All-Atom Multiscale Simulation of Cowpea Chlorotic Mottle Virus Capsid Swelling

Yinglong Miao,[†] John E. Johnson,[‡] and Peter J. Ortoleva*[,†]

Center for Cell and Virus Theory, Department of Chemistry, Indiana University, Bloomington, Indiana 47405, Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037

Received: March 14, 2010; Revised Manuscript Received: June 25, 2010

An all-atom multiscale computational modeling approach, molecular dynamics/order parameter extrapolation (MD/OPX), has recently been developed for simulating large bionanosystems. It accelerates MD simulations and addresses rapid atomistic fluctuations and slowly varying nanoscale dynamics of bionanosystems simultaneously. With modules added to account for water molecules and ions, MD/OPX is applied to simulate the swelling of cowpea chlorotic mottle virus (CCMV) capsid solvated in a host medium in this study. Simulation results show that the N-terminal arms of capsid proteins undergo large deviations from the initial configurations with their length extended quickly during the early stage of capsid swelling. The capsid swelling is a symmetry-breaking process involving local initiation and front propagation. The capsid swelling rate is ∼0.25 nm/ns (npn) during the early stage of the simulation, and propagation of the structural transition across the capsid is roughly 0.6 npn. The system conditions that affect swelling of the capsid are analyzed. Prospects for creating a phase diagram for CCMV capsid swelling and using predictions to guide experiments are discussed.

I. Introduction

Cowpea chlorotic mottle virus (CCMV) is a member of the bromovirus group of the Bromoviridae family. Its genome consists of four positive-sense single-stranded RNA molecules, two of which are encapsulated separately in two virions, and the remaining two are packaged together into a third type of particle. The crystal structure of wild-type CCMV was solved at 3.2 Å resolution by X-ray crystallography.[1] Its capsid consists of 180 chemically identical protein subunits that form a 286-Å-diameter icosahedral shell displaying a $T = 3$ quasisymmetry. Each protein subunit is composed of 190 amino acids taking three quasiequivalent positions on the capsid surface. As such, one asymmetric unit (i.e., protomer) of the capsid includes three subunits (A, B, and C). In the X-ray crystal structure, there are 41 N-terminal residues in subunit A and 26 residues in both B and C that are not visible. Figure 1a shows the three quasi-3-fold related subunits with the missing N-terminal residues added and optimized.[2] The icosahedral capsid can also be divided into 12 pentamers and 20 hexamers with 5 A subunits in each pentamer and 3 B and 3 C subunits in each hexamer (Figure 1b).

Native CCMV is well-known to undergo a reversible swelling transition (∼10% increase in capsid diameter) when the pH is increased from 5.0 to 7.0 or above in the absence of divalent cations and at low ionic strength ($I = 0.2$M) (Figure 1c).[1,3–5] In the expansion scheme proposed by Liu et al.,[6] its swollen state can be generated by taking the pentamers and hexamers through rigid-body transformations from their native configurations as the following: translating pentamers by 24 Å radially outward and rotating them counterclockwise by 9° around their 5-fold axes and translating hexamers by 21 Å radially outward and rotating them counterclockwise by 8° around their 3-fold axes. Many other viruses undergo similar structural transitions (STs) in response to changes in their microenvironment, such as *Nudaurelia capensis* ω virus (NωV),[7,8] HK97 bacteriophage[9] and poliovirus.[10,11] Because the purified RNA and coat proteins of CCMV can reassemble in vitro to produce infectious virions,[12] and the empty CCMV capsid assembled from its protein subunits displays many behaviors similar to those of the complete virus,[4,13] it is an excellent system for studying protein−protein and protein−RNA interactions, which provide important information on the STs, self-assembly and disassembly of viruses and other biological nanostructures.

Theoretical approaches including (1) computational molecular dynamics (MD),[14–17] (2) single-scale coarse-grained models,[18,19] (3) symmetry-constrained models,[15] (4) normal-mode analysis (NMA),[20–22] and (5) Poisson−Boltzmann (PB) solvers[2,23,24] have been applied to simulate viral properties and dynamics, but they have suffered from one or more limitations as reviewed earlier.[23,25,26] In the following, we focus on the studies on viruses, notably CCMV.

MD has been applied to study RNA release from CCMV by constructing a hypothetical short single-stranded RNA, 5′-R(PGpGpApCpUpUpCpGpGpUpCpC)-3′.[17] In the MD simulation, only one asymmetric unit (i.e., a protomer) of CCMV capsid was included. Results indicate that the RNA fragment loses its secondary structure and moves into the channel along the three quasi-3-fold axis of CCMV capsid by free diffusion. However, the conclusions are not based on simulations of whole-capsid, full-length RNA behaviors.

Coarse-grained models were used to simulate a set of viral capsids,[19] although they have not been directly applied to CCMV. In the simulations, shape-based clusters of atoms were lumped into structureless beads for a reduced description of the viral capsids to allow more efficient sampling and a larger time step than that used in MD. Although the simulations provide insights into the stability of studied viral capsids, these models require parameter recalibration before simulation starts. They lack viral atomistic details needed to address local atomic interactions (e.g., the binding of divalent cations to capsid

* Corresponding author. E-mail: ortoleva@indiana.edu. Website: http://sysbio.indiana.edu.
[†] Indiana University.
[‡] The Scripps Research Institute.

10.1021/jp102314e
Published on Web 08/09/2010

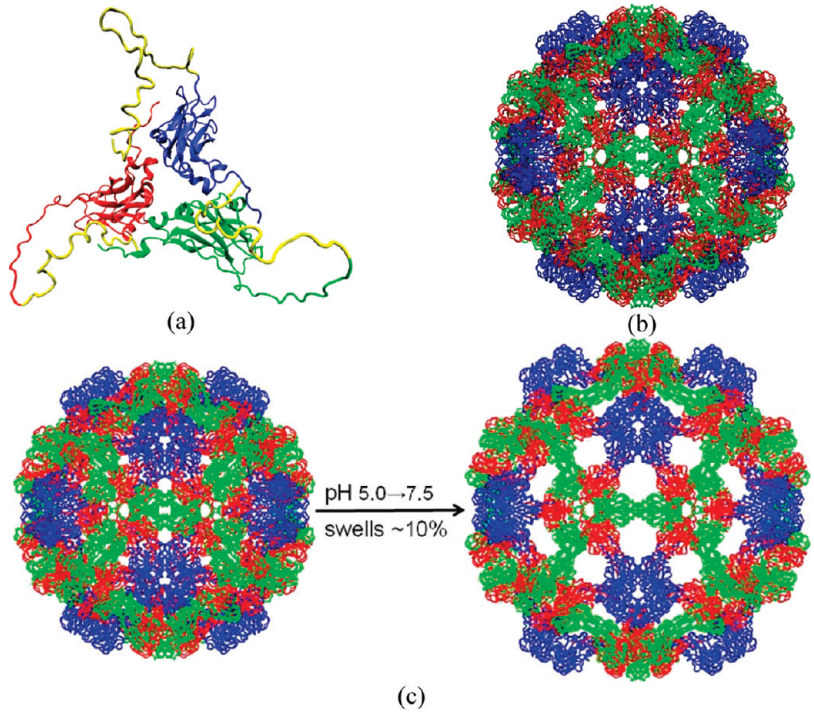

Figure 1. (a) X-ray crystal structure of CCMV protomer with the missing N-terminal residues added and optimized. Three quasi-3-fold related protein subunits are colored in blue (A), red (B), and green (C), respectively, and the N-terminal arms are in yellow. (b) Native CCMV capsid organized in 12 pentamers and 20 hexamers with 5 blue A subunits in each pentamer and 3 red B and 3 green C subunits in each hexamer. (c) Native CCMV swells ~10% as the pH is increased from 5.0 to 7.5 in the absence of divalent cations.

proteins and small-molecule drug/virus interactions) and miss the interscale feedback whereby nanoscale structural variables affect the statistical distribution of atomic-scale fluctuations, which in turn mediate the entropic and free energy effects that drive the dynamics of the nanoscale features. This feedback loop is central to a complete understanding of the dynamics of viruses as well as other bionanosystems.[27,28]

NMA was used to explore low-frequency, large-scale structural changes of CCMV capsid during its swelling.[20] It provided candidate structures of intermediate and swollen CCMV, and a putative pathway for the swelling. In the study, icosahedral symmetry was used to facilitate normal mode calculations via group theory, and various methods, including elastic network models and the use of block rotation/translation were applied to reduce the number of degrees of freedom. However, an icosahedral virus does not necessarily maintain the symmetry during its STs.[14,28,29] It is likely that initiation of instability is local, starting with the motion or deformation of a single structural unit and then propagates across the virus. On the basis of single-well (harmonic) potential approximations, NMA cannot capture the local, highly nonlinear, diffusive (friction-dominated) character of viral STs. Neither can it be readily used to study viral interactions with drugs or cell surface receptors for similar reasons.

Electrostatics of CCMV capsid were previously modeled by solving the PB equation.[24] The electrostatic interaction between RNA and CCMV capsid was simulated with a coarse-grained RNA model and a Monte Carlo approach.[2] In the model, each RNA nucleotide segment was treated as a sphere with a charge of $-0.25e$, and no connectivity was enforced between the spheres. Results showed that there is a very strong interaction between RNA segments and the highly positive N-terminal residues of the capsid proteins. RNA segments were predicted to form a shell close to the capsid with the highest densities associated under protein dimers. These high-density regions are connected to each other in a continuous net of triangles. Medium density of RNA was found under protein pentamers. These studies revealed interesting features, but they do not capture the structural constraints imposed by RNA connectivity. Furthermore, they have the limitations of coarse-grained models and do not provide a general, parameter calibration-free approach to whole-virus modeling. In addition, it is assumed that the electrical field is fixed (i.e., is imposed by the capsid), rather than being coevolved with the RNA (and possible capsid ST) during a simulation.

Therefore, although the above approaches reveal interesting features of CCMV stability and dynamics, they do not meet the challenge of simulating the whole-virus dynamics. Viruses such as CCMV are normally composed of millions of atoms, and their STs involve both atomistic fluctuations and nanoscale processes (e.g., large-scale translation and rotation of protein capsomeres) spanning across multiple time and length scales. It is natural to adopt a multiscale approach to study their dynamics. Multiscale analysis has been of interest since work on Brownian motion by Einstein.[30] It was shown that the wandering of a nanoparticle arises due to the interplay of short-scale random collisions and large-scale, slow motions of the whole nanoparticle created by the separation in the magnitude of the mass of an atom versus that of the nanoparticle.

In earlier studies, Fokker-Plank (FP) and Smoluchowski equations were derived from the Liouville equation for nanoparticles without internal atomistic structure.[31–38] The development of projection operators[39] is not directly applicable because it requires integration over atomistic configurations and thereby makes it difficult to capture internal structure and involves memory kernels that are not readily evaluated using standard MD. We have advanced the multiscale approach to develop an all-atom multiscale analysis (AMA) theory for dynamical nanosystems[25] by (1) accounting for atomic-scale fluctuations of the system internal structure; (2) introducing general,

TABLE 1: Simulations Showing CCMV Capsid Swelling Is Strongly Dependent on System Conditions

simulation	pH[a]	$[Mg^{2+}]$ (mol/L)	water boundary[b] (Å)	ensemble[c]	N_{atoms}[d]	performance (days/ns)[e]	time (ns)	increase in capsid diameter (Å)[f]	increase in capsid radii (Å)[g] outer	increase in capsid radii (Å)[g] av
sim1	7.0	0	7	NPT	2 411 918	2.60	3	10.23	5.74	1.35
sim2	7.0	0	15	NPT	2 845 915	2.95	3	10.45	7.61	1.58
sim3	7.0	0	15	NVT	2 845 915	2.36	5	26.08	18.05	10.48
sim4	5.0	0	15	NPT	2 846 275	2.95	3	8.54	6.20	1.94
sim5	5.0	0.05	15	NPT	2 841 943	2.95	3	6.86	4.17	1.31

[a] The pH of the system is set such that amino acid residues in the capsid are protonated/deprotonated corresponding to their pK_a. [b] The water boundary is the minimum distance between atoms on the capsid surface and faces of the water box. [c] NPT means a constant pressure (1 atm) and temperature (298.15 K) ensemble, and NVT is constant volume (unchanged from dimensions of the initial system prepared with VMD) and temperature (298.15 K). [d] There are 522 000 atoms in CCMV capsid, and the total number of atoms in the systems ranges from ~2.41 to ~2.85 million, depending on the thickness of the water boundary and salt concentration used for solvation. [e] Performance here is for running simulations with 256 processors on the Indiana University Big Red cluster. [f] The capsid diameter is calculated as the maximum distance between two C_α atoms in the capsid. [g] The capsid outer and average radii are calculated as the average and maximum distance from heavy atoms in the capsid backbone to their center of mass, respectively.

automatically generated collective modes and order parameters (OPs); (3) constructing ensembles of atomic configurations constrained to fixed values of the slowly varying OPs, as needed to construct thermal-average forces and friction/diffusion coefficients for their stochastic equations; (4) applying perturbation expansion to the solution of *N*-atom Liouville equation in terms of a smallness parameter (length or time scale ratio of one atom to the entire system), and (5) deriving stochastic equations (FP or Smoluchowski) for the probability density of coarse-grained system OPs. The AMA approach captures the cross-talk between the OPs and atomistic variables rigorously,[25,40–42] and provides the conceptual framework for our whole-virus simulation approach.

On the basis of concepts from AMA theory of nanosystem dynamics, we recently developed a multiscale molecular dynamics/order parameter extrapolation (MD/OPX) approach for simulating viruses and other large bionanosystems.[26,28] In the present implementation, it is optimized for nonenveloped viruses or the core of genomic complex of enveloped viruses, such dengue virus.[43,44] OPs constructed with orthogonal polynomials of atomic coordinates of a reference configuration or atomic displacements between two known configurations are introduced to capture the system, slowly varying nanoscale features. Replica short MD runs with random atomic velocity initializations are implemented to estimate the ensemble-average rate of change in OPs, which is then used to extrapolate the system over a time period that is much longer than the 10^{-14}-s time scale of fast atomic vibrations and collisions. Since the time scale for the evolution of OPs is much larger than that of atomic vibrations and collisions, the OP extrapolation time can be many orders of magnitude greater than the MD simulation time step. The resulting MD/OPX algorithm addresses rapid atomistic fluctuations and slowly varying coherent dynamics of bionanosystems simultaneously and accelerates their MD simulations. Its underlying all-atom description of bionanosystems enables the use of a universal interatomic force field, avoiding recalibration with each new application as needed for coarse-grained models. The equation free multiscale analysis (EFM) approach developed by Kevrekidis et al.[45,46] shares much of the flavor of MD/OPX; that is, short bursts of MD simulations can be used to extrapolate coarse variables (i.e., OPs here) over large time intervals and thus project the system over a long time. It has been applied to an alanine dipeptide for molecular dynamics study by using the dipeptide dihedral angle ($N-C_\alpha-C-N$) as its coarse variable.[46] However, the dipeptide is a very small system, and thus, its dihedral angle does not appear to be slowly varying compared with our automatically constructed OPs for large bionanosystems.

In our earlier study,[28] MD/OPX was used to simulate the swollen state of CCMV capsid in vacuum (i.e., gas phase). The swollen CCMV capsid shrinks significantly during 200 ns simulation, and it was found to be an energy-driven, symmetry-breaking process that starts locally and then propagates across the capsid. It involves large-scale cooperative motions of protein pentamers and hexamers in the capsid. In this study, MD/OPX is applied to the swelling of the CCMV capsid solvated in a host medium at pH 7.0 and ionic strength $I = 0.2$M. We first identify key elements of a model needed to simulate the swelling of the native CCMV capsid through a short-time MD simulation, then MD/OPX is applied to capture the long-time dynamics of CCMV capsid swelling with the mechanisms of the capsid swelling and its dependence on system conditions is discussed, and conclusions are drawn.

II. Computational Methods

CCMV Capsid Structure and MD Simulation Protocols. The X-ray crystal structure of CCMV protomer (i.e., asymmetric unit) was downloaded from the online Protein Data Bank (PDB ID: 1CWP). The missing N-terminal residues of protein subunits were added and further optimized with residues in the crystal structure fixed by using NAMD.[48] Therefore, they do not introduce unphysical stress to the capsid (refer to the scheme provided in Zhang et al.[2]). The native CCMV capsid was generated by reorienting 60 copies of the protomer according to the icosahedral symmetry. The capsid was then solvated in a water cube with different water boundaries by using visual molecular dynamics (VMD)[47] (see Table 1). Protonation states of the amino acid residues were set according to a specific pH; for example, 5.0 or 7.0. Because amino acids usually have one protonation state over a range of pH values, a simulation model prepared at pH 7.0 actually works over a pH range; for example, 7.0−8.0. K^+ and Cl^- were added randomly to neutralize the solvated system, achieving an ionic strength of 0.2 M.

With the above resulting system and its structural topology, NAMD[48] was applied to run molecular dynamics simulation by using the CHARMM22 force field[49,50] for capsid proteins and TIP3P model for water molecules.[51] The simulation proceeded with an initial energy-minimization of the entire system for 10 000 steps using the conjugate gradient algorithm until the potential energies level off. Following the energy minimization, the system was gradually heated to 298.15 K and equilibrated for 20 ps, after which the system temperature and

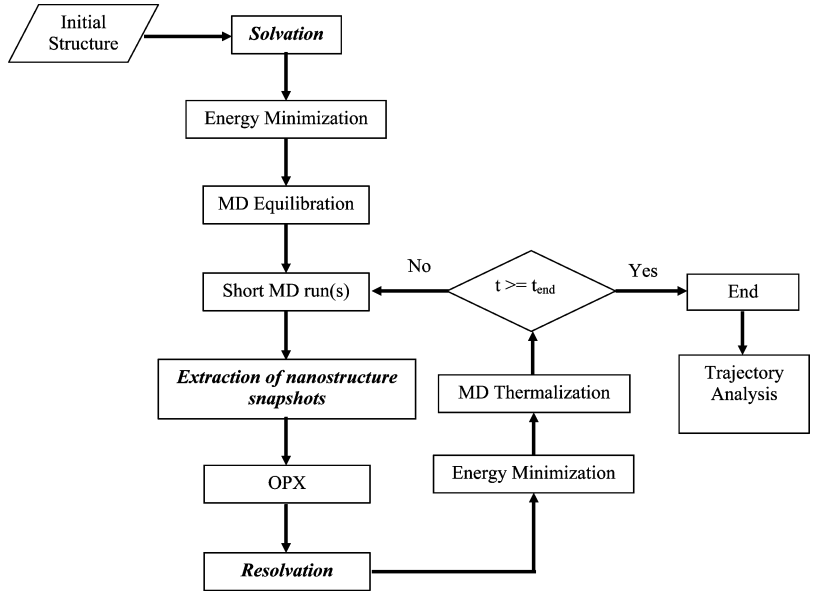

Figure 2. Schematic flowchart of MD/OPX for simulation of nanostructures solvated in a host medium with new modules for "solvation", "extraction of nanostructure snapshots", and "resolvation" integrated.

pressure (in the case of pressure control) stay stable around specified values. The product MD trajectory was then obtained for investigating the capsid structural changes.

Periodic boundary conditions were applied to remove surface effects of the system. Either a constant volume/temperature (NVT) or a constant pressure/temperature (NPT) ensemble can be used for running MD. The temperature set in the simulations in this study was 298.15 K, and the pressure was 1 atm for the NPT simulation. The volume of the system was determined by the lattice vectors specified for its unit cell. The lattice vectors were initially calculated according to the dimensions of the resulting system prepared with VMD. They were kept constant for NVT simulation while scaled isotropically in the NPT simulation by applying the Nose−Hoover Langevin piston pressure control. The integration time step used in our MD simulation was 1 fs, and a multiple-time-stepping algorithm[48] was employed in the following manner: bonded interactions were computed for every time step; short-range nonbonded interactions, every two timesteps; and long-range electrostatic interactions, every four timesteps. A cutoff distance of 12 Å was used for van der Waals and short-range electrostatic interactions, and the long-range electrostatic interactions were computed with the particle-mesh Ewald summation method using a grid point density of 1/Å. SHAKE[52] constraints were applied on hydrogen−oxygen bonds in water molecules.

To investigate the controlling factors of CCMV capsid swelling, five MD simulations were performed on the native state of CCMV capsid that was solvated in a host medium under various conditions. The complete CCMV capsid with the missing N-terminal residues of protein subunits added has 522 000 atoms, and the total number of atoms in the simulated systems ranges from ∼2.41 to ∼2.85 million, depending on the thickness of water boundary and salt concentration used for solvation. Simulations were run on the Indiana University Big Red cluster, and the performance goes from 2.36 to 2.95 days/ns when using 256 processors. A summary of the characteristics of prepared systems and the simulations is listed in Table 1.

MD/OPX for Simulating Solvated Systems. Although earlier MD/OPX studies were focused on simulating nanostructures in vacuum (i.e., gas phase),[26,28] new modules have been developed to account for water molecules and ions for simulating nanostructures solvated in a host medium. A flowchart of MD/OPX with new modules for "solvation", "extraction of nanostructure snapshots", and "resolvation" integrated is shown in Figure 2. MD/OPX is implemented by using NAMD[48] as the MD platform. VMD is used for "solvation" and "resolvation" of simulated nanostructures and trajectory analysis.

MD/OPX starts with an initial atomic structure that can be determined from X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and cryo-electron microscopy. For system preparation, the structure is solvated in water and neutralized under certain pH and ionic strength (see example procedure described above for CCMV capsid). The resulting system along with its structural topology is then taken as input to NAMD for initial energy minimization and equilibration, after which the simulation is performed in MD/OPX cycles until it reaches the end of specified time. In each MD/OPX cycle, short MD runs are used to generate the time evolution of the system atomic configuration. With several snapshots chosen from the system trajectory, the nanostructure is extracted from the system without water molecules and ions for calculating its structural OPs (see the Appendix) and, therefore, estimating their rates of changes, that is, time derivatives, which are then used to extrapolate OPs over long time. The atomic configuration of the nanostructure is then reconstructed from the newly obtained OPs and resolvated in the host medium as explained below. Energy minimization and MD thermalization are applied to the resulting system before running short MD for another cycle.

In the "resolvation" module, the OPX resulting nanostructure is first put back to the host medium by using the final configuration of the preceding short MD run, which keeps the same number of water molecules and ions (i.e., same system structural topology) as needed for continuing the simulation in NAMD. Water molecules and ions that are within a certain distance of the nanostructure (denoted d_s) are removed and then redistributed as follows, since many of them may overlap with the nanostructure or be embedded in it. First, a water box that is large enough to solvate the nanostructure with at least a water boundary of d_s (340^3 Å^3 for simulation of CCMV capsid swelling) is prepared and equilibrated under 298.15 K and 1 atm by running MD for 100 ps. With the nanostructure placed in the center of the equilibrated water box, the coordinates of

water molecules, which are selected as within d_s and not overlapping with the nanostructure, are retrieved to replace those of the above removed water molecules and ions. Note that the number of the retrieved water molecules may not be exactly the same as that needed for redistribution due to the water density difference between the two systems. Thus d_s is adjusted iteratively by small increments for selecting water molecules in the equilibrated water box until they are enough for the redistribution of the removed water molecules and ions. The ions are moved to positions that are nearest to their original ones, and the overlapping water molecules are removed. This not only generates the final configuration that has no overlapping between the nanostructure and host particles, but also maintains local density and pressure of the host medium for the nanostructure across the OPX time step without introducing unphysical stress.

MD/OPX Simulation of CCMV Capsid Swelling. MD/OPX simulation of CCMV capsid solvated in a water cube with 15 Å water boundary continued with the resulting structure of its MD simulation (sim2 in Table 1) at 1 ns as described. It proceeded with cycles of short MD run and projection of the atomic structure by extrapolating its structural OPs. In each cycle, one 500 fs MD run was obtained by running NAMD, and 5 snapshots were chosen from the trajectory, with which 27 capsid OPs were constructed by using Legendre polynomials of atomic coordinates over order (0, 1, 2) in the X, Y, and Z directions (see the Appendix). The capsid OPs were extrapolated over long time intervals by using least-squares fitting. The newly obtained OPs were taken to reconstruct atomic configuration of the capsid for resolvation in the water cube as described above, followed by a 5000-step energy minimization, 6-stepwise graduate heating to 298.15 K, and 5 ps MD thermalization. Atomic accelerations of the energy-minimized and MD-thermalized configurations were computed as indicators to determine whether the structure is mature for the next MD/OPX cycle (i.e., no unphysical large forces on the atoms and atomic accelerations need to be smaller than a tolerance), and the time step for extrapolation of OPs is adjusted adaptively to ensure the simulation is stable. As explained above, resolvation in MD/OPX is implemented in such a way that the local density and pressure of the host medium for the capsid structure is maintained across the OPX time step, and thus, there is no significant configurational change in the capsid and host medium interface during maturation (i.e., energy minimization, graduate heating and thermalization) of the resolvated system. The 5 ps MD run is used to thermalize the system, rather than structurally equilibrate it. It is short, allowing large acceleration of MD/OPX over direct MD.

By repeating such cycles, native CCMV capsid (denoted Ncap) was simulated for 3 ns. Analysis of the MD/OPX trajectory showed that the capsid swells in the water cube (see Results and Discussion). At 1.5 ns, the capsid expands to a limit such that it starts to interact with the replicas of its neighboring cells. Therefore, the capsid was extracted from the trajectory snapshot and solvated in a larger water cube (25 Å water boundary) for further swelling simulation. This capsid intermediate (denoted Icap) was simulated with MD/OPX by following the same procedure as used for Ncap. MD/OPX trajectory of Icap lasting 17.04 ns was used for investigating swelling of the capsid.

The OP time step obtained from adaptive MD/OPX cycles of simulating Ncap and Icap was found to undergo periodic oscillations, and its average was determined to be 3.5 ps. With this, the ratio of average OP time step to the time interval of the short MD run is 70. In the case of simulating Icap solvated in a water cube with a 25 Å water boundary, there are 3 614 545 atoms in the system. The overall speedup of MD/OPX over MD is 2.14 times faster when using 128 processors on the Indiana University Big Red cluster and 1.58 with 512 processors, and the resulting performance is 2.50 days/ns and 0.91 days/ns, respectively. The reason for downgrading of the speedup when using more processors is that extraction of trajectory snapshots and resolvation of the capsid executed by calling VMD on only one node (4 processors) take up a larger portion of the MD/OPX running time, whereas parallel MD is accelerated by running on more processors.

In the above MD/OPX simulation, 5000 steps of energy minimization were applied to the entire system as a maturation process of the OPX resulting configuration to remove its unphysical structures; for example, overlapping atoms and stretched bonds/angles. Instead, energy minimization can be applied to the OPX resulting configuration of the capsid only before resolvation. This is not implemented in the present MD/OPX because the latest NAMD[48] has problems with embedding a NAMD run (for energy minimization of the nanostructure only) inside NAMD for the top level MD/OPX simulation. However, when this is resolved, the speedup of MD/OPX over MD can be improved to 4.40 by using 128 processors or 2.67 with 512 processors.

Analysis of CCMV Capsid: Volume, SASA, and Water Density. To measure the volume of CCMV capsid (i.e., protein shell), the capsid snapshot extracted from simulation trajectories is placed at the center of the MD-equilibrated homogeneous water box that has been used for "resolvation" in MD/OPX. The average density of water in the box is computed as 975.51 kg/m^3. By selecting water molecules within 1.4 Å of the capsid, the volume of the capsid is considered as the total volume of selected water molecules, which can be calculated by summing their masses and dividing it by the water density. The volume of the capsid cavity is determined as the total volume of water molecules that are not within 1.4 Å of the capsid and for which the radial distance to the capsid COM is smaller than the capsid average radius. Once the volume of the capsid and that of its cavity are determined, the volume for the space outside of the capsid is calculated by subtracting them from the total volume of the water box. Because these volume calculations are performed by immersing the capsid in a water box, they are applicable to nanostructures of any shape without symmetry constraints and not affected by the symmetry-breaking process of CCMV capsid swelling described below. With this, the average density of water inside the capsid cavity and outside of the capsid can be calculated, and their time evolution can be studied along the simulation trajectories. For calculating the capsid solvent-accessible surface area (SASA), the radius of protein atoms in the capsid is extended by 1.4 Å to determine whether they are exposed to the solvent.

III. Results and Discussion

Insights into CCMV Capsid Swelling through Short-Time MD Simulation. To identify key elements needed to carry out long-time MD/OPX simulation of CCMV capsid swelling, short-time MD simulation was first run on native CCMV capsid that was solvated in a host medium at pH 7.0 and ionic strength $I = 0.2$ M. Native CCMV capsid was generated by reorienting 60 copies of its protomer (i.e., asymmetric unit) according to the icosahedral symmetry and solvated in the center of a water cube with 7 Å boundary in each direction and neutralized with K^+ and Cl^- to $I = 0.2$ M by using VMD. A schematic

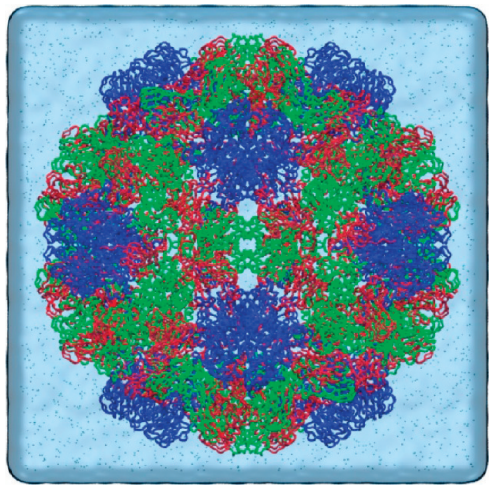

Figure 3. Schematic representations of CCMV capsid solvated in a water cube used for our simulations with the water box created using VMD "volmap" plugin and ions represented in CPK.

representation of the resulting system is shown in Figure 3a. Then the system along with its structural topology file was taken as input for NAMD[48] simulation by using the CHARMM22 force field[49,50] for capsid proteins and TIP3P model for water molecules[51] (see Computational Methods). The simulation proceeded with initial energy minimization, thermalization with graduate heating, and MD equilibration under constant temperature (298.15 K) and pressure (1 atm) control (i.e., NPT ensemble). The system was equilibrated after 20 ps when the system temperature and pressure fluctuate around target values. A product MD run lasting 3 ns was obtained to investigate structural changes of the capsid to identify essential features needed to model swelling of the capsid.

Although the capsid average radius (i.e., the average distance between atoms in the capsid backbone and their COM) might better account for the distribution of capsid atoms, the capsid size is normally described by its diameter in experiments.[1,8,53] In addition, N-terminal arms of viral capsid proteins are known to be highly dynamic, fluctuating in the capsid interior,[1,54] and the capsid minimum radius (i.e., the minimum distance between atoms in the capsid backbone and their COM) is not suitable for measuring the capsid size. Therefore, we use the capsid diameter calculated as the maximum distance between two C_α atoms in the capsid along with the maximum radius of the capsid backbone (i.e., the maximum distance from atoms in the capsid backbone to their COM) to characterize CCMV capsid size and thus describe swelling of the capsid in the following.

Analysis of the MD trajectory shows that the capsid size stays roughly constant during initial energy minimization, thermalization, and NPT equilibration with the outer radius of the capsid backbone increased by only 0.12 Å, although the average and inter radii of the capsid backbone are deceased by 1.60 and 4.40 Å, respectively. The capsid starts to swell after equilibration. For the resulting structure at 3 ns, its outer radius is 5.74 Å greater than the native CCMV capsid and 1.35 Å greater for the average radius. The capsid diameter is increased by 10.23 Å. It was also found to be 13.99 Å thicker than the native state with its maximum radius (i.e., the maximum distance of capsid backbone atoms to the capsid COM) increased from 142.49 to 148.23 Å and the minimum radius decreased from 82.32 to 73.85 Å (Figure 4a). This indicates that the capsid proteins undergo fluctuations during capsid swelling. The thickness of the water layer surrounding the capsid backbone in the *X*, *Y*, and *Z* directions, which are calculated by averaging the

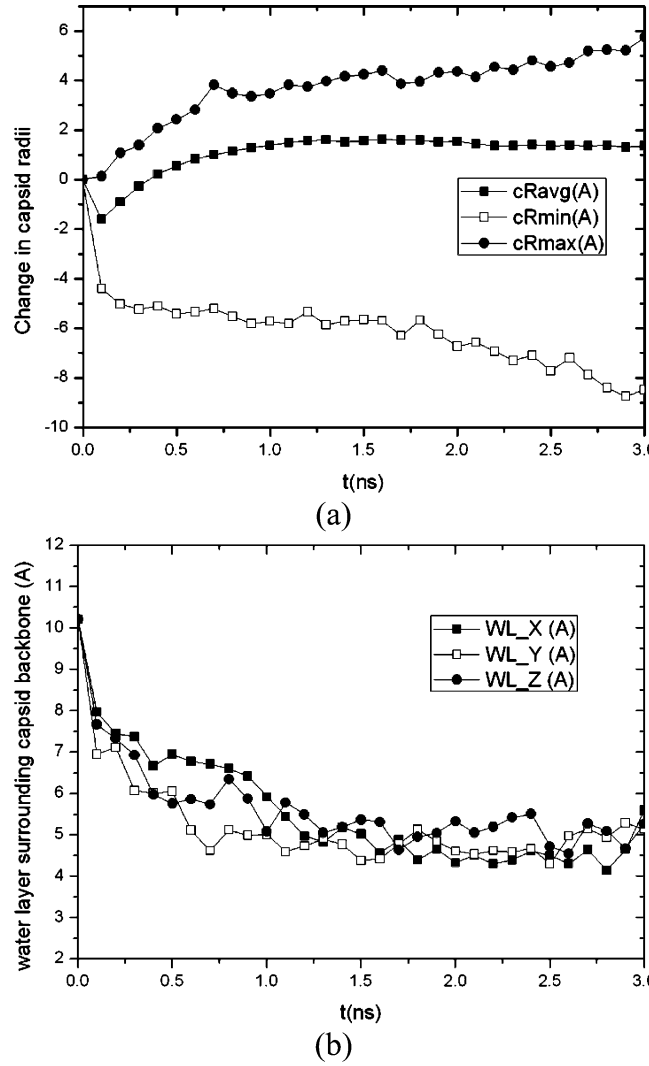

Figure 4. MD simulation (3 ns) of native CCMV capsid solvated in a water cube with an initial 7 Å water boundary: (a) changes in the average, minimum, and maximum radii of the capsid backbone from the starting configuration and (b) averaged thickness of water layer surrounding the capsid backbone in the *X*, *Y*, and *Z* directions.

minimum distance between non-hydrogen atoms in the capsid backbone (excluding those in protein side chains) and the faces of water box on two sides of each direction, were found to decrease as the capsid swells (Figure 4b). As a result, the distance between the capsid surface and that of its replicas in neighboring periodic cells decreases. When it reaches the cutoff distance of nonbonded atomic interactions (usually 12 Å for short-range van der Waals and electrostatic interactions in MD simulation), the capsid starts to interact with the replicas of its neighboring cells, which slows down swelling of the capsid.

The above MD simulation shows that the native CCMV capsid swells in 3 ns when placed in a host medium at pH 7.0. However, the simulation takes 2.60 days to finish 1 ns by using 256 processors on the Indiana University Big Red cluster. And as the capsid swells, the water box needs to be expanded to appropriately simulate swelling of the capsid (see discussions below). With the system size largely increased, the performance of MD simulation is downgraded, and it becomes even more time-consuming to simulate the swelling of CCMV capsid. Since viral STs involve processes taking place across multiple time and length scales, in the following, we apply multiscale MD/OPX to CCMV capsid swelling, which accelerates MD for long-

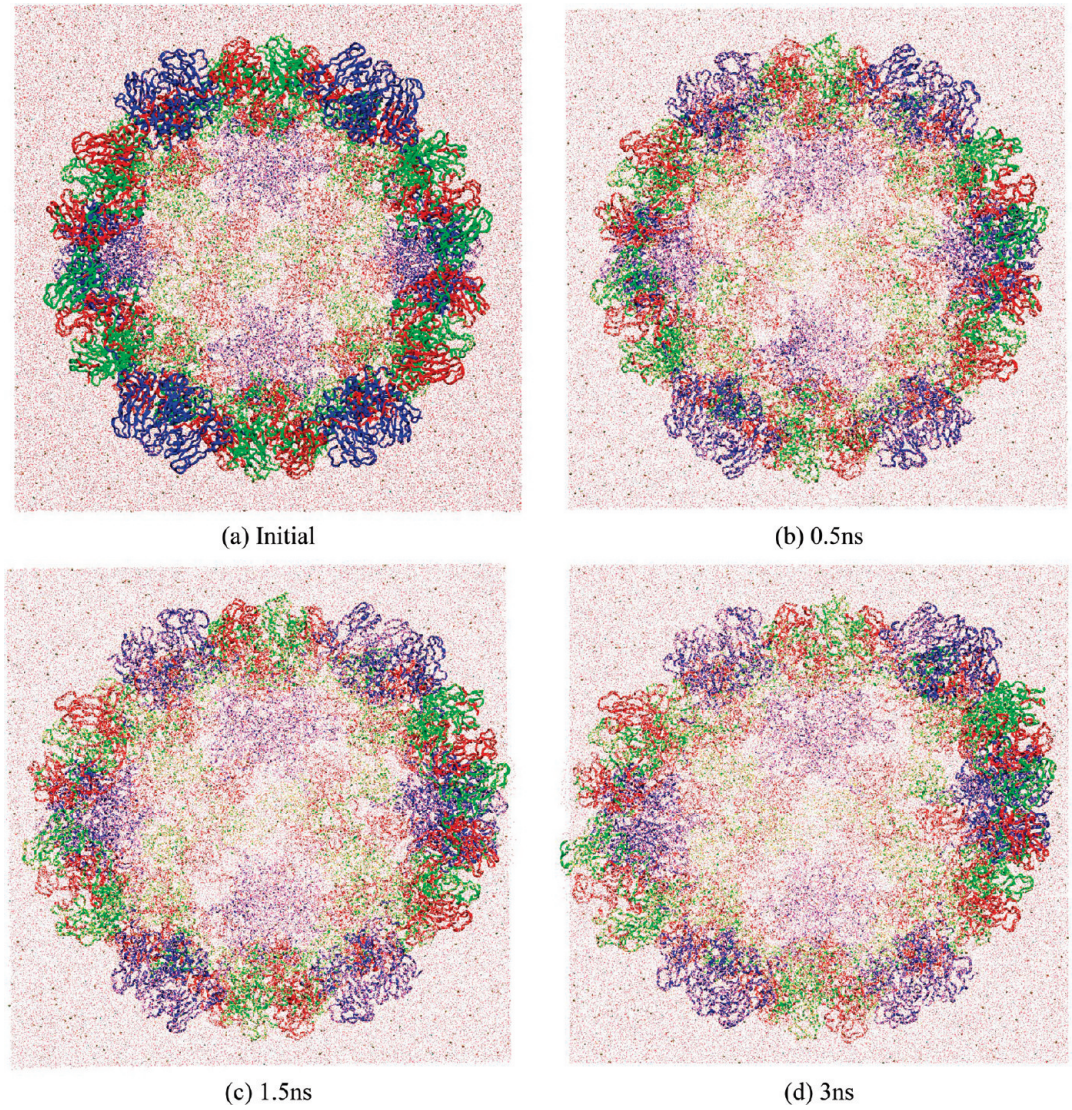

Figure 5. MD/OPX trajectory snapshots showing the expansion of native CCMV capsid (Ncap) in a water cube with an initial 15 Å water boundary: (a) the starting configuration, (b) 0.5 ns, (c) 1.5 ns, and (d) 3 ns.

time simulation and addresses the atomistic fluctuations and nanosystem coherent motions simultaneously.

MD/OPX Simulation of CCMV Capsid Swelling. Since earlier MD/OPX studies,[26,28] new modules have been added to account for water molecules and ions and thus achieve simulating bionanosystems in a host medium. Here, MD/OPX is used to simulate swelling of the CCMV capsid in a host medium. Becausethe native CCMV capsid is expected to expand by ~24 Å to reach the size of its swollen state,[1] simulation of its swelling is divided into the following two phases for efficiency: (1) increasing the water boundary from the 7 Å used in the above MD simulation to 15 Å for solvating native CCMV capsid (Ncap) and running MD/OPX on the resulting system until the capsid swells to the size limit of the water cube, and (2) extracting the intermediate capsid structure (Icap) obtained from step 1 and solvating it in a larger water cube for further simulation of capsid swelling.

For the first phase of the MD/OPX simulation, we started with native CCMV capsid (Ncap) solvated in a water cube of 15 Å water boundary. MD/OPX trajectory lasting 3 ns was obtained to study the structural changes of the capsid (see Computational Methods). Figure 5 shows trajectory snapshots of the capsid in water cube at (a) the starting configuration, (b) 0.5 ns, (c) 1.5 ns, and (d) 3 ns. The capsid was seen to expand in the MD/OPX trajectory. Quantitative analysis of the MD/OPX trajectory shows the capsid swells with the outer radius of its backbone increased by 7.58 Å in 3 ns and 2.92 Å for the average radius (Figure 6a). The capsid swells fast up to 2 ns at a rate of ~0.22 npn in the outer radius, and the swelling slows down thereafter. The predicted capsid swelling rate and the time course of capsid radii are in good agreement with the direct MD simulation (sim2 in Table 1 as described below). The rmsd's of atomic positions between capsid snapshots along the trajectory and its starting structure increases from 0 to 12 Å in 3 ns (Figure 6b). Calculations show the thickness of the water layer surrounding the capsid backbone decreases as a result of capsid swelling. It reaches 6 Å (half of the cutoff distance for atomic nonbonded interactions) at 1.5 ns (Figure 6c), within which the capsid starts to interact with the replicas of its neighboring cells. Therefore, it is necessary to extract the capsid structure from the trajectory snapshot at 1.5 ns and solvate it in a larger water cube for further swelling simulation.

With the intermediate capsid structure (Icap) extracted from the above MD/OPX trajectory at 1.5 ns, we solvated the capsid in a larger water cube (25 Å water boundary) and then performed 17.04 ns MD/OPX simulation. Analysis of the output

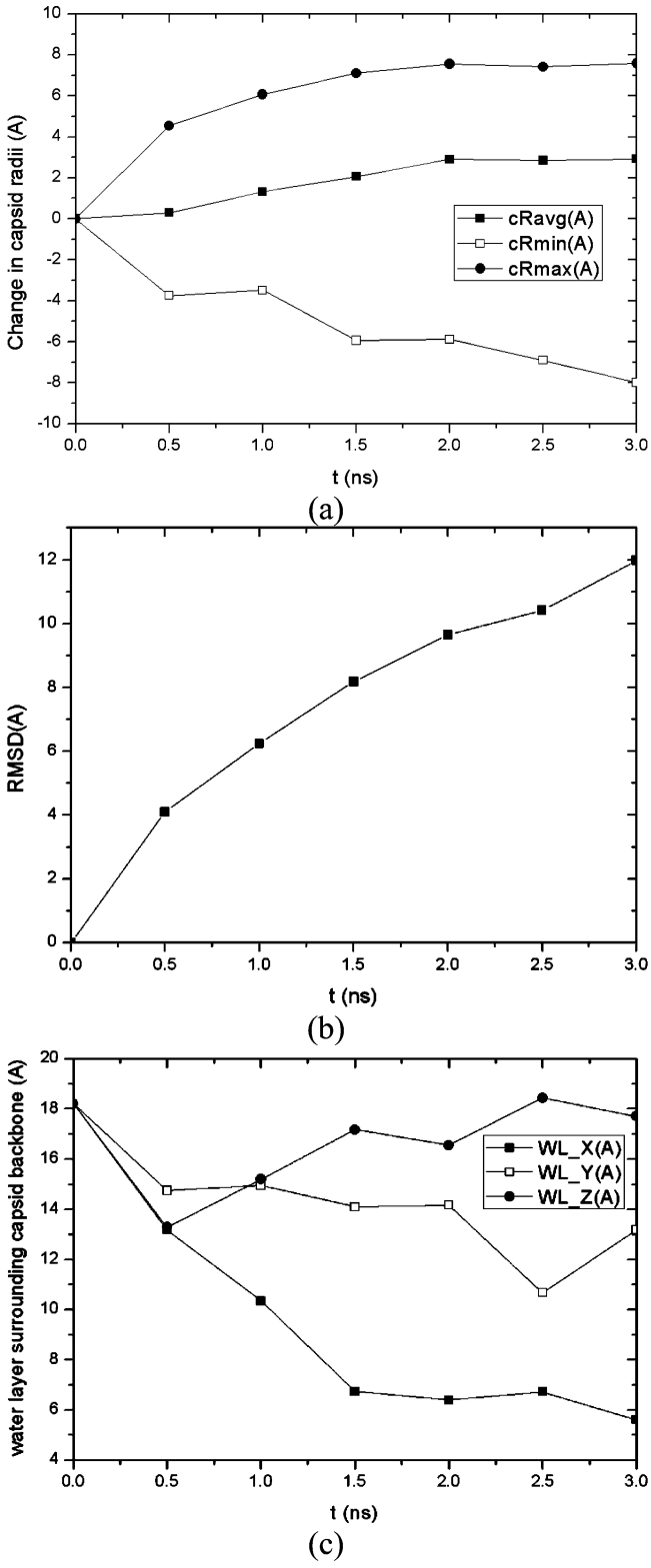

Figure 6. Analysis of the 3 ns MD/OPX trajectory of native CCMV capsid (Ncap): (a) changes in the average, minimum, and maximum radii of the capsid backbone from its starting configuration, (b) rmsd of atomic positions between snapshots of capsid backbone along the trajectory and the starting configuration, and (c) averaged thickness of the water boundary surrounding the capsid backbone in the *X*, *Y*, and *Z* directions.

MD/OPX trajectory shows the capsid swells by 20.41 and 15.95 Å, with small fluctuations in its backbone maximum and minimum radii, respectively, and 13.93 Å in the capsid average radius during 17.04 ns (Figure 7a). The rmsd of atomic positions between capsid snapshots along the trajectory and its starting structure (Icap) increases from 0 to 16.05 Å in 17.04 ns (Figure 7b). The capsid was found to swell fast during the early stage of the simulation, with a quick increase in both the capsid average radius and rmsd. Notably, the capsid swells by ~12.94 Å in its outer radius over 5.16 ns, i.e., ~0.25 npn. The swelling slows down as the capsid changes toward its near-equilibrium state; that is, the swollen form of the capsid.

Combining the above two MD/OPX trajectories into one, we obtained 18.54 ns dynamics of CCMV capsid swelling from its native state. Total increase in the capsid outer radius is 26.49 Å and 14.93 Å in the average radius. Analysis of the entire trajectory (Figure 8) showed that the capsid cavity volume increases from 4.36×10^6 to 6.86×10^6 Å^3 (by 57%) during the course of swelling. The capsid SASA was found to increase from 1.20×10^6 to 1.47×10^6 Å^2 in 18 ns with moderate fluctuations. These indicate that regions of capsid proteins buried in the capsid interior in its native state are exposed to the solvent during capsid swelling, which also leads to an increase in the capsid volume.

CCMV Capsid Swelling Mechanisms. In our earlier study,[28] cooperative motions of protein capsomeres were described for the swollen form of a CCMV capsid simulated in vacuum (i.e., gas phase). To characterize the capsomere cooperative motions during the swelling of the CCMV capsid in a host medium, we calculated the average COM translation distances and rotation angles of 12 pentamers and 20 hexamers along the trajectory obtained from the above MD/OPX simulation. As shown in Figure 9, pentamers translate radially outward by 15.10 Å on average in 18.54 ns and 14.24 Å for hexamers; they also undergo counter-clockwise rotation by 2.92° and 2.28°, respectively, with fluctuations along the trajectory. Although the 18.54 ns MD/OPX simulation has not completely captured the entire swelling process of the CCMV capsid, these predictions agree with experimental data on the expansion of the CCMV capsid in water from its native to the swollen state through rigid-body transformations of the capsomers.[1,6] Therefore, the swelling of the CCMV capsid in a host medium involves large-scale translation and rotation of the pentamers and hexamers, which was also found in the shrinkage of swollen CCMV capsids in vacuum studied earlier.[28] The capsomeres undergo cooperative motions through strongly coupled allosteric interactions during viral capsid STs.

We now explore the swelling mechanisms of CCMV capsid by observing the atomistic motions of the capsid. Figure 10 shows the interior view of the capsid trajectory snapshots at (a) the initial native state and (b) 1.8, (c) 3, (d) 6, (e) 12, and (f) 18 ns. Atoms in the capsid are represented by their displacements from the original positions with a blue-white-red color scale (i.e., 0 Å for blue, 24 Å for red, and white as the midpoint) to investigate conformational changes in the capsid. N-terminal arms of the capsid proteins are highlighted as ribbons, and they are colored according to the displacements of atoms, as well. It is seen that atomic displacements are not constrained to the icosahedral symmetry, that is, atoms in the capsid do not change their color simultaneously in a symmetry-preserving manner. Especially, the N-terminal arms change their color from blue to red during the early stage of capsid swelling (see trajectory snapshots at 1.8 and 3 ns and the movie in the Supporting Information).

Quantitative calculations show that the atomic rmsd between trajectory snapshots and the initial PDB structure for C_α atoms in the N-terminal arms of capsid proteins (the first 26 residues in each protein subunit) displays much greater increase than

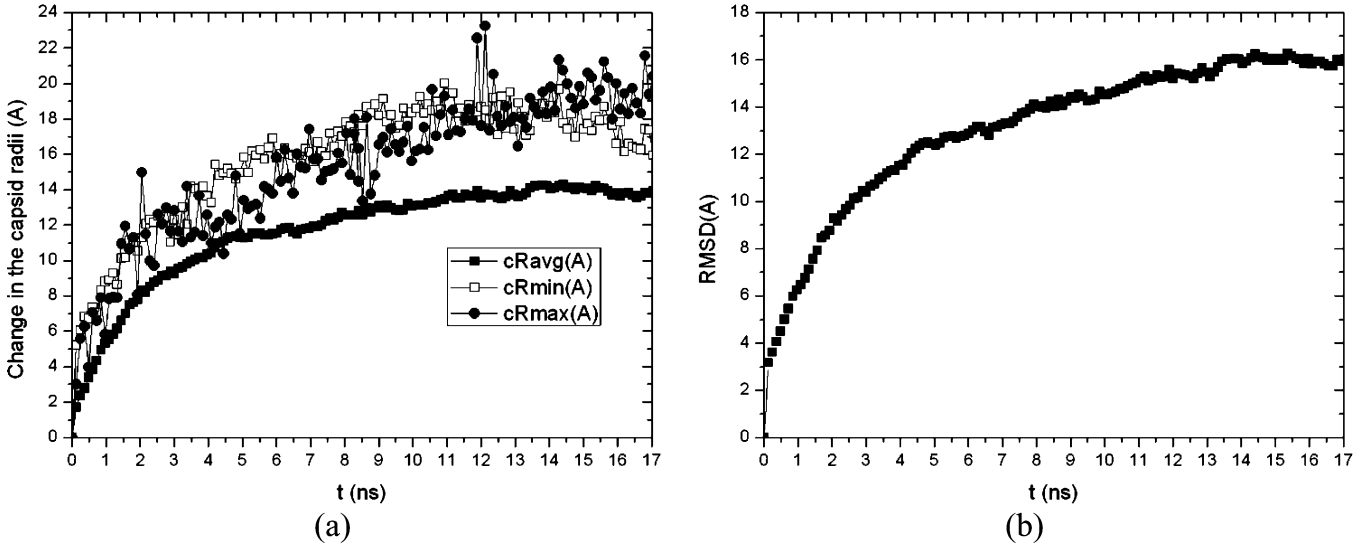

Figure 7. MD/OPX (17.04 ns) simulation of the swelling of capsid intermediate (Icap): (a) time course of the change in the capsid average, minimum, and maximum radii and (b) rmsd of atomic positions between snapshots of capsid backbone along the trajectory and the initial structure.

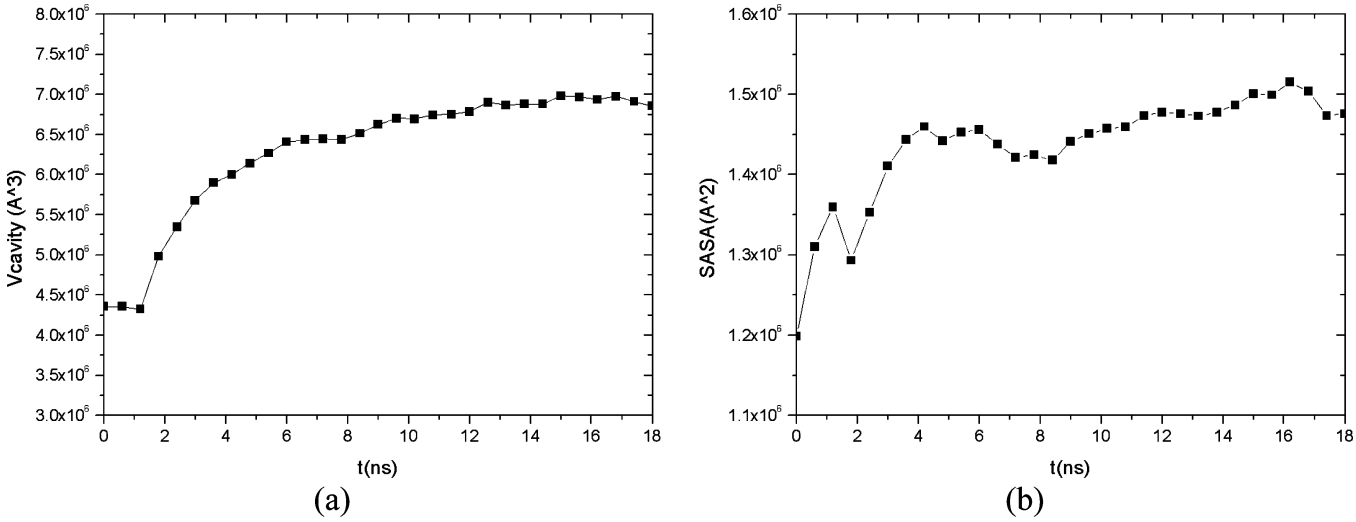

Figure 8. Analysis of the CCMV capsid along the entire 18.54 ns MD/OPX trajectory: (a) the volume of capsid cavity and (b) capsid SASA.

that for the rest of the capsid protein backbone, with their difference being 2.78 Å at 0.6 ns, increased to 5.32 Å at 8.4 ns, and becoming nearly constant thereafter (Figure 11a). This indicates that the N-terminal arms undergo larger deviations from their initial configurations during capsid swelling, especially in the early stage. The change in the N-terminal arm length (i.e., the distance between the C_α atom of the first and 26th residues) as averaged over 180 capsid proteins is found to increase quickly by 4.18 Å at 0.6 ns, reach 6.82 Å at 4.8 ns, and remain constant afterward (Figure 11b). In contrast, change in the capsid average radius increases much more slowly during the early stage of simulation and keeps increasing near the end. This shows that the N-terminal arms extend their length much more quickly than global expansion of the capsid.

Similar to the shrinkage of the swollen CCMV capsid in vacuum simulated earlier,[28] the ST of the CCMV capsid during its swelling involves "local initiation and front propagation" (see the movie in the Supporting Information). The nucleation takes ~6 ns to occur, as seen by the color change of two entire capsomers from blue to red in the trajectory snapshot (i.e., the upper two shown in Figure 10d). This is 14 ns shorter than the nucleation time needed for the shrinkage of the swollen CCMV capsid in vacuum. One reason is that N-terminal arms of capsid proteins were added in the present simulation of CCMV capsid swelling in a host medium, and their fast motions accelerate the nucleation of capsid ST. After nucleation, the capsid ST then propagates across the capsid indicated by spreading of the red areas. The propagation of cooperative motions from one capsomere to its neighbors takes place roughly on a time scale of 10 ns (see snapshots Figure 10d−f and the movie in the Supporting Information). Thus, the propagation speed can be calculated as ~0.6 npn, with the average capsomere distance of ~6 nm, which is similar to the ST in the swollen CCMV capsid during its shrinkage in vacuum. Therefore, STs of viral capsids start with local conformational changes that break icosahedral symmetry (e.g., N-terminal motions of capsid proteins), followed by global structural changes that involve large-scale cooperative translations and rotations of capsid structural units (e.g., capsomeres or protein subunits) through strong allosteric interactions.

Dependence of CCMV Capsid Swelling on System Conditions. The above simulations show that the CCMV capsid swells over tens of nanoseconds. Although there have not been swelling experiments performed on empty CCMV capsids, such STs for

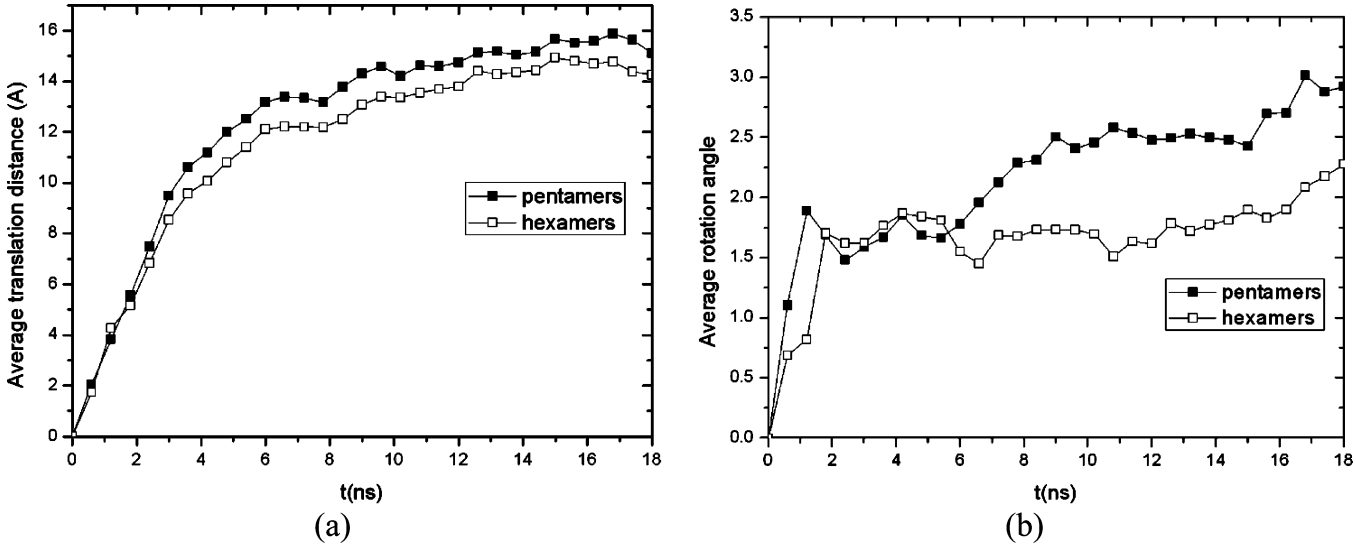

Figure 9. Cooperative motions of pentamers and hexamers during 18.54 ns CCMV capsid swelling: time courses of (a) their average COM translation distances and (b) their average rotation angles calculated through structure fitting to the initial configurations in native CCMV capsid.

the complete virus particles (i.e., capsids embedded with viral RNA) were observed to occur on time scales of milliseconds to minutes in experiments.[55–57] In the following, we present a list of simulations that indicate the swelling of the CCMV capsid is strongly dependent on system conditions. Five different simulations starting from the capsid native state were given in Table 1. The capsid was solvated in a water cube with two different water boundaries, that is, 7 (sim1) and 15 Å (sim2−sim5). The water boundary here is the minimum distance between atoms on the capsid surface and the faces of water cube. Note that sim1 has been used to identify key elements for long-time simulation of the capsid selling as discussed above. There are 522 000 atoms in the CCMV capsid, and the total number of atoms in these systems ranges from ∼2.41 to ∼2.85 million, depending on the thickness of the water boundary and salt concentration used for solvation.

Periodic boundary conditions were applied for all MD simulations (see Computational Methods). Constant volume and temperature control (i.e., the NVT ensemble) was used for sim3, and the others were performed under constant pressure and temperature control (i.e., NPT ensemble). The specified temperature was 298.15 K, and the pressure was 1 atm for NPT simulations. With the above settings, NAMD was used to simulate the five systems to determine the controlling factors of CCMV capsid swelling. It was found that the unit cell for systems prepared with VMD "solvate" plugin by using its default parameters (notably 2.4 Å for the minimum distance between water molecules and the solute) need to be compressed to reach 1 atm pressure in NPT simulations (e.g., sim1 and sim2). Without pressure control, the systems would have a pressure much lower than 1 atm in NVT simulation (e.g., sim3). The analysis of the output trajectories is as follows.

The CCMV capsid was found to swell much faster and more dramatically in sim3 under NVT control (i.e., constant volume, low pressure) than in NPT simulations. Although the capsid swells by 7.61 Å in its outer radius and only 1.58 Å in the average radius during 3 ns NPT simulation in sim2, the capsid as simulated under NVT control swells with the outer and average radius increased by 15.33 and 10.03 Å in 3 ns, and 18.05 and 10.48 Å over 5 ns in sim3. During the 5 ns significant swelling, the capsid thickens by 16.91 Å with the capsid SASA increased from 1.20×10^6 to 1.51×10^6 Å^2; the volume of capsid cavity is increased by 45% from 4.09×10^6 to 5.95×10^6 Å^3 (see calculation details in Computational Methods and Excel spreadsheet the in Supporting Information). In the early stage of the simulation, water molecules flux from the capsid cavity into the capsid to solvate proteins, and small regions of vacuum are formed inside the capsid. As the capsid continues swelling, water molecules flux from the capsid exterior into the capsid cavity. The average density of water inside the capsid cavity decreases significantly from 839.56 to 670.07 kg/m^3 at the end of the 5 ns simulation, whereas the average density of water outside of the capsid stays roughly constant.

When pressure control is applied as in sim1 and sim2, the system is compressed to reach 1 atm pressure, and the swelling rate of the CCMV capsid is decreased. The capsid cavity volume and its SASA are also found to increase accompanying the capsid swelling. The density of water in the capsid cavity decreases during early stage of the simulation due to swelling of the capsid and the flux of water molecules into the capsid to solvate proteins. It then increases as the capsid swelling slows down and the water molecules flux from the capsid exterior into the capsid cavity under the pressure control. In the case of sim2, the average density of water inside the capsid cavity decreases from 886.00 to 841.39 kg/m^3 at 1 ns and then increases to 866.82 kg/m^3 at 3 ns.

To understand the different swelling behaviors of the CCMV capsid found in the NPT and NVT simulations, we compare the system pressure profiles of sim2 and sim3. Because no algorithm has been developed to analyze the radial pressure of a system like CCMV capsid here, we divided the simulation box into horizontal slices of ∼10 Å thickness in the *Z* direction and calculated the local pressure of each slice by using the pressure profile analysis module provided in NAMD[48] for the two simulations. Calculation results showed that the average pressure difference between slices that sandwich CCMV capsid at the bottom and the top (both from the capsid cavity surface to the capsid exterior) is 88 atm at 0.1 ns in sim3, which is ∼13 atm greater that in sim2. Although the pressure profile in NAMD is calculated by following from the algorithm proposed for lipid bilayers[58] and it could lead deviations for the CCMV capsid solvated in a water box here, these calculations suggest the pressure gradient across the capsid in sim3 is higher than that in the latter. As the capsid swells, the pressure of the system

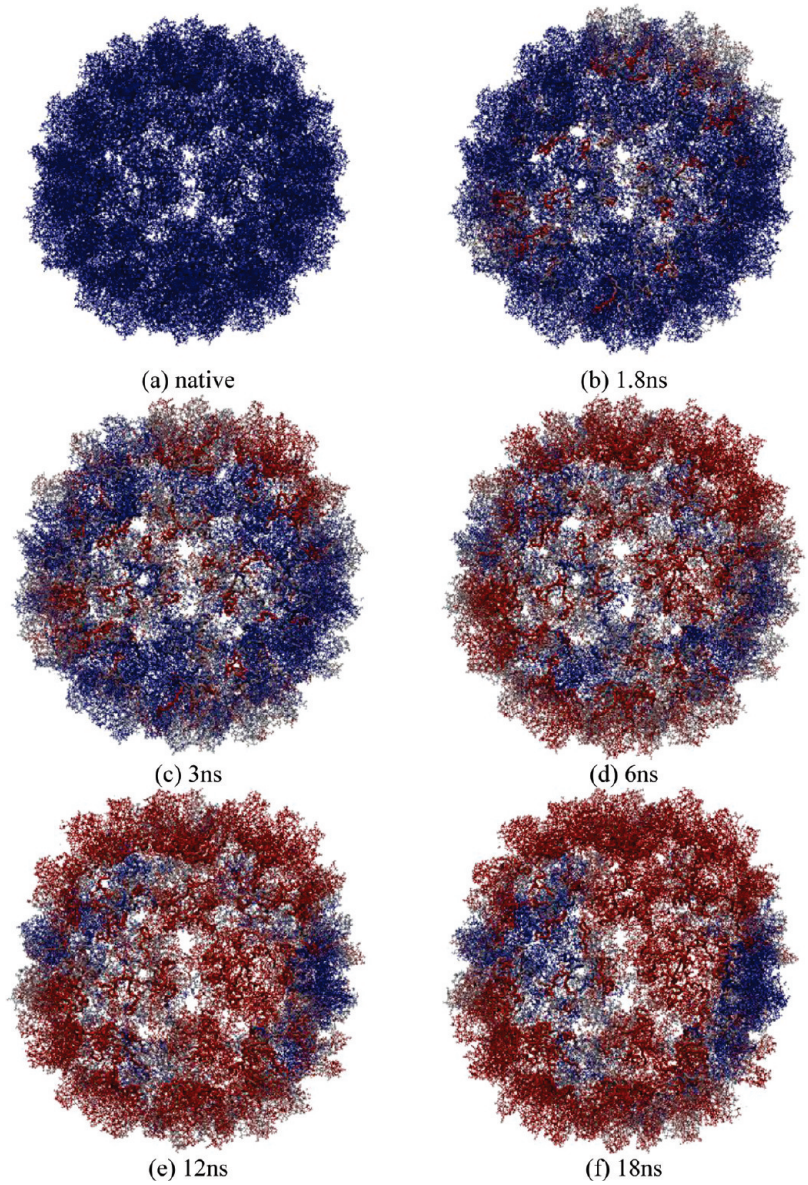

Figure 10. Interior view of the capsid swelling trajectory snapshots with atoms colored by their displacements from the original positions in a BWR color scale (0 Å for blue, 24 Å for red, and white as the midpoint): (a) initial native state with all atoms in blue and (b) 1.8, (c) 3, (d) 6, (e) 12, and (f) 18 ns. N-terminal arms of the capsid proteins are highlighted as ribbons, and they are colored according to the displacements of atoms, as well. They change their color from blue to red during the early stage of the simulation, and their fast fluctuations precede and induce global expansion of the capsid. Swelling of the CCMV capsid is a symmetry-breaking process involving local initiation and front propagation.

slice right below or above the capsid surface (i.e., pressure of water outside of the capsid) keeps increasing. It takes 0.6 ns to decrease the pressure gradient across the capsid to zero in sim2, but 1 ns (i.e., 0.4 ns longer) in sim3. As a result, the capsid swells faster and more dramatically in NVT simulation sim3 than in NPT simulation sim2.

Complete wild-type CCMV virions swell radially by ~10% when the solution pH is raised to 7.0 at low ionic strength ($I < 0.2$ M) in the absence of divalent cations. The transition can be reversed when the pH is lowered from 7.0 to 5.0.[1] The virus stays stable in its native compact form at pH 5.0. Although there has been no swelling experiment performed on empty CCMV capsids, here, we adjust the system pH from 7.0 (sim1 and sim2) to 5.0 and investigate the pH dependence of CCMV capsid stability. Because only the pK_a of the histidine side chain (6.04) lies in this pH range among all amino acids, the side chain of 360 histidines in the CCMV capsid (i.e., two in each protein subunit) are protonated corresponding to the pH change. As a result, the net charge of the CCMV capsid increases from $+900e$ to $+1260e$. The capsid is solvated in a water cube and neutralized with 2282 K^+ and 3542 Cl^- for NPT-ensemble simulation. Trajectory analysis shows that the size of the capsid characterized by its outer radius undergoes a 1.41 Å smaller increase from the initial X-ray crystal structure during 3 ns at pH 5.0 (sim4) than that at pH 7.0 (sim2), although the capsid average radius displays a slightly greater increase (0.36 Å) at pH 5.0 than that obtained at pH 7.0 (see Figures S2 and S3 in the Supporting Information). The capsid diameter is found to increase from 284.69 to 294.43 Å in 2 ns and level off thereafter with moderate fluctuations at pH 5.0. The increase in the capsid diameter is also 1.91 Å smaller than that obtained from the pH 7.0 simulation (see Figure S1 in Supporting Information). This indicates the empty CCMV capsid becomes less likely to swell (more stable) when the pH is lowered from 7.0 to 5.0, which agrees with the experimental results on complete CCMV virions.

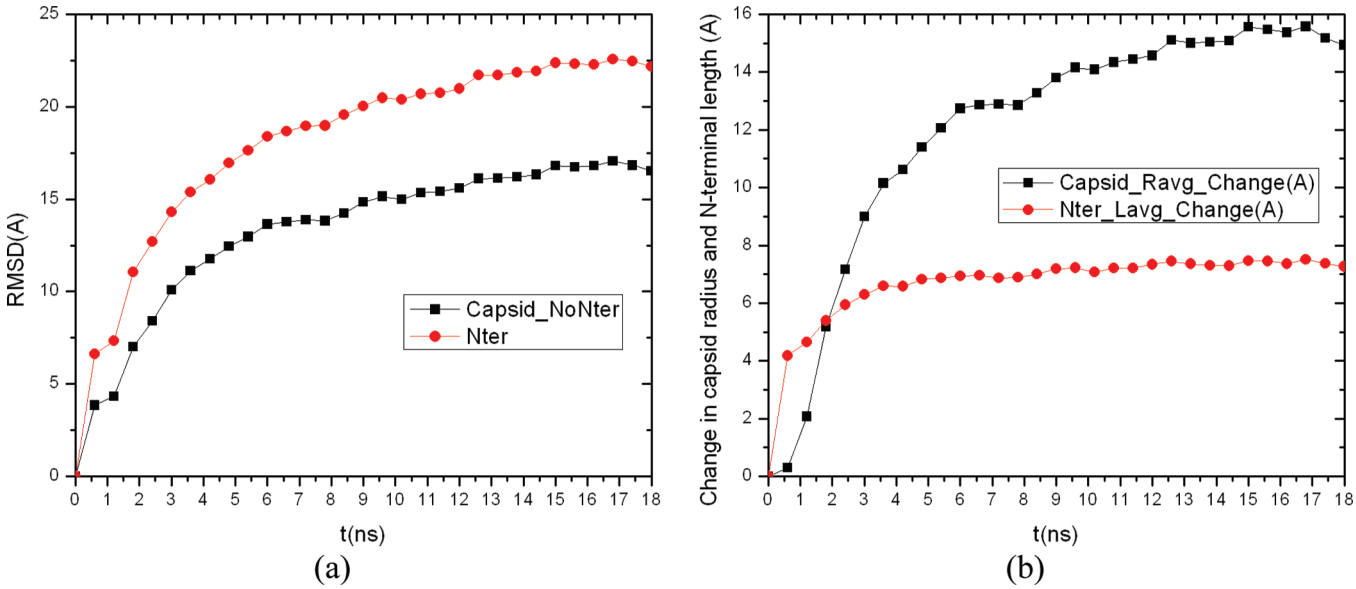

Figure 11. (a) Atomic rmsd between trajectory snapshots and the initial PDB structure for C_α atoms in the N-terminal arms (i.e., the first 26 residues) of capsid proteins as compared with that for the rest of capsid protein backbone and (b) the change in N-terminal arm length as averaged over 180 capsid proteins as compared with the change in capsid average radius along the simulation.

Although the empty CCMV capsid is shown to be more stable at pH 5.0 than that at pH 7.0 in the above analysis, the capsid still displays swelling in the early stage of the simulation. Divalent cations (e.g., Ca^{2+} and Mg^{2+}) are known to bind the CCMV capsid proteins, stabilize the capsid, and eliminate swelling of the complete virus.[1] To examine their effects on the behavior of the empty CCMV capsid, we added 0.05 M $MgCl_2$ to the system prepared in sim4 at pH 5.0 and ran another 3 ns MD simulation on the resulting system, that is, at pH 5.0 and 0.05 M $MgCl_2$. The results show that the capsid is further stabilized, with its diameter increased to only 291.07 Å in 1 ns and leveled off thereafter (Figure S1). A number of Mg^{2+} ions are found to diffuse into the capsid (Figure S4), 51 at 1 ns and 60 (~9% of the total) at 3 ns. Especially at the end of the simulation, 11 out of the 60 Mg^{2+} ions are bound to the divalent cation interprotein binding sites, which are composed of residues Glu81, Gln85, Lys84, and Glu148 from a protein and Gln149* and Asp153* from its adjacent protein.[1,24] A typical example is shown in Figure S5a. Time courses of the distances between the Mg^{2+} ion in the center and its ligand atoms show that the Mg^{2+} ion is diffused into the binding site from bulk water and then bound there (Figure S5b). Such binding of divalent cations to the sites between the capsid proteins stabilizes the capsid against swelling.[1,6,53]

Although the MD simulations performed here are fairly short (due to the size of the CCMV capsid) and longer trajectories are needed to conclude on the capsid stability and swelling dynamics, our results indicate that the swelling of the CCMV capsid in a host medium is strongly dependent on the system conditions (e.g., the pressure gradient across the capsid, pH, and the presence of divalent cations). We show that the empty CCMV capsid becomes more stable as the solution pH is lowered from 7.0 to 5.0 and further stabilized when divalent cations are added to the system. This argues for the correctness of the conditions under which our above long-time MD/OPX simulation was carried out.

In addition to the factors discussed above, discrepancies between the time scale for capsid swelling obtained from our simulations and that for the swelling of the complete virus particles in experiments include the following: (1) The free energy landscape is populated by many minima. Thus, along the pathway connecting the original to the final structure, the system can reside for appreciable time at these minima. In the transition from one minimum to the next, many transition states and associated energy barriers are encountered. (2) The free energy landscape can be so populated with minima that the system is driven by a higher-order entropy, that is, the system loses track of where it is going, although it is driven by an overall energy gradient; this is a complex analog of a two-well potential with a broad, shallow well and a narrow shallow one; in that case, the system could prefer to be in the relatively high energy well. (3) The dynamics of the viral capsid is likely highly frictional; thus, despite a large free-energy gradient, the kinetics can be slow. (4) Attractive capsid−RNA interactions will repress capsid swelling, which requires overcoming the free energy barriers.

IV. Conclusions

Although STs of CCMV (e.g., pH-induced swelling) have been studied since 1975,[3,4,53,59,60] the exact transition mechanisms remain unclear.[5] With new modules added to account for water molecules and ions, MD/OPX is used to simulate the swelling of the native CCMV capsid in a host medium. Insights into the stability of CCMV capsid were first obtained through short-time MD simulation. With essential features identified to allow the swelling of CCMV capsid, MD/OPX simulation was run on the capsid with its long-time dynamics captured. We validate the algorithm whereby the dynamics of capsid OPs is computed by using capsid atomic variables, even though the capsid is immersed in a host medium. Furthermore, this algorithm is based on extrapolation of the capsid configuration, followed by resolvation and rethermalization, and the host fluid dynamics is accounted for indirectly, since it interacts with the capsid as captured explicitly through the short MD run.

MD/OPX simulation results show that the N-terminal arms of capsid proteins undergo large deviations from the initial configurations, with their length extended quickly during the early stage of capsid swelling. Similar to the shrinkage of swollen CCMV capsid in vacuum,[28] swelling of native CCMV

capsid in a host medium is also symmetry-breaking, involving local initiation and front propagation. This suggests that it is not appropriate to use symmetry-constrained models to study the pathways and mechanisms of viral STs.[15,20]

The pH-induced swelling of CCMV has been suggested to be a first-order phase transition at low salinity ($I = 0.2$ M) with hysteresis found in its titration curve.[59] However, divalent cations, such as Mg^{2+} and Ca^{2+}, are known to bind CCMV capsid proteins at the quasi-3-fold axes and, thus, stabilize the capsid and the complete virus.[1,4] By adding magnesium ions to the solution ($[Mg^{2+}] = 0.01$ M), hysteresis in the titration curve of CCMV can be abolished,[3,59] and thus, CCMV swelling becomes fully reversible. This indicates that virus swelling can become a second-order phase transition in the presence of divalent cations. It was also found that the effect of a high concentration in potassium ions (e.g., 0.8 M KCl) is identical to that of magnesium for eliminating hysteresis in the CCMV titration curve. Furthermore, as salinity is increased from 0.2 to 0.3 M, the hysteresis loop size is reduced by a factor of 2. Although the above controlling factors for the swelling of complete CCMV have been determined, those for the empty CCMV capsid swelling are still under investigation.[53]

Our all-atom multiscale simulations can be used to create a phase diagram for CCMV capsid mapping out regions in the space of pH, salinity, concentration of divalent cations, and temperature wherein first- and second-order transitions take place. Because capsid fluctuations are believed to play an important role in vial STs, identifying conditions under which critical fluctuations are realized will be of great interest. The simulations will be used to suggest the experimental conditions wherein large fluctuations and the most dramatic STs occur. From our simulations, swelling of CCMV capsid was found to start locally and then spread across the capsid. However, one might suggest that viral STs occur simultaneously. If this is the case, there should be instantaneous communication of proteins across the viral capsid; for example, as follows from Coulomb interactions. Therefore, decreasing salinity will increase the length scale of effective interactions, which is limited to the Debye length. In this way, we will determine theoretically if there is an observable transition between the nucleation/propagation and the global/simultaneous ST mechanism as salinity changes. This suggests that at low salinity, the system should be highly cooperative and global so that critical fluctuations would be repressed; alternatively, the high salinity system is expected to act more like a two- or three-dimensional system and display critical behavior.

As found through our MD/OPX simulation, the N-terminal arms of CCMV capsid proteins are highly dynamic. They undergo large deviations from the initial configurations with their length extended quickly during the early stage of CCMV capsid swelling. Such motions are found to take place in other viral capsids. They are believed to play important roles in packaging viral genome during virus maturation or engineered nanomaterial synthesis by using viral capsids as molecular containers.[54] For example, the N-terminus of the poliovirus capsid protein VP1 that is entirely internal in the native virion becomes externalized upon cell attachment, and the exposed N-terminus of VP1 was shown to be responsible for liposome binding.[10,61] These conformational changes can be predicted from our simulations and used to guide experimental studies; for example, chemical labeling.[62] If the protein terminal domains are protruding/retracting in a fluctuating manner, then the degree of labeling will increase when such fluctuations are large. Furthermore, the depth of penetration of the labeling should increase as the critical fluctuation intensity increases.

Although the water molecules and ions have been modeled explicitly in the simulation of the CCMV capsid swelling in this study, an implicit solvent approach can be used instead. It will simplify MD/OPX making the workflow similar to the simulation of nanostructures in vacuum. With this, the performance obtained from our previous study on simulating the shrinkage of swollen CCMV capsid for 200 ns is expected.[28] Given the capability of MD/OPX to simulate viral STs with their multiscale features captured, we suggest it will facilitate computer-aided design of antiviral drugs and vaccines and nanocapsules for delivery of therapeutic agents to diseased tissues.

Acknowledgment. This project was supported in part by the National Institutes of Health (NIH) through Stanford University via SimBios, National Science Foundation through the Collaborative Research in Chemistry Program, NIH National Institute of Biomedical Imaging and Bioengineering, and Indiana University College of Arts and Sciences, and the METACyt project through the Center for Cell and Virus Theory. Y. Miao thanks D. Zhang for generously providing the complete CCMV protomer structure and J. Tillotson for help on running jobs on the Indiana University Big Red supercomputer.

Appendix: Construction of Order Parameters from Atomic Configurations.

Consider a nanostructure embedded in a box (a cube for simplicity of presentation) of volume L^3. Basis functions $u_k(x)$ (e.g., orthogonal polynomials or harmonic functions) labeled with integer index k are introduced such that

$$\int_{-L/2}^{L/2} \mathrm{d}x\, u_k(x)\, u_{k'}(x) = \delta_{kk'} \tag{1}$$

for Kronecker delta $\delta_{kk'}$. Composite functions $U_{\underline{k}}(\vec{s})$ are defined such that

$$U_{\underline{k}}(\vec{s}) = u_{k_1}(x)\, u_{k_2}(y)\, u_{k_3}(z) \tag{2}$$

where $\vec{s} = (x, y, z)$ for a box centered at $\vec{s} = \vec{0}$. By using $U_{\underline{k}}(\vec{s})$, system collective modes can be constructed as $\vec{\mathrm{U}}_{\underline{k}}(\vec{s}^0) = \{\vec{\mathrm{U}}_{\underline{k}}(\vec{s}_1^0), \ldots, \vec{\mathrm{U}}_{\underline{k}}(\vec{s}_N^0)\}$ with $\vec{s}^0 = \{\vec{s}_1^0, \cdots, \vec{s}_N^0\}$, and the nanostructure $\vec{s}$ is considered to be deformation of a reference configuration $\vec{s}^0$:

$$\vec{s}_i = \sum_{\underline{k}} U_{\underline{k}}(\vec{s}_i^0)\, \vec{\Psi}_{\underline{k}} + \vec{\sigma}_i \tag{3}$$

where $\vec{\sigma}_i$ is the residual due to finite truncation of the $\underline{k}$ sum and $\vec{\Psi}_{\underline{k}}$ are system order parameters (OPs). Assume m_i is the mass of atom i and $m = 1/N\sum_{i=1}^{N} m_i$. We determine $\vec{\Psi}_{\underline{k}}$ to minimize the mass-weighted root-mean-square residuals (RMSR), that is, $[(1/N)\sum_{i=1}^{N}(m_i/m)\vec{\sigma}_i^2]^{1/2}$, which yields

$$\sum_{k} B_{qk}\vec{\Psi}_{k} = \frac{L^3}{N}\sum_{i=1}^{N}\frac{m_i}{m}U_q(\vec{s}_i^{\,0})\,\vec{s}_i, \qquad B_{qk} = \frac{L^3}{N}\sum_{i=1}^{N}\frac{m_i}{m}U_q(\vec{s}_i^{\,0})\,U_k(\vec{s}_i^{\,0}) \quad (4)$$

The $\vec{\sigma}_i$ contribution is neglected in arriving at this definition of $\vec{\Psi}_k$ as $\vec{\sigma}_i$ fluctuates with i and, hence, with space, whereas the basis functions that capture overall nanostructural features vary smoothly by design. When most of the space in the system is occupied with atoms, the i sum is essentially a Monte Carlo integration. The orthogonality of the normalized basis functions implies that $B_{qk} \approx \delta_{qk}$, and eq 4 can be approximated as

$$\vec{\Psi}_q \approx \frac{L^3}{N}\sum_{i=1}^{N}\frac{m_i}{m}U_q(\vec{s}_i^{\,0})\,\vec{s}_i \quad (5)$$

Therefore, dimensionality reduction for nanostructures from the N-atom configuration to M OPs is achieved through

$$\vec{s}_i, \{i = 1, ..., N\} \overset{\vec{u}_k}{\rightleftharpoons} \vec{\Psi}_k, \qquad \{k = k_1, ..., k_M\} \quad (6)$$

In the above formulation, atomic coordinates of a reference configuration are used to construct the nanostructure collective modes, and their summation given by eq 3 results in a new configuration that the system can evolve to. Instead, atomic displacements calculated from two known configurations of the nanostructure can be used to construct its collective modes, and their summations weighted by OPs would describe system coherent motions directly.

Supporting Information Available: Animation of the MD/OPX simulation trajectory and the simulation data relevant to our discussions in Section III are provided. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Speir, J. A.; et al. Structures of the Native and Swollen Forms of Cowpea Chlorotic Mottle Virus Determined by X-Ray Crystallography and Cryoelectron Microscopy. *Structure* **1995**, *3* (1), 63–78.

(2) Zhang, D. Q.; et al. Electrostatic interaction between RNA and protein capsid in cowpea chlorotic mottle virus simulated by a coarse-grain RNA model and a Monte Carlo approach. *Biopolymers* **2004**, *75* (4), 325–337.

(3) Adolph, K. W. Structural Transitions of Cowpea Chlorotic Mottle Virus. *J. Gen. Virol.* **1975**, (28), 147–154.

(4) Liepold, L. O.; et al. Structural transitions in Cowpea chlorotic mottle virus (CCMV). *Phys. Biol.* **2005**, (2), S166–S172.

(5) Speir, J. A.; et al. Enhanced Local Symmetry Interactions Globally Stabilize a Mutant Virus Capsid That Maintains Infectivity and Capsid Dynamics. *J. Virol.* **2006**, *80* (7), 3582–3981.

(6) Liu, H. J.; et al. Pseudo-atomic models of swollen CCMV from cryo-electron microscopy data. *J. Struct. Biol.* **2003**, *142* (3), 356–363.

(7) Canady, M. A.; et al. Large conformational changes in the maturation of a simple RNA virus, *Nudaurelia capensis ω* virus (NωV). *J. Mol. Biol.* **2000**, (299), 573–584.

(8) Canady, M. A.; Tsuruta, H.; Johnson, J. E. Analysis of Rapid, Large-Scale Protein Quaternary Structural Changes: Time-Resolved X-ray Solution Scattering of *Nudaurelia capensis ω* Virus (NωV) Maturation. *J. Mol. Biol.* **2001**, (311), 803–814.

(9) Wikoff, W. R.; et al. Time-resolved molecular dynamics of bacteriophage HK97 capsid maturation interpreted by electron cryo-microscopy and X-ray crystallography. *J. Struct. Biol.* **2006**, 300–306.

(10) Belnap, D. M.; et al. Molecular tectonic model of virus structural transitions: the putative cell entry states of poliovirus. *J. Virol.* **2000**, *74* (3), 1342–1354.

(11) Hogle, J. M. Poliovirus cell entry: Common structural themes in viral cell entry pathways. *Annu. Rev. Microbiol.* **2002**, (56), 677–702.

(12) Bancroft, J. B.; Hiebert, E. Formation of an Infectious Nucleoprotein from Protein and Nucleic Acid Isolated from a Small Spherical Virus. *Virology* **1967**, *32* (2), 354-&.

(13) Brumfield, S.; et al. Heterologous expression of the modified coat protein of Cowpea chlorotic mottle bromovirus results in the assembly of protein cages with altered architectures and function. *J. Gen. Virol.* **2004**, *85*, 1049–1053.

(14) Freddolino, P. L.; et al. Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure* **2006**, *14* (3), 437–449.

(15) Speelman, B.; Brooks, B. R.; Post, C. B. Molecular Dynamics Simulations of Human Rhinovirus and an Antiviral Compound. *Biophys. J.* **2001**, *80*, 121–129.

(16) Phelps, D. K.; Rossky, P. J.; Post, C. B. Influence of an antiviral compound on the temperature dependence of viral protein flexibility and packing: a molecular dynamics study. *J. Mol. Biol.* **1998**, *276* (2), 331–337.

(17) Isea, R.; Aponte, C.; Cipriani, R. Can the RNA of the cowpea chlorotic mottle virus be released through a channel by means of free diffusion? A test in silico. *Biophys. Chem.* **2004**, *107* (2), 101–106.

(18) Harries, D.; May, S.; Gelbart, W. M.; Ben-Shaul, A. Structure, Stability and Thermodynamics of Lanellar DNA−Lipid Complexes. *Biophys. J.* **1998**, *75*, 159.

(19) Arkhipov, A.; Freddolino, P. L.; Schulten, K. Stability and dynamics of virus capsids described by coarse-grained modeling. *Structure* **2006**, (14), 1767–1777.

(20) Tama, F.; Brooks, C. L. The mechanism and pathway of pH induced swelling in cowpea chlorotic mottle virus. *J. Mol. Biol.* **2002**, *318* (3), 733–747.

(21) Tama, F.; Brooks, C. L. Diversity and identity of mechanical properties of icosahedral viral capsids studied with elastic network normal mode analysis. *J. Mol. Biol.* **2005**, *345* (2), 299–314.

(22) van Vlijmen, H. W. T.; Karplus, M. Normal mode calculations of icosahedral viruses with full dihedral flexibility by use of molecular symmetry. *J. Mol. Biol.* **2005**, *350* (3), 528–542.

(23) Sayyed-Ahmad, A.; Miao, Y.; Ortoleva, P. Poisson-Boltzmann theory of bionanosystems. *Commun. Comput. Phys.* **2008**, *3* (5), 1100–1116.

(24) Konecny, R.; et al. Electrostatic properties of cowpea chlorotic mottle virus and cucumber mosaic virus capsids. *Biopolymers* **2006**, *82* (2), 106–120.

(25) Miao, Y.; Ortoleva, P. All-atom multiscaling and new ensembles for dynamical nanoparticles. *J. Chem. Phys.* **2006**, *125* (4), 44901.

(26) Miao, Y.; Ortoleva, P. J. Molecular dynamics/order parameter extrapolation for bionanosystem simulations. *J. Comput. Chem.* **2008**, *30* (3), 423–437.

(27) Pankavich, S.; et al. Stochastic dynamics of bionanosystems: Multiscale analysis and specialized ensembles. *J. Chem. Phys.* **2008**, *128* (23), 234908.

(28) Miao, Y.; Ortoleva, P. J. Viral structural transition mechanisms revealed by multiscale molecular dynamics/order parameter extrapolation simulation. *Biopolymers* **2009**, *93* (1), 61–73.

(29) Rader, A. J.; Vlad, D. H.; Bahar, I. Maturation dynamics of bacteriophage HK97 capsid. *Structure* **2005**, *13* (3), 413–421.

(30) Einstein, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annal. Phys.* **1905**, (17), 549–560.

(31) Chandrasekhar, S. Dynamical Friction. I. General Considerations: the Coefficient of Dynamical Friction. *Astrophys. J.* **1943**, *97*, 255–262.

(32) Deutch, J. M.; Oppenheim, I. The Concept of Brownian Motion in Modern Statistical Mechanics. *Faraday Discuss. Chem. Soc: London* **1987**, 1–20.

(33) Deutch, J. M.; et al. Light-Scattering from Systems with Chemical Oscillations and Dissipative Structures. *J. Chem. Phys.* **1972**, *57* (10), 4327–4332.

(34) Shea, J. E.; Oppenheim, I. Fokker-Planck equation and Langevin equation for one Brownian particle in a nonequilibrium bath. *J. Phys. Chem.* **1996**, *100* (49), 19035–19042.

(35) Shea, J. E.; Oppenheim, I. Fokker-Planck equation and non-linear hydrodynamic equations of a system of several Brownian particles in a non-equilibrium bath. *Physica A* **1997**, *247* (1−4), 417–443.

(36) Ortoleva, P. J. Nanoparticle dynamics: a multiscale analysis of the Liouville equation. *J. Phys. Chem. B* **2005**, *109* (45), 21258–66.

(37) Peters, M. H. Fokker-Planck equation, molecular friction, and molecular dynamics for Brownian particle transport near external solid surfaces. *J. Stat. Phys.* **1999**, *94* (3−4), 557–586.

(38) Peters, M. H. Fokker-Planck equation and the grand molecular friction tensor for coupled rotational and translational motions of structured Brownian particles near structured surfaces. *J. Chem. Phys.* **1999**, *110* (1), 528–538.

(39) Zwanzig, R., *Nonequilibrium Statistical Mechanics*; Oxford University Press: New York, 2001.

(40) Pankavich, S.; et al. Stochastic Dynamics of Bionanosystems: Multiscale Analysis and Specialized Ensembles. *J. Chem. Phys.* **2008**; 029822.

(41) Shreif, Z.; Ortoleva, P. Curvilinear All-Atom Multiscale (CAM) Theory of Macromolecular Dynamics. *J. Stat. Phys.* **2008**, *130*, 669–685.

(42) Miao, Y.; Ortoleva, P. J. Viral structural transitions: an all-atom multiscale theory. *J. Chem. Phys.* **2006**, *125* (21), 214901.

(43) Kuhn, R. J.; Rossmann, M. G. Structure and assembly of icosahedral enveloped RNA viruses. *Virus Struct. Assem.* **2005**, *64*, 263–284.

(44) Shreif, Z., et al., *Enveloped viruses understood via multiscale simulation: Computer-aided vaccine design*; Springer: Netherlands, 2008; Vol. 15; pp 363−380.

(45) Kevrekidis, I. G.; Gear, C. W.; Hummer, G. Equation-free: The computer-aided analysis of comptex multiscale systems. *AICHE J.* **2004**, *50* (7), 1346–1355.

(46) Hummer, G.; Kevrekidis, I. G. Coarse molecular dynamics of a peptide fragment: Free energy, kinetics, and long-time dynamics computations. *J. Chem. Phys.* **2003**, *118* (23), 10762–10773.

(47) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14* (1), 33–38.

(48) Phillips, J. C.; et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, (26), 1781–1802.

(49) Mackerell, A. D.; et al. Self-Consistent Parameterization of Biomolecules for Molecular Modeling and Condensed Phase Simulations. *FASEB J.* **1992**, *6* (1), A143–A143.

(50) MacKerell, A. D.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102* (18), 3586–3616.

(51) Jorgensen, W. L.; et al. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.

(52) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327–341.

(53) Lavelle, L.; Michel, J. P.; Gingery, M. The disassembly, reassembly and stability of CCMV protein capsids. *J. Virol. Methods* **2007**, *146* (1−2), 311–316.

(54) Douglas, T.; et al. Protein Engineering of a Viral Cage for Constrained Nanomaterials Synthesis. *Adv. Mater.* **2002**, *14* (6), 415–418.

(55) Taylor, D. J.; et al. Correlation of chemical reactivity of *Nudaurelia capensis* omega virus with a pH-induced conformational change. *Chem Commun (Cambridge)* **2003**, (22), 2770–2771.

(56) Lee, K. K.; et al. Evidence that a Local Refolding Event Triggers Maturation of HK97 Bacteriophage Capsid. *J. Mol. Biol.* **2004**, *340* (3), 419–433.

(57) Henzler-Wildman, K.; Kern, D. Dynamic personalities of proteins. *Nature* **2007**, *450* (7172), 964–972.

(58) Lindahl, E.; Edholm, O. Spatial and energetic-entropic decomposition of surface tension in lipid bilayers from molecular dynamics simulations. *J. Chem. Phys.* **2000**, *113* (9), 3882–3893.

(59) Jacrot, B. Studies on the assembly of a spherical plant virus: II. The mechanism of protein aggregation and virus swelling. *J. Mol. Biol.* **1975**, *95* (3), 433–446.

(60) Klug, W. S.; et al. Failure of viral shells. *Phys. Rev. Lett.* **2006**, *97* (22), 228101.

(61) Fricks, C. E.; Hogle, J. M. Cell-Induced Conformational Change in Poliovirus−Externalization of the Amino Terminus of Vp1 Is Responsible for Liposome Binding. *J. Virol.* **1990**, *64* (5), 1934–1945.

(62) Beardsley, R. L.; Running, W. E.; Reilly, J. P. Probing the struture of Caulobacter crescentus ribosome with chemical labeling and mass spectrometry. *J. Proteome Res.* **2006**, *5*, 2935–2946.

JP102314E