capri Rounds 6−12. *Proteins: Struct., Funct., Genet.* **2007**, *69*, 793−800.

(35) Daily, M. D.; Masica, D.; Sivasubramanian, A.; Somarouthu, S.; Gray, J. J. Capri Rounds 3−5 Reveal Promising Successes and Future Challenges for Rosettadock. *Proteins: Struct., Funct., Genet.* **2005**, *60*, 181−186.

(36) Gray, J. J.; Moughon, S. E.; Kortemme, T.; Schueler-Furman, O.; Misura, K. M. S.; Morozov, A. V.; Baker, D. Protein-Protein Docking Predictions for the Capri Experiment. *Proteins: Struct., Funct., Genet.* **2003**, *52*, 118−122.

(37) Gray, J. J.; Moughon, S.; Wang, C.; Schueler-Furman, O.; Kuhlman, B.; Rohl, C. A.; Baker, D. Protein-Protein Docking with Simultaneous Optimization of Rigid-Body Displacement and Side-Chain Conformations. *J. Mol. Biol.* **2003**, *331*, 281−299.

(38) Blok, J.; Air, G. M. Variation in the Membrane-Insertion and Stalk Sequences in 8 Subtypes of Influenza Type-a Virus Neuraminidase. *Biochemistry* **1982**, *21*, 4001−4007.

(39) Gerl, M. J.; Sampaio, J. L.; Urban, S.; Kalvodova, L.; Verbavatz, J. M.; Binnington, B.; Lindemann, D.; Lingwood, C. A.; Shevchenko, A.; Schroeder, C.; et al. Quantitative Analysis of the Lipidomes of the Influenza Virus Envelope and Mdck Cell Apical Membrane. *J. Cell Biol.* **2012**, *196*, 213−221.

(40) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D., Jr.; Pastor, R. W. Update of the Charmm All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **2010**, *114*, 7830−7843.

(41) Klauda, J. B.; Monje, V.; Kim, T.; Im, W. Improving the Charmm Force Field for Polyunsaturated Fatty Acid Chains. *J. Phys. Chem. B* **2012**, *116*, 9424−9431.

(42) Jo, S.; Kim, T.; Im, W. Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. *PLoS One* **2007**, *2*, e880.

(43) Jo, S.; Lim, J. B.; Klauda, J. B.; Im, W. Charmm-Gui Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophys. J.* **2009**, *97*, 50−58.

(44) Durrant, J. D.; McCammon, J. A. Autoclickchem: Click Chemistry in Silico. *PLoS Comput. Biol.* **2012**, *8*, e1002397.

(45) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(46) Mackerell, A. D., Jr.; Feig, M.; Brooks, C. L., 3rd. Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. *J. Comput. Chem.* **2004**, *25*, 1400−1415.

(47) Stone, J. E.; Phillips, J. C.; Freddolino, P. L.; Hardy, D. J.; Trabuco, L. G.; Schulten, K. Accelerating Molecular Modeling Applications with Graphics Processors. *J. Comput. Chem.* **2007**, *28*, 2618−2640.

(48) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R.

J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.

(49) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(50) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with Namd. *J. Comput. Chem.* **2005**, *26*, 1781−1802.

(51) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.

(52) Ascher, D.; Dubois, P. F.; Hinsen, K.; James, J. H.; Oliphant, T. *Numerical Python*; UCRL-MA-128569 ed.; Lawrence Livermore National Laboratory: Livermore, CA, 1999.

(53) Dubois, P. F.; Yang, T.-Y. Extending Python with Fortran. *Comput. Sci. Eng.* **1999**, *1*, 66−73.

(54) Oliphant, T. E. *Guide to Numpy*; Brigham Young University: Provo, UT, 2006.

(55) Durrant, J. D.; de Oliveira, C. A.; McCammon, J. A. Povme: An Algorithm for Measuring Binding-Pocket Volumes. *J. Mol. Graphics Modell.* **2011**, *29*, 773−776.

(56) Roe, D. R.; Cheatham, T. E. Ptraj and Cpptraj: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.

(57) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chem., Int. Ed.* **1999**, *38*, 236−240.

(58) Christen, M.; Hunenberger, P. H.; Bakowies, D.; Baron, R.; Burgi, R.; Geerke, D. P.; Heinz, T. N.; Kastenholz, M. A.; Krautler, V.; Oostenbrink, C.; et al. The Gromos Software for Biomolecular Simulation: Gromos05. *J. Comput. Chem.* **2005**, *26*, 1719−1751.

(59) Brenke, R.; Kozakov, D.; Chuang, G. Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S. Fragment-Based Identification of Druggable 'Hot Spots' of Proteins Using Fourier Domain Correlation Techniques. *Bioinformatics* **2009**, *25*, 621−627.

(60) Kozakov, D.; Hall, D. R.; Chuang, G. Y.; Cencic, R.; Brenke, R.; Grove, L. E.; Beglov, D.; Pelletier, J.; Whitty, A.; Vajda, S. Structural Conservation of Druggable Hot Spots in Protein-Protein Interfaces. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13528−13533.

(61) Els, M. C.; Air, G. M.; Murti, K. G.; Webster, R. G.; Laver, W. G. An 18-Amino Acid Deletion in an Influenza Neuraminidase. *Virology* **1985**, *142*, 241−247.

(62) Hoffmann, T. W.; Munier, S.; Larcher, T.; Soubieux, D.; Ledevin, M.; Esnault, E.; Tourdes, A.; Croville, G.; Guerin, J. L.; Quere, P.; et al. Length Variations in the Na Stalk of an H7n1 Influenza Virus Have Opposite Effects on Viral Excretion in Chickens and Ducks. *J. Virol.* **2012**, *86*, 584−588.

(63) Munier, S.; Larcher, T.; Cormier-Aline, F.; Soubieux, D.; Su, B.; Guigand, L.; Labrosse, B.; Cherel, Y.; Quere, P.; Marc, D.; et al. A Genetically Engineered Waterfowl Influenza Virus with a Deletion in the Stalk of the Neuraminidase Has Increased Virulence for Chickens. *J. Virol.* **2010**, *84*, 940−952.

(64) Steven, A. C.; Cardone, G.; Butan, C.; Winkler, D. C.; Heymann, J. B. Chapter 4: Cryo-Electron Tomography of Pleiomorphic Viruses. In *Structural Virology*; Agbandje-McKenna, M.; McKenna, R., Eds.; RSC Biomolecular Sciences; Royal Society of Chemistry:: Cambridge, UK, 2011; Vol. 21, pp 62−82.

(65) Mattos, C.; Ringe, D. Locating and Characterizing Binding Sites on Proteins. *Nat. Biotechnol.* **1996**, *14*, 595−599.

(66) Mattos, C.; Bellamacina, C. R.; Peisach, E.; Pereira, A.; Vitkup, D.; Petsko, G. A.; Ringe, D. Multiple Solvent Crystal Structures: Probing Binding Sites, Plasticity and Hydration. *J. Mol. Biol.* **2006**, *357*, 1471−1482.

(67) Landon, M. R.; Amaro, R. E.; Baron, R.; Ngan, C. H.; Ozonoff, D.; McCammon, J. A.; Vajda, S. Novel Druggable Hot Spots in Avian Influenza Neuraminidase H5n1 Revealed by Computational Solvent

Mapping of a Reduced and Representative Receptor Ensemble. *Chem. Biol. Drug Des.* **2008**, *71*, 106−116.

(68) Moscona, A. Global Transmission of Oseltamivir-Resistant Influenza. *N. Engl. J. Med.* **2009**, *360*, 953−956.

(69) LeGoff, J.; Rousset, D.; Abou-Jaoude, G.; Scemla, A.; Ribaud, P.; Mercier-Delarue, S.; Caro, V.; Enouf, V.; Simon, F.; Molina, J. M.; et al. I223r Mutation in Influenza a(H1n1) Pdm09 Neuraminidase Confers Reduced Susceptibility to Oseltamivir and Zanamivir and Enhanced Resistance with H275y. *PLoS One* **2012**, *7*, e37095.