Structure 14, 1767–1777, December 2006 ©2006 Elsevier Ltd All rights reserved DOI 10.1016/j.str.2006.10.003

Stability and Dynamics of Virus Capsids Described by Coarse-Grained Modeling

Anton Arkhipov,[1,2] Peter L. Freddolino,[1,3] and Klaus Schulten[1,2,3,*]

[1] Beckman Institute for Advanced Science and Technology

[2] Department of Physics

[3] Center for Biophysics and Computational Biology

University of Illinois at Urbana-Champaign

Urbana, Illinois 61801

Summary

We report a study of the structural dynamics of viral capsids, simulated on a microsecond timescale, by employing a coarse-graining molecular dynamics method. The method was calibrated against an all-atom simulation of one complete virus. Among the studied capsids, some collapsed rapidly, while others were found to be stable. Interlocking between coat proteins is found to be a key factor determining the stability of the capsids.

Introduction

Viral capsids are multiprotein assemblies that serve as shells for the genetic material and replicative machinery of viruses (Fields et al., 1996; Flint et al., 2004; Levine, 1991; Bamford et al., 2005). Viral capsids vary in shape, composition, and size but are always built from multiple copies of a few distinct proteins (sometimes a single protein), usually arranged in a symmetrical fashion. Many viral capsids are icosahedral, in which case the capsid is formed by repeated tiling of an asymmetric unit (made of one or several proteins) on the faces of the icosahedron. Icosahedral capsids are classified by their T-number (Caspar and Klug, 1962), which is the number of identical protein conformations found in a single asymmetric unit.

One of the most important roles of a capsid is to protect its contents, but the capsid also needs to become unstable to release the virulence factors into the host cell. Thus, the stability of the capsid and the transitions between stable and unstable structures are key issues in the life cycle of a virus. Many viral capsids are known structurally (see, for example, the VIPER database; Shepherd et al., 2006), and recent advances in cryo-electron microscopy (Lander et al., 2006; Jiang et al., 2006) have made it possible to visualize nonsymmetric elements of capsids. However, a static structure does not reveal dynamical properties. Unfortunately, in experiments studying capsid dynamics in real time, the spatial and temporal resolution are usually too low to distinguish movements of individual capsid proteins.

Advances in numerical computations have made it possible to simulate parts of viruses (Phelps et al., 2000; Reddy et al., 1998; Perryman et al., 2004; Speelman et al., 2001) and, recently, even a complete virus (Freddolino et al., 2006), with single-atom resolution. All-atom modeling is currently limited to a 10 nm size and a nanosecond duration. However, the size of most viral capsids is much larger than 10 nm, and structural transitions of interest occur on millisecond (or longer) timescales. Thus, there is a need to develop a new computational technique that allows one to cover the required length and timescales. An analogous mismatch between computational capabilities and in vivo length and timescales is also found in many other macromolecular assemblies of biological interest, e.g., the cytoskeleton, motor proteins, and bacterial flagella.

A general approach to overcome the spatial and timescale limitations in computational modeling is to reduce the number of degrees of freedom covered in a simulation. One has to both identify the "important" degrees of freedom and derive the equations determining the evolution of the system when only these are taken into account. Such an approach allows one to describe larger systems over longer time intervals, but this comes at the cost of reduced accuracy. One solution involves so-called "coarse graining" and has been recently applied to the study of viral capsids (Rader et al., 2005; Tama and Brooks, 2005; Tama et al., 2004; Konecny et al., 2006; LaMarque et al., 2004; Zhang et al., 2004; Arsuaga et al., 2002), ribosomes (Wang et al., 2004; Trylska et al., 2005), individual protein chains (Tozzini, 2005; Tozzini and McCammon, 2005; Bahar et al., 1997), and actin filaments (Chu and Voth, 2005), as well as lipid and protein-lipid aggregates (Shelley et al., 2001; Marrink et al., 2004; Bond and Sansom, 2006). Building upon these advances, as well as our own experience in simulations of a virus (Freddolino et al., 2006) and coarse graining of proteins and lipids (Shih et al., 2006a, 2006b), we have developed a new, to our knowledge, coarse-grained (CG) molecular dynamics (MD) method and have applied it to the study of the structural dynamics of several icosahedral viral capsids.

Due to the computational challenges posed by viral systems, most simulations performed so far are of a small section of a capsid with icosahedral symmetry imposed (Phelps et al., 2000; Speelman et al., 2001) and, thus, are not applicable to the study of large-scale intersubunit motions. Previous coarse-graining efforts on virus capsids (Rader et al., 2005; Tama and Brooks, 2005; Konecny et al., 2006) have focused on reproducing very special properties of the systems, but they did not simulate long-term viral dynamics. The work of Zhang et al. (2004) involved an electrostatic calculation with only the nucleic acid coarse grained (and only then because the complete structure was missing), whereas Tama and Brooks (2005) and Rader et al. (2005), although addressing interesting issues of the flexibility of viral capsids or the maturation of the bacteriophage HK97 capsid, have used elastic network modeling rather than true CG dynamics and, thus, were only able to illustrate normal modes and interpolate between existing structures, rather than simulate dynamics.

The CG model presented here overcomes the limitations of previous approaches used in CG simulations

*Correspondence: kschulte@ks.uiuc.edu

Table 1. Simulated Systems

Simulation Name	Description (Reference; PDBcode)	T/N^{P}_{assym}	NCG/D_{max}	Δt (fs)	t_{sim} (μs)	Result
simSTMV	STMV capsid (Larson et al., 1998; 1A34)	1/1	675/18 nm	250	5	Collapse
simSTMVfull	FullSTMV virion (Larson et al., 1998; 1A34)	1/1	892/18 nm	200	5	Stable
simSPMV	SPMV capsid (Ban and McPherson, 1995; 1STM)	1/1	600/17 nm	250	25	Collapse
simSTNV	STNV capsid (Jones and Liljas, 1984; 2BUK)	1/1	852/20 nm	250	7	Collapse
simBMV	BMV capsid (Lucas et al., 2002; 1JS9)	3/3	2,541/28 nm	170	5	Stable
simBMVcut	BMV capsid, N-terminal tails cut away (Lucas et al., 2002)	3/3	1,929/28 nm	200	5	Collapse
simPV	Poliovirus capsid (1HXS)	1/4	4,518/33 nm	170	11	Stable
simB174	Bacteriophage ϕX174 procapsid (Dokland et al., 1999; 1CD3)	1/7	6,006/36 nm	200	5	Stable
simRV	Reovirus core (Reinisch et al., 2000; 1EJ6)	1/8	20,883/75 nm	120	1.5	Stable

The names of simulations are given in the left column. "T" is the T-number for an icosahedral capsid, "N^{P}_{assym}" is the number of proteins per asymmetric unit of a capsid, "NCG" is the number of CG particles used to simulate a given system, "D_{max}" is the maximum diameter of a capsid in all-atom representation, "Δt" is the time step used, and "t_{sim}" is the total simulation time.

of viruses by employing a flexible algorithm for representing the system with a very small number of CG particles, and then by using a true dynamical description of the CG system. In this model, the shape of a protein is reproduced by several point masses (about 200 atoms are represented by a single particle), which interact through effective potentials. This level of coarse graining allows us to reach the microsecond timescale. The CG simulations of the viral capsids reveal a range of flexibility among the studied systems; for example, some of the capsids collapse when simulated without their genetic material. The simulations provide a wealth of previously unavailable information on the dynamics of viral capsids and appear to generate realistic insight, as our results agree well with the data from all-atom simulations of a complete virus (Freddolino et al., 2006).

There are many macromolecular systems other than the viral capsids whose dynamical properties remain largely unknown which could be investigated by CG techniques. Our CG method has been developed in a general form, such that it can be applied readily to other systems, provided that the structure of the studied complex is available.

Results and Discussion

We have studied the stability of several viral capsids by employing a novel, to our knowledge, CG MD method (see Experimental Procedures). The diameters of the simulated capsids were in the range of 10–100 nm; the dynamics was investigated over time intervals of 1.5–25 μs.

Simulated Systems

The investigated systems are listed in Table 1. All structures were obtained from the VIPER database (Shepherd et al., 2006; Natarajan et al., 2005) and were coarse grained as described in Experimental Procedures (the satellite tobacco mosaic virus RNA structure [Larson et al., 1998; Larson and McPherson, 2001] was taken from the model used in our previous all-atom simulation of this virus, as described in Freddolino et al. [2006]). The CG scheme is illustrated in Figure 1. In the CG model, the unit proteins are represented by CG beads, with an average ratio of 200 atoms per bead; thus, the shape of the protein is roughly preserved, and the CG subunits behave as elastic objects, with Coulombic and Lennard-Jones-like interactions between the separate subunits (see Experimental Procedures).

The choice of the viruses for the study was made on the basis of structure availability and potential interest in terms of capsid stability. First, we chose the three satellite plant viruses (STMV) (Dodds, 1998; Larson et al., 1998; Routh et al., 1995), the satellite panicum mosaic virus (SPMV) (Ban and McPherson, 1995; Day et al., 1994), and the satellite tobacco necrosis virus (STNV) (Jones and Liljas, 1984; Liljas et al., 1982). These are small ($\sim$20 nm in diameter) viruses, each consisting of a short ($\sim$1000 nucleotides) RNA genome packed inside a T=1 capsid composed of 60 identical proteins (Ban et al., 1995). They produce no infection on their own, but rely on the preinfection of a plant by their respective helper viruses (Francki, 1985; Valverde et al., 1991). The complete virions of STMV, SPMV, and STNV are very stable structures; however, the empty capsids have not ever been observed. This suggests that these capsids are unable to maintain their shape without the support of the RNA core, although this has not been proven. We found that all three capsids collapsed, following pathways similar to the one observed for STMV in the all-atom simulations (Freddolino et al., 2006).

Our next choice was the brome mosaic virus (BMV) (Lucas et al., 2001, 2002), a T=3 plant RNA virus $\sim$28 nm in diameter that is composed of 180 identical protein units, with 60 units in each of the 3 distinct conformations. Empty BMV capsids are not observed in vivo (although they have been created in vitro), and the unit proteins seem to interact weakly with each other (Lucas et al., 2002). Numerous experiments illustrated that BMV disassembly is a multistep process culminating in cleavage of the N-terminal tails of the capsid proteins (Larson et al., 2005; Lucas et al., 2001; Cuillel et al., 1981; Bancroft et al., 1967; Pfeiffer and Hirth, 1974), at which point the capsid falls apart or restructures into various possible conformations, including a T=1 form. We simulated the BMV capsid and found that, even without the RNA, it remained stable, although flexible. To mimic the proteolytic cleavage, we built a CG model of the BMV capsid with the N-terminal tails removed at residue 57, in line with experimental observations on the cleavage location (Lucas et al., 2002). In the resulting simulation (simBMVcut), the capsid collapsed over a period of 5 μs.

Finally, three larger and more complex assemblies were simulated: the poliovirus capsid, the bacteriophage

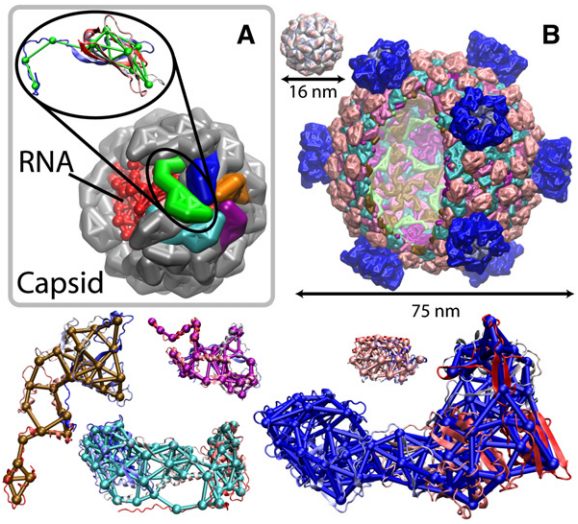

Figure 1. Coarse-Graining Scheme

(A) Complete STMV particle in the CG representation (simulation simSTMVfull). The capsid (gray) consists of 60 identical proteins; 5 proteins around the 5-fold symmetry axis are shown in various colors. Part of the capsid is removed to demonstrate the RNA core of the virus (red). The inset shows a unit protein of STMV in both the all-atom and the CG (green) representations.

(B) The CG model of the reovirus core (partially open to show the inside), the largest system simulated (simRV). The STMV particle (upper left) is shown to scale for comparison. At the bottom, several proteins (drawn to scale) from the reovirus core are presented (all-atom versus CG, the coloring of the CG models corresponds to the colors used in the snapshot of the full particle).

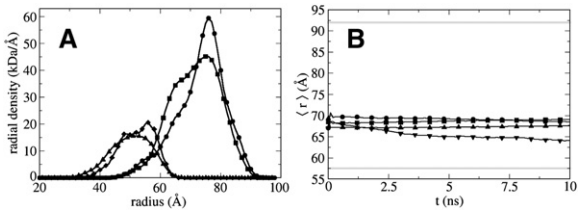

Figure 2. Comparison between Simulations of STMV in the All-Atom and CG Representations

(A) Radial distribution of mass in the complete STMV particle. Mass distribution of the protein is denoted by circles (CG) and squares (all-atom); mass distribution of RNA is denoted by diamonds (CG) and triangles (all-atom).

(B) Average radii, <r>, of the capsid as functions of time in the simulations of a complete STMV (circles for CG, squares for the all-atom) and of an empty STMV capsid (upside-down triangles for CG, right-side-up triangles for all-atom). The approximate outer and inner edges of the capsid in the crystal structure are designated by two gray, horizontal lines.

ϕX174 procapsid, and the reovirus core (an internal capsid particle, or ICP). Poliovirus and reovirus infect humans, whereas bacteriophage ϕX174 infects bacteria.

The poliovirus capsid (Miller et al., 1996; PDB code: 1HXS; Hogle et al., 1985, 1987; Wien et al., 1997; Filman et al., 1989) is a T=1 icosahedron ~33 nm in diameter and is composed of 60 copies of 4 different polypeptides. Although infectious particles form only when the RNA genome is packed in the capsid, empty capsids can and do form in vivo (Flint et al., 2004). Therefore, one expects that the empty poliovirus capsid taken from the crystal structure of the virus is stable, but the possibility of a serious rearrangement of unit proteins to account for the lack of RNA cannot be excluded.

The bacteriophage ϕX174 procapsid (Dokland et al., 1999) is a precursor particle formed during the assembly of the virus. The procapsid includes the proteins that form the mature capsid as well as additional proteins forming the scaffold around the capsid; the scaffold proteins leave the procapsid during and after the packing of the single-stranded DNA genome, which leads to the maturation of the particle and formation of an infectious virion (Dokland et al., 1997; Uchiyama and Fane, 2005; Ilag et al., 1995; McKenna et al., 1992, 1994). The bacteriophage ϕX174 procapsid is a T=1 icosahedron that is ~36 nm in diameter. It is built up of 4 proteins and contains 60 copies of 3 of them and 240 copies of the fourth (in 4 distinct conformations). As the procapsid forms on its own and serves to facilitate genome packing, it is expected to be stable.

The reovirus core (Reinisch et al., 2000) is one of the largest viral structures available, and it is ~75 nm in diameter; it is a T=1 structure composed of five different proteins (Nibert et al., 1996). Altogether, 480 protein units comprise the particle. This assembly represents a biochemical machine performing the synthesis, modification, and export of the viral messenger RNA, and it is known to be a stable particle (Fields, 1996).

CG simulations of the reovirus core, the bacteriophage ϕX174 procapsid, and the poliovirus capsid demonstrated that these are indeed stable structures.

Comparison with All-Atom Studies

Previously, we reported the first, to our knowledge, all-atom MD study of a complete virus (Freddolino et al., 2006), STMV. Dynamics of the complete virus, the capsid without the RNA, and the RNA core alone were simulated on a timescale of 10 ns, and we found that while the complete STMV and the RNA core were stable, the capsid without RNA collapsed rapidly. We compared these data with the results of our CG simulations of the same system to test the quality of our CG model (Figures 2 and 3 and Table 2; also see the Supplemental Data available with this article online). Below we refer to the capsid without the RNA as "empty capsid," although the capsid in its all-atom description is completely filled with water (see Freddolino et al., 2006).

In Figure 2A, the radial distribution of mass of the complete STMV particle is shown. The presented distribution is an average over 100 snapshots of the system (one snapshot per 0.1 ns) from the first 10 ns of either the CG or all-atom simulation. The particle's structure and motion are slightly affected by the coarse graining, as the peak of the protein mass distribution is higher for the CG simulation, indicating that the STMV capsid proteins are shifted to the periphery of the particle. However, the shapes of the CG and all-atom distributions are similar, and for a significant part of the sampled radii the distributions overlap closely. Most notably, the CG distributions are essentially the same as the all-atom ones in the vicinity of the capsid or RNA core edges, including the part of the protein density that overlaps with the RNA density (this protein density corresponds mainly to the tails of the capsid monomers).

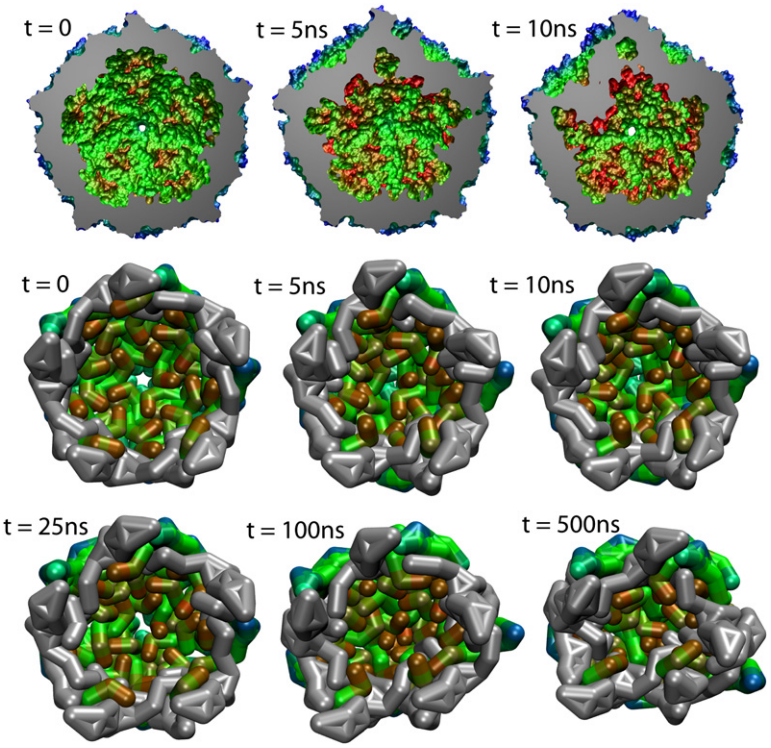

Figure 3. Collapse of the Empty STMV Capsid in the All-Atom and CG Simulations

The three snapshots on the top are from the all-atom simulations (the simulation was 10 ns long), and the six on the bottom are from the CG simulation. Coloring is according to the distance from the center (red for the closest, blue for the most distant), and the regions of the capsid that are closest to the cutting plane are drawn in gray. At the earliest stage of the collapse (first 10 ns), in both all-atom and CG simulations, two groups of trimers on the opposite sides of a 2-fold symmetry axis are shifting toward the center. Further deformation of the capsid involves significant movements of other parts of the capsid (as the CG simulation suggests, bottom row).

In Figure 2B, the average distance, <*r*>, between the center of the STMV particle and the centers of mass of the protein monomers is presented as a function of time. For both the complete STMV and the empty capsid, the difference between <*r*> in the CG or all-atom simulations does not exceed 4 Å, which should be compared to the thickness of the capsid (35 Å, denoted by the gray lines in Figure 2B). The decrease of <*r*> in the simulation of an empty capsid is due to the collapse mentioned above. The radial root-mean-square deviation (rmsd) and minimal distance, r_{min}, between the subunit centers and the center of the capsid, presented in Table 2, are also found to be in satisfactory agreement between the two levels of description, given the resolution difference. The employed coarse-graining method replaces an STMV unit protein by only 11 beads, which by itself changes the distribution of mass to some extent. To take this into account, we present the values of r_{min} for both the initial structures (t = 0, these are the structures obtained after the minimization) and the structures obtained after 5 ns of dynamics (Table 2, also see Figure 2B for comparison of values of <*r*> at different points in time). One finds a 1–2 Å difference in <*r*> and r_{min} between the CG and all-atom structures; the difference remains the same for the complete virus and increases up to several Å for the empty capsid after several nanoseconds. The radial rmsd in the case of the complete virus is higher for the CG simulation, while for the empty capsid this value is essentially the same in both representations. Judging by the presented numbers, the dynamics of the CG structure follows the trends observed for the all-atom structure, with deviations of a few Å. As the radial rmsd values show, the mobility of the protein subunits is somewhat overestimated by the CG model, but the overall structural arrangement remains the same.

Table 2. Motions of the Subunits' Centers of Mass in the All-Atom and CG Simulations of STMV

Quantity	Complete, All-Atom	Complete, CG	Capsid, All-Atom	Capsid, CG
r_{min} (t = 0 ns) (Å)	67.6	65.8	68.2	66.0
r_{min} (t = 5 ns) (Å)	67.2	65.2	50.4	58.4
Radial rmsd (t = 5 ns) (Å)	0.82	2.26	3.68	3.47

For each STMV protein unit, the distance between its center of mass and the center of the particle is computed, and the rmsd of this value is shown (radial rmsd), together with the minimal distance, r_{min}, at times t = 0 and t = 5 ns. Average distances are shown in Figure 2B.

Thus, it appears that the overall shape and size of the complete virion are reproduced well with the CG model. We also compare CG and all-atom results in a case in which large-scale motions are involved—the collapsing empty STMV capsid. As shown in Figure 3, the CG model reproduces the pattern of the collapse and the timescale on which it occurs. The all-atom simulation reveals that the collapse involves a shifting of two trimers of unit proteins toward the center of the capsid, the trimers being located on opposite sides of a 2-fold symmetry axis (top three snapshots in Figure 3). After a rapid initial shift (first 5 ns), the trimers slow down, but they continue to slide toward the center (5–10 ns). In three independent all-atom simulations, it was found that the pattern of the collapse is always the same, but the part of the initially symmetric capsid, where the collapse occurs, is different. The same behavior was observed in three CG simulations of the empty capsid: the collapse started from the shifting of two trimers toward the center

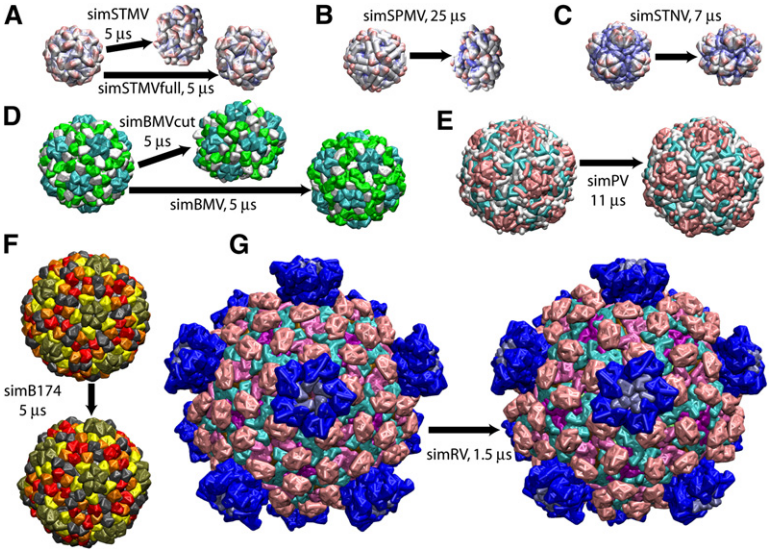

Figure 4. Dynamics of the Capsids

(A–G) The initial and final structures for each simulation are shown (all particles are drawn to scale). (A) The STMV capsid collapses when simulated without RNA (simulation simSTMV), but the full virion is stable (simSTMVfull). (B and C) The SPMV and STNV capsids collapse. (D) The BMV capsid is highly flexible, but stable (simBMV); however, when the interlocking N-terminal tails of the unit proteins are removed, the capsid collapses (simBMVcut). (E–G) (E) The poliovirus capsid, (F) the procapsid of the bacteriophage ϕX174, and the (G) reovirus core are stable structures.

(central row in Figure 3), and in each simulation this happened at different locations on the capsid. Also, for the CG and all-atom structures in Figure 3, the initial shifting of trimers occurs at different spots on the capsid, though the visual orientation in the two cases shown is chosen such that the collapsing trimers are at the top.

With the CG model, one can simulate dynamics over much longer times than is possible with an all-atom representation. The longest CG simulation of the empty STMV capsid covered 5 μs, and the results suggest a possible scenario for the collapse after the first 10 ns. The bottom row of snapshots in Figure 3 shows that after the initial shifting of trimers, the structure is generally stable for tens of nanoseconds, but then other parts of the capsid slide toward the center one after another; the sliding events are separated by tens and hundreds of nanoseconds. It was already observed in the all-atom simulation that the speed of the collapse is not constant. Dynamics of the CG system featured the same behavior, but while in the all-atom study only one structural transition was observed (initial shifting of trimers), several transitions were observed in the CG simulation. It seems that the collapsing capsid goes through a set of metastable states. Transitions between each of the states occur relatively quickly, in nanoseconds, while the time spent by the system in each metastable state can be much longer, namely, hundreds of nanoseconds. Our CG simulations of capsids of other viruses (see below) suggest that this behavior is specific to the particular virus (also see movies in Supplemental Data).

Because of the level of reduction employed in the CG model, one can only aim to reproduce coarse properties, such as the size and shape of a macromolecular assembly, and modes of large-scale motions. Despite these shortcomings, however, CG simulations reveal rather detailed dynamics of very large systems and do so on timescales of microseconds. As the comparison presented above shows, the shape, size, and peculiarities of movements are reproduced by the CG model in the case of STMV, with a maximum difference of several Å in the positions of the protein subunits. Since each CG bead represents a part of the protein 10–20 Å in size, this deviation does not seem particularly substantial. The comparison of the collapse stages for the empty capsid demonstrates that the CG model also reproduces the characteristic times of the dynamics, suggesting that the timescale of motions observed in the simulations (described below) is realistic.

Capsid Dynamics for Coarse-Grained Description

Results of the CG simulations are presented in Figure 4 (also see Supplemental Data). For each simulation, the initial state is shown, with an arrow leading toward an image showing the final state. For STMV (Figure 4A), we performed simulations of the empty capsid (simSTMV) and of the complete virus (simSTMVfull). As expected according to previous all-atom simulations (Freddolino et al., 2006), the empty STMV capsid collapsed, and the full virion was stable.

The SPMV capsid was unstable as well (Figure 4B). Over the first 5 μs, the capsid became somewhat disordered, but it maintained its size well. However, subsequently, the total collapse occurred in a few sharp steps. At each step, one or a few unit proteins slipped toward the center, which produced a rapid shift in the arrangement of neighboring subunits. This scenario is somewhat different from the STMV collapse, which occurred more smoothly. The final SPMV capsid structure was disordered and flat; the internal volume was almost filled up by the unit proteins. The STNV capsid also collapsed (Figure 4C), in yet another manner: in the very beginning of the simulation (~200 ns), the coat proteins shifted radially and almost uniformly toward the center. The resulting structure was asymmetric and significantly squeezed in comparison with the initial one, although the arrangement of unit proteins did not change much; the internal volume shrank to less than half of its initial value. After the rapid initial implosion, the structure remained the same throughout the rest of the simulation.

The complete capsid of BMV (Figure 4D) was stable overall, but it was very flexible. Empty BMV capsids have been observed (not in vivo), which means that this structure may remain stable under certain

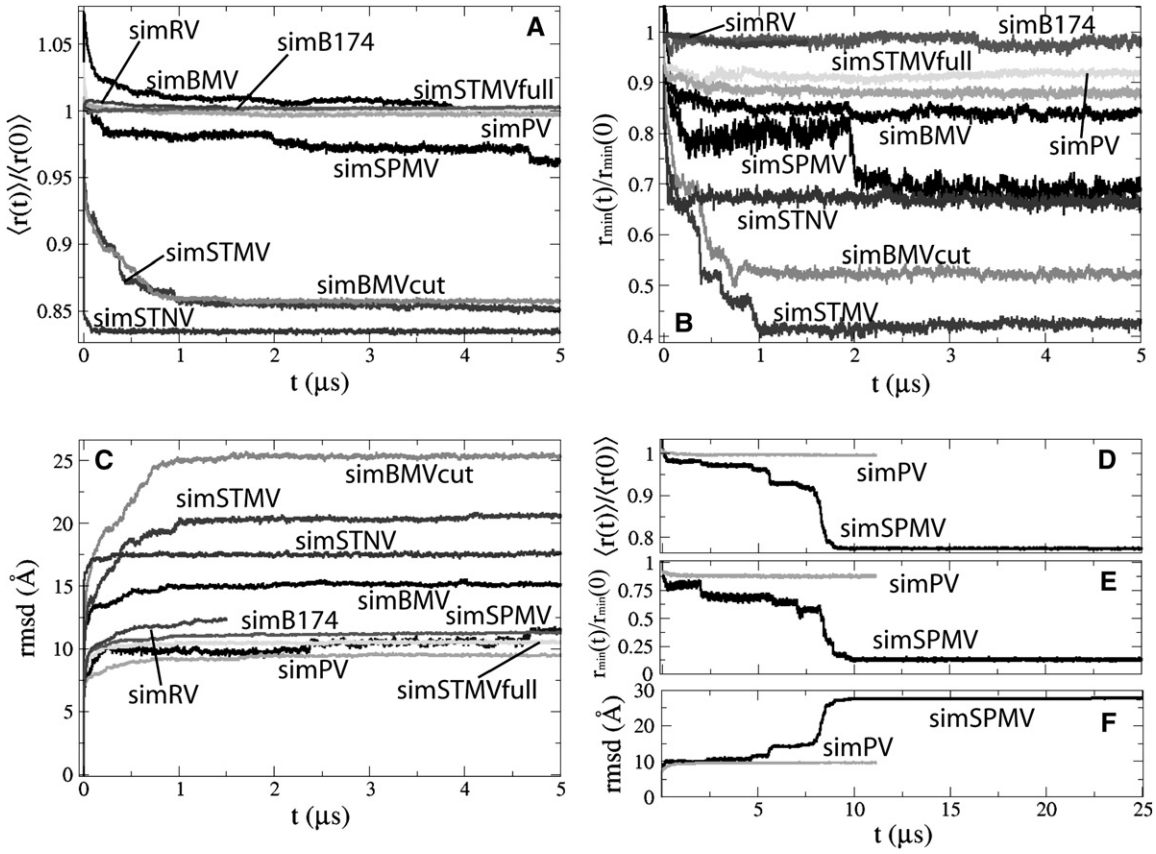

Figure 5. Quantitative Characterization of the Stability of Simulated Systems

(A–F) The average radius, $\langle r \rangle$, of the capsid; the minimal radius, r_{min}; and the rmsd per CG particle are shown in (A), (B), and (C), respectively, as functions of time, t, for each of the simulations. The evolution of these quantities over longer time periods (for simulations simPV and simSPMV, which were run for up to 11 and 25 μs, respectively) is shown in (D), (E), and (F). The radii, $\langle r \rangle$, and r_{min} are divided by their initial (t = 0) values, i.e., for the stable systems, the values of $\langle r(t) \rangle / \langle r(0) \rangle$ and $r_{min}(t)/r_{min}(0)$ remain around unity over the simulation time, while for the unstable systems, these values drop significantly as the capsids collapse.

conditions, and results of our simulation agree with this fact (Larson et al., 2005). It is possible that the flexibility of the BMV capsid is a key property that allows the capsid to swell at elevated salt concentrations, which makes the particle sensitive to protease and RNase. Subsequent cleavage of the protein tails may bring the system to a state similar to the one in the beginning of simulation simBMVcut, and our modeling showed that such a structure is unstable. In simulation simBMVcut, the initially spherical capsid was brought into a discoidal shape, with a ratio of the short dimension to the long one of ~0.6.

The capsids in simulations simPV, simB174, and simRV remained very stable (Figures 4E–4G). For each of the stable structures (including the complete BMV capsid), the unit proteins moved asymmetrically and not negligibly around their initial positions (this can be noticed best in Figure 4 near the 5-fold symmetry axes of the capsids; the proteins around these axes move slightly away from the initially symmetric arrangement). We also noticed that the stable capsids tended to become less spherical, making their icosahedral edges sharper; this minute but noticeable modification happened over a very short initial time interval of a few nanoseconds. This behavior could be due to (1) imprecisions in our model (although it does agree with experimental observations wherever comparison is possible); (2) the fact that the crystal structures were solved with icosahedral symmetry imposed, whereas no such restraints are present in silico or in vivo; or (3) some other difference between crystallographic and physiological conditions.

Quantitative Characterization of the Dynamics

In Figure 5, we present the average and minimum radii of the capsids (calculated as an average distance and minimal distance, respectively, between any of the CG beads and the center of mass), as well as the rmsd per CG bead, as functions of time. For ϕX174 and reovirus, the radii do not change appreciably during the simulations. The minimal radii of the STMV capsid (when the RNA core is present) and of the BMV and poliovirus capsids drop slightly as the coat proteins rearrange during the initial stage of the simulations. The minimal radius of the BMV capsid drops below the values for all other stable capsids, and it exhibits more noise throughout the simulation than the others. Also, the average radius of the BMV capsid increases by almost 10% over the first few nanoseconds, but then resumes the initial value. Apparently, the initial structure was too tight, and the capsid extended to compensate for this;

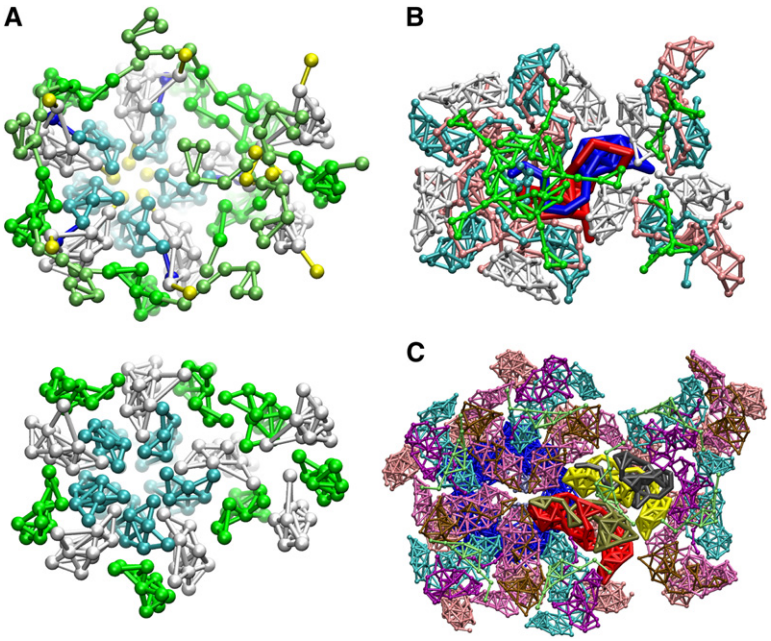

Figure 6. Interlocking of Unit Proteins Enhancing Stability of a Capsid

(A–C) Proteins around the 5-fold and 3-fold symmetry axes of the (A) BMV capsid, (B) poliovirus capsid, and (C) reovirus core are shown (as viewed from inside the capsids). In (A), the BMV capsid (top) is compared to one for which the proteins' tails are cut (bottom; simulation simBMVcut). The tails missing in the bottom image are color coded in the top image: tails of green (white, cyan) proteins are lime (yellow, blue). In (B) and (C), a few protein units are highlighted in color; the licorice representation is used to demonstrate the interlocking.

subsequently, some of the unit proteins rearranged slightly, which allowed the capsid to shrink back to its initial size.

The collapses of the unstable capsids follow slightly different pathways in each case, as demonstrated by the radius plots shown in Figure 5. The STNV capsid shrinks rapidly (first 200 ns) into a conformation that remains stable as indicated in Figure 5. The SPMV capsid collapse proceeds in steps, as can be discerned in Figure 5, through long plateaus in the $\langle r(t) \rangle$ and $r_{min}(t)$ curves with sharp drops in between. The most significant deformation happens between 5 and 10 μs (Figures 5D and 5E), when the minimal radius drops almost down to zero. The collapse of the STMV and BMV capsids occurs in a similar fashion: the unit proteins fall toward the center in a series of steps during about 1 μs, until a metastable state is reached; the structure is then strongly deformed, but it does not change much further.

All stable capsids have a similar rmsd value of ~1 nm, which is reached after about 300 ns and then remains constant. The BMV capsid rmsd is 1.5 times higher. The rmsd of all the unstable capsids is higher than that of any of the stable ones; the BMV capsid with the protein tails removed exhibits the largest rmsd, which might be just due to the fact that it is the largest of the collapsing capsids. Since the CG beads represent the centers of mass of groups of atoms, the rmsd for the CG beads, although not directly comparable to the rmsd of the atoms, nevertheless represents a crude estimate for atomic rmsd. The rmsd for stable capsids, about 1 nm, is somewhat high by the standards of structural analysis, but in this regard one needs to take into account the coarse-graining attribute. Most important, of course, is not the actual value, but the fact that the rmsd values level out for stable structures.

Although the capsids of the small plant viruses in our study collapsed into tighter structures, this should not be taken as a prediction of the behavior of these proteins in nucleic acid-free solution, because the starting conformation is not accessible to these capsid proteins in the absence of the genome. The study of the aggregation of capsid proteins in the presence and absence of genomic material is, however, a likely future application for the methods described here, extending prior theoretical studies as reported by Rapaport (2004).

Interlocking of Unit Proteins

Our simulations demonstrate that the unstable capsids are those of small viruses with the simplest organization of the coat proteins, where the genetic material of the virus is a key stabilizing factor. The larger and more complex capsids feature intricate networks of interlocking protein tails, and this is obviously correlated with the observed stability of these capsids even in such a crude representation as ours. Examples of the aforementioned interlocking are shown in Figure 6. The tails of the BMV coat proteins intertwine loosely, the main entanglement being at the 3-fold symmetry axis; however, this is enough to hold the capsid in the icosahedral shape. When these tails are cut, the coat proteins loose the main connection factor, and the capsid collapses. The poliovirus and the reovirus exhibit far stronger contacts; the unit protein tails embrace the bulk parts of the other units, and in the case of the reovirus core even exhibit a "thread and needle" arrangement. The importance of the protein tails intertwining is intuitively understandable, but the CG simulations were crucial in demonstrating that the capsids lacking the interlocking tails do collapse, while the ones with interlocking remain stable.

We computed energetic characteristics of contacts between the subunits in each system, both in the all-atom and the CG representations. Due to the crude description, using the CG model one cannot identify hydrogen bonding or salt bridging between the subunits; instead, these types of interactions are implicitly present in the nonbonded interaction potentials. The energies of intersubunit interactions in the CG simulations immediately after minimization can be found in the second to

Table 3. Interactions between the Protein Units in the Simulated Systems

System	AA Energy (kcal/mol per Å^2)	Hydrogen Bonds ($\times 10^{-3}$ per Å^2)	Salt Bridges ($\times 10^{-3}$ per Å^2)	Buried Surface	CG Energy (kcal/mol per Å^2)	Result
Reoviruscore	−1.21	75.3	23.9	35.7	−0.20	Stable
Poliovirus	−0.61	39.1	9.1	8.5	−0.44	Stable
ϕX174 procapsid	−0.44	25.0	11.4	8.8	−0.32	Stable
BMV	−0.36	25.4	8.2	6.0	−0.28	Stable
BMV, no tails	−0.22	19.5	3.4	4.8	−0.12	Collapse
STMV	−0.38	27.3	4.9	5.4	−0.10	Collapse
SPMV	−0.20	12.4	2.2	4.6	−0.03	Collapse
STNV	−0.19	10.2	5.3	4.2	−0.07	Collapse

In each case, the presented values are calculated individually for interactions between each protein chain of one asymmetric unit (or one face of the icosahedral capsid for STMV, SPMV, and STNV) and all other chains within 10 Å of the first chain, and contributions from each chain are summed. Note that "AA Energy" refers to nonbonding interactions between distinct chains in the all-atom system, and that "CG Energy" gives the same quantity in the CG system (calculated for the conformation obtained right after the minimization of the CG system). For the all-atom representation, the number of hydrogen bonds and salt bridges between distinct protein chains is reported. "Buried Surface" refers to the buried surface area of hydrophobic side chains. All values are divided by the area of the asymmetric unit (or face), to account for the different sizes of the systems. The last column states the overall behavior of the system (cf. Table 1).

last column of Table 3. For the all-atom structures, more detailed information can be obtained. The buried hydrophobic surface area, the nonbonded interaction energies, and the number of hydrogen bonds and salt bridges between the adjacent protein subunits are reported in Table 3. To obtain these values, the all-atom crystal structures were minimized in vacuum for 10,000 steps (12,000 steps in the case of reovirus, for which only other chains with atoms within 15 Å of the asymmetric unit were included); energy calculations were performed by using CHARMM22 force fields with a dielectric constant of 20. To compute interaction energies (in both all-atom and CG models), hydrogen bonds, and salt bridges, interactions between each protein chain in an asymmetric unit (or face of the icosahedral capsid for SPMV, STMV, and STNV) with all other chains within 10 Å of the former were taken into account. The reported buried hydrophobic surface area is also for one asymmetric unit (or face). All values were divided by the area of the asymmetric unit (or face), calculated as $4\pi R^2/N_u$, where R is the average radius of the capsid and N_u is the number of asymmetric units (faces) composing the capsid.

The energy of interaction between the protein units in the CG model correlates well with the stability of the capsids. Indeed, the absolute value of interaction energy for all collapsed capsids is at least about two times smaller than the minimal value of that for any of the stable capsids (0.12 kcal mol^{-1} Å^{-2} for BMV with N-terminal tails removed versus 0.2 kcal mol^{-1} Å^{-2} for the reovirus core; note that all interaction energies calculated here were attractive, so larger energies imply a stronger interaction). A similar trend is demonstrated by the values for the all-atom models. The absolute values of interaction energies and the number of hydrogen bonds per unit area are smaller in the case of collapsing capsids than for the stable ones, with the exception of STMV, which demonstrates a hydrogen bond count higher than that for the ϕX174 procapsid and the complete BMV capsid and an absolute value of interaction energy per unit area higher than that for BMV. However, the number of salt bridges per unit area for BMV is almost twice as high as that for STMV, and the other stable capsids have even more salt bridges. Finally, the area of buried hydrophobic surface per viral unit area is generally higher for the stable capsids than for the unstable ones. Interestingly, the ratio of the area of exposed hydrophobic surface to that of the buried one (not shown) is approximately the same for all systems (about 1:4) and is not correlated to the results of our CG simulations. Hence, it seems that in terms of the buried surface, it is not the exposure of hydrophobic surface, but rather the thickness of the protein shell that adds to the stability of the structure.

The data in Table 3 show that when the tails are removed, the absolute value of the energy of interaction between subunits in the CG model is less than half of that for the intact BMV capsid. The all-atom model also suggests a key role for the interaction between the N-terminal tails of BMV subunits in establishing the attractive contacts, as the interaction energy, the number of hydrogen bonds, and the salt bridges are reduced significantly upon removal of the tails.

Our results show that overall the number of attractive contacts between adjacent subunits is higher in stable capsids than in unstable ones. The capsids found to be stable in this study are composed of more protein units than those that collapsed, and the stable capsids are also thicker. Thus, it seems that for the considered viruses, the large number of protein units and the entanglement between them (Figure 6) establish many attractive intersubunit contacts that enhance the stability of the capsids.

Conclusions

We have presented a series of simulations employing a novel, to our knowledge, CG MD method and addressing the question of the dynamic stability of viral capsids. The sizes of the investigated systems were in the range of 10–100 nm, and our simulations revealed the dynamics of these systems on the timescale of 1–10 μs. The CG method suggested here allows one to perform otherwise impossible MD simulations of macromolecular assemblies. Remarkably, we find that the CG simulations demonstrated the same level of performance on parallel computers as all-atom ones do, i.e., the proposed CG method has the same efficiency as the well-developed atomistic MD techniques, but it operates on 100–1000 times larger length and timescales.

We have found that while the BMV and poliovirus capsids, the bacteriophage ϕX174 procapsid, and the reovirus core are stable without their respective genetic material, the empty capsids of all three studied satellite viruses collapse. Whenever a comparison is possible, these and our other more detailed and quantitative results agree with information about the studied viruses known from previous experiments and from the characterization of the STMV particle in all-atom simulations (Freddolino et al., 2006).

Our simulations also provide detailed dynamical information about the structure of the capsids on a microsecond timescale, revealing behavior that previously could only be indirectly obtained from experimental data. For example, the empty capsids of the three studied satellite viruses have not been observed in experiments, which might be due to an intrinsic instability, to the fact that the capsids cannot self-assemble without the help from other molecules, or because of some other, yet unknown reason. Our computations suggest that these capsids are inherently unstable, i.e., even if preformed artificially, they would collapse rapidly. The BMV capsid was found to be stable when composed of complete proteins, but it was found to be unstable if the protein N-terminal tails were removed. This supports the hypothesis about the pathway of in vitro disintegration and rearrangement of the BMV coat proteins proposed by Larson et al., (2005). Finally, we found that the stable capsids exhibit a range of flexibilities; the BMV capsid is the most flexible, the (much more rigid) poliovirus capsid is less flexible, and the bacteriophage ϕX174 procapsid and reovirus core are the least flexible.

The developed shape-representing coarse graining seems to be an attractive option for applications to the stability, assembly, and disassembly of viruses; flexibility and assembly of the cytoskeleton; mechanisms of cellular motility; motions in motor proteins; large-scale mechanics of the nuclear pore complex; and chromatophores of photosynthetic bacteria.

Experimental Procedures

Our CG method is designed to model the large-scale motions of macromolecular assemblies, and it represents the unit molecules that the studied systems consist of by as few point-like particles as possible. The method allows one to reproduce the shape of the unit proteins and to simulate the movements arising due to large-scale interactions between protein domains. A key limitation of the method is that internal conformational change and protein refolding are not allowed.

In the framework of our method, a molecule (capsid protein for the present application) is represented by a number of CG particles ("beads") that are considered to be point masses. The level of coarse graining is a matter of choice. We selected a resolution of ~200 atoms per CG bead, since this furnishes a substantial reduction in the number of particles; however, even the smallest capsid proteins were represented by 5–10 CG beads, a number that allows one to reproduce overall shapes. Biomolecules, and proteins in particular, have a variety of shapes. Often, a single protein is comprised of both compact domains and disordered and elongated linkers or tails, and the compact regions and tails are often equally important. To our knowledge, all existing CG methods assign the CG beads to represent a fixed group of atoms or a fixed number of atoms, but this is not efficient for the coarse graining of molecules with complex shapes, because with such an approach either the tails are misrepresented, or too many CG beads are used for the compact domains. Efficient topology-conserving algorithms have been developed for neural computations (see, for example, Ritter et al., 1992), and we take advantage of one of them, the topology-representing network (Martinetz and Schulten, 1991, 1994; Martinetz et al., 1993).

Consider a molecule whose all-atom structure is known. The molecule consists of N_a atoms with coordinates $\mathbf{r}_n$ and masses m_n; $n = 1, 2, \ldots, N_a$. One seeks to reproduce the shape and inertia of the molecule with the CG beads. The mass distribution, $p_n(\mathbf{r}) = m_n\delta(\mathbf{r} - \mathbf{r}_n)/M$ ($M = \Sigma_n m_n$), is used as a target probability distribution for the evolving map. A given number, N, of CG beads are assigned their initial positions randomly. The CG beads are considered as nodes of the neural network (Martinetz et al., 1993), and S adaptation steps are performed. At each step, the program carries out the following procedures. First, the n^{th} atom of the full structure is chosen randomly, according to an a priori probability distribution, $p_n = m_n/M$; coordinates of the atom, $\mathbf{r}_n = \mathbf{v}$, are used as an input pattern to adapt the neural network, as suggested in Martinetz et al. (1993). Second, for each CG bead, i ($i = 1, 2, \ldots, N$), one determines the number, k_i, of CG beads, j, that obey the condition $|\mathbf{v} - \mathbf{R}_j| < |\mathbf{v} - \mathbf{R}_i|$, where $\mathbf{R}_j$ is the position vector of the j^{th} CG bead. Third, positions of the CG beads are updated ($i = 1, 2, \ldots, N$), according to the rule:

$$\mathbf{R}_i^{new} = \mathbf{R}_i^{old} + \varepsilon \exp[-k/\lambda](\mathbf{v} - \mathbf{R}_i^{old}). \quad (1)$$

Parameters ε and λ are adapted at each step, according to the functional form $f_s = f_0(f_S/f_0)^{s/S}$, where s is the current step and $\lambda_0 = 0.2N$, $\lambda_S = 0.01$, $\varepsilon_0 = 0.3$, and $\varepsilon_S = 0.05$. We use $S = 200N$, which is usually enough to obtain a well-converged distribution of CG beads. A CG model built for a single molecule is transferred to other identical molecules by applying respective center of mass translation and rigid-body rotations to the initial CG model; this ensures uniformity of the CG representation for all protein subunits.

Once the CG beads are placed, an all-atom "domain," the so-called Voronoi cell, is determined for each bead; this domain contains a set of atoms such that every atom from this set is closer to the considered bead than to any other bead. The total mass of each domain is then assigned to the corresponding CG particle. A bond between two CG beads is established if the distance between them is smaller than a cutoff distance (we used 17–20 Å; occasionally, when a protein had long tails, this value was increased to 25 Å). The proteins belonging to separate chains in the all-atom representation were not bonded in the CG representation.

A CHARMM-style force field (MacKerell et al., 1998) is used to represent the interaction potential of CG particles. The values of the bond lengths and angles found in the CG model at the initial coarse graining of the crystal structure are used as the equilibrium bond lengths and angles in the corresponding harmonic potentials. The spring constants are set to 20 kcal/(mol Å^2) and 20 kcal/(mol rad^2) for bonds and angles, respectively. This choice has been motivated by analysis of the behavior of capsid unit proteins in the all-atom simulation of the complete STMV virion (Freddolino et al., 2006). For a unit protein, the domains corresponding to single CG beads were identified, and a relative distance between the domains' centers of mass was analyzed. Assuming that each bond and angle is an independent harmonic oscillator, we extracted the spring constants from the mean distances (angles) and respective rmsds. Obtained values were on the same order of magnitude as the ones we use uniformly in the CG model.

In regard to nonbonded interactions, a CG bead should represent an object approximately the size of the group of atoms (Voronoi cell) that the bead stands for. Therefore, the LJ radius for each CG bead is calculated as the radius of gyration of the respective all-atom domain, extended by 1 Å. The well depth of the LJ potential is always set to 4 kcal/mol. The total charge of the Voronoi cell is assigned to the corresponding CG particle and is used to account for the Coulomb interactions between the CG beads.

The choice of adding 1 Å to the radius of gyration is dictated by the LJ radius of individual atoms in the CHARMM force field, which is typically 1–2 Å, but it was also tested by comparing the stability of the complete STMV particle in the CG representation with that in the all-atom representation (Freddolino et al., 2006). In varying the CG bead LJ radius, we found that the use of 1 Å minimized deviations of the STMV protein subunits relative to those in the all-atom simulation. Although only one parameter from the all-atom study was used to optimize the CG model, we found that other quantities, not involved in the parameterization, are reproduced by the CG

model when compared with the all-atom simulation of the complete STMV (see Results and Discussion).

The dynamics of the CG system is described through the Langevin equation:

$$d\mathbf{p}/dt = \mathbf{F} - \gamma\mathbf{p} + \sigma\mathbf{G}(t), \quad (2)$$

where **p** is the momentum of a bead, **F** is the force acting on the bead due to the chosen interaction potentials, γ is a damping constant, **G**(t) is a univariate Gaussian random process, and σ is connected to γ through the dissipation-fluctuation theorem, $\sigma^2 = 2\gamma k_B T/m$, with m being the bead's mass, k_B the Boltzmann constant, and T the temperature. Use of the Langevin equation allows one to introduce a simple, implicit solvent model, in which the viscous damping and random forces from the solvent are described by the second and third terms, respectively, on the right-hand side of Equation 2. For a single bead without force, Equation 2 describes free diffusion. We performed all-atom MD simulations of the drift of protein segments, similar in size to the ones represented by the CG beads, in explicit water, and we found the effective γ value for the center of mass of a protein segment to be $\sim 20\ ps^{-1}$; this value is used in the CG simulations. The relative dielectric constant is set to 80 everywhere. "Ions" (CG beads carrying a charge of $\pm 20|e|$ and representing 20 Na^+ or Cl^- ions each) were added to neutralize the total charge of the capsids, when necessary.

All simulations presented here used the parallel MD program NAMD 2.5 (Phillips et al., 2005). The CG simulations (Table 1) were run with a single time step of 120 fs, depending on the system. The time step was chosen such that the numerical solution remained stable (not to be confused with the stability of the capsid in each simulation). For the numerical scheme used in NAMD to integrate the equations of motion, the maximal time step for the solution to remain stable is usually determined by the minimal ratio of mass to the spring constant for all bonds in the system (see Phillips et al., 2005 for details). As the minimal masses of CG beads varied from one system to another, slightly different time steps could be used in each case. A 30 Å cutoff was used for the nonbonded interactions. Our CG model requires one to simulate $\sim$1500 times fewer degrees of freedom than in the case of an all-atom representation. A $\sim$200-fold increase in the integration time step allowed us to simulate events on a timescale of tens of microseconds. The largest system studied (simRV) would require a simulation of $\sim$60 million atoms for 1.5 μs, both size and duration currently unfeasible with traditional MD even on thousands of processors; however, in the CG representation, this simulation has been performed on a cluster of 48 processors with a run time of several days. Analysis and visualization were performed with VMD (Humphrey et al., 1996).

Supplemental Data

Supplemental Data include movies of each simulation performed for this study and are available at http://www.structure.org/cgi/content/full/14/12/1767/DC1/.

Acknowledgments

The authors thank Alexander McPherson and John E. Johnson for many suggestions and discussions. This work was supported by a grant from the National Institutes of Health (PHS-5-P41-RR05969). The authors are grateful for the supercomputer time provided by the Pittsburgh Supercomputer Center and the National Center for Supercomputing Applications via National Resources Allocation Committee grant MCA93S028.

Received: June 30, 2006
Revised: October 2, 2006
Accepted: October 5, 2006
Published: December 12, 2006

References

Arsuaga, J., Tan, R.K.Z., Vazquez, M., Sumners, D.W., and Harvey, S.C. (2002). Investigation of viral DNA packaging using molecular mechanics models. Biophys. Chem. *101–102*, 475–484.

Bahar, I., Kaplan, M., and Jernigan, R.L. (1997). Short-range conformational energies, secondary structure propensities, and recognition of correct sequence-structure matches. Proteins *29*, 292–308.

Bamford, D.H., Grimes, J.M., and Stuart, D.I. (2005). What does structure tell us about virus evolution? Curr. Opin. Struct. Biol. *15*, 655–663.

Ban, N., and McPherson, A. (1995). The structure of satellite panicum mosaic virus at 1.9 Å resolution. Nat. Struct. Biol. *2*, 882–890.

Ban, N., Larson, S.B., and McPherson, A. (1995). Structural comparison of the plant satellite viruses. Virology *214*, 571–583.

Bancroft, J.B., Hills, G.J., and Markham, R. (1967). A study of self-assembly process in a small spherical virus. Formation of organized structures from protein subunits in vitro. Virology *31*, 354–379.

Bond, P.J., and Sansom, M.S.P. (2006). Insertion and assembly of membrane proteins via simulation. J. Am. Chem. Soc. *128*, 2697–2704.

Caspar, D.L.D., and Klug, A. (1962). Physical principles in the construction of regular viruses. Cold Spring Harb. Symp. Quant. Biol. *27*, 1–24.

Chu, J.W., and Voth, G.A. (2005). Allostery of actin filaments: molecular dynamics simulations and coarse-grained analysis. Proc. Natl. Acad. Sci. USA *102*, 13111–13116.

Cuillel, M., Jacrot, B., and Zulauf, M. (1981). A T=1 capsid formed by the protein of brome mosaic virus in the presence of trypsin. Virology *110*, 63–72.

Day, J., Ban, N., Patel, S., Larson, S.B., and McPherson, A. (1994). Characterization of crystals of satellite panicum mosaic virus. J. Mol. Biol. *238*, 849–851.

Dodds, J.A. (1998). Satellite tobacco mosaic virus. Annu. Rev. Phytopathol. *36*, 295–310.

Dokland, T., McKenna, R., Ilag, L.L., Bowman, B.R., Incardona, N.L., Fane, B.A., and Rossmann, M.G. (1997). Structure of a viral procapsid with molecular scaffolding. Nature *389*, 308–313.

Dokland, T., Bernal, R.A., Burch, A., Pletnev, S., Fane, B.A., and Rossmann, M.G. (1999). The role of scaffolding proteins in the assembly of the small, single-stranded DNA virus ϕX174. J. Mol. Biol. *288*, 595–608.

Fields, B.N. (1996). Reoviridae. In Fields Virology, B.N. Fields, D.M. Knipe, and P.M. Howley, eds. (Philadelphia: Lippincott-Raven), pp. 1553–1555.

Fields, B.N., Knipe, D.M., and Howley, P.M. (1996). Fields Virology, Third Edition (Philadelphia: Lippincott Williams and Wilkins).

Filman, D.J., Syed, R., Chow, M., Macadam, A.J., Minor, P.D., and Hogle, J.M. (1989). Structural factors that control conformational transitions and stereotypes specificity in type 3 poliovirus. EMBO J. *8*, 1567–1579.

Flint, S.J., Enquist, L.W., Racaniello, V.R., and Skalka, A.M. (2004). Principles of Virology, Second Edition (Washington, D.C.: ASM Press).

Francki, R.I.B. (1985). Plant virus satellites. Annu. Rev. Microbiol. *139*, 151–174.

Freddolino, P.L., Arkhipov, A.S., Larson, S.B., McPherson, A., and Schulten, K. (2006). Molecular dynamics simulations of the complete satellite tobacco mosaic virus. Structure *14*, 437–449.

Hogle, J.M., Chow, M., and Filman, D.J. (1985). Three-dimensional structure of poliovirus at 2.9 Å resolution. Science *229*, 1358–1365.

Hogle, J.M., Chow, M., and Filman, D.J. (1987). The structure of poliovirus. Sci. Am. *255*, 42–49.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. *14*, 33–38.

Ilag, L.L., Olson, N.H., Dokland, T., Music, C.L., Cheng, R.H., Bowen, Z., McKenna, R., Rossmann, M.G., Baker, T.S., and Incardona, N.L. (1995). DNA packaging intermediates of bacteriophage ϕX174. Structure *3*, 353–363.

Jiang, W., Chang, J., Jakana, J., Weigele, P., King, J., and Chiu, W. (2006). Structure of ε15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. Nature *439*, 612–616.

Jones, T.A., and Liljas, L. (1984). Structure of satellite tobacco necrosis virus after crystallographic refinement at 2.5 Å resolution. J. Mol. Biol. *177*, 735–767.

Konecny, R., Trylska, J., Tama, F., Zhang, D., Baker, N.A., Brooks, C.L., and McCammon, J.A. (2006). Electrostatic properties of cowpea chlorotic mottle virus and cucumber mosaic virus capsids. Biopolymers *82*, 106–120.

LaMarque, J.C., Le, T.L., and Harvey, S.C. (2004). Packaging double-helical DNA into viral capsids. Biopolymers *73*, 348–355.

Lander, G.C., Tang, L., Casjens, S.R., Gilcrease, E.B., Prevelige, P., Poliakov, A., Potter, C.S., Carragher, B., and Johnson, J.E. (2006). The structure of an infectious P22 virion shows the signal for headful DNA packaging. Science *312*, 1791–1795.

Larson, S.B., and McPherson, A. (2001). Satellite tobacco mosaic virus RNA: structure and implications for assembly. Curr. Opin. Struct. Biol. *11*, 59–65.

Larson, S.B., Day, J., Greenwood, A., and McPherson, A. (1998). Refined structure of satellite tobacco mosaic virus at 1.8 Å resolution. J. Mol. Biol. *277*, 37–59.

Larson, S.B., Lucas, R.W., and McPherson, A. (2005). Crystallographic structure of the T=1 particle of brome mosaic virus. J. Mol. Biol. *346*, 815–831.

Levine, A.J. (1991). Viruses (New York: W.H. Freeman).

Liljas, L., Unge, T., Jones, T.A., Fridborg, K., Lovgren, S., Skoglund, U., and Strandberg, B. (1982). Structure of satellite tobacco necrosis virus at 3.0 Å resolution. J. Mol. Biol. *159*, 93–108.

Lucas, R.W., Kuznetsov, Y.G., Larson, S.B., and McPherson, A. (2001). Crystallization of brome mosaic virus and T=1 brome mosaic virus particles following a structural transition. Virology *286*, 290–303.

Lucas, R.W., Larson, S.B., and McPherson, A. (2002). The crystallographic structure of brome mosaic virus. J. Mol. Biol. *317*, 95–108.

MacKerell, A.D., Jr., Bashford, D., Bellott, M., Dunbrack, R.L., Jr., Evanseck, J., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B *102*, 3586–3616.

Marrink, S.J., de Vries, A.H., and Mark, A.E. (2004). Coarse grained model for semiquantitative lipid simulations. J. Phys. Chem. B *108*, 750–760.

Martinetz, T., and Schulten, K. (1991). A 'neural gas' network learns topologies. In Artificial Neural Networks, T. Kohonen, K. Mäkisara, O. Simula, and J. Kangas, eds. (Amsterdam: Elsevier), pp. 397–402.

Martinetz, T., and Schulten, K. (1994). Topology representing networks. Neur. Netw. *7*, 507–522.

Martinetz, T.M., Berkovich, S.G., and Schulten, K. (1993). "Neural gas" for vector quantization and its application to time-series prediction. IEEE Trans. Neur. Netw. *4*, 558–569.

McKenna, R., Xia, D., Willingmann, P., Ilag, L.L., Krishnaswamy, S., Rossmann, M.G., Olson, N.H., Baker, T.S., and Incardona, N.L. (1992). Atomic structure of single-stranded DNA bacteriophage ϕX174 and its functional implications. Nature *355*, 137–143.

McKenna, R., Ilag, L.L., and Rossmann, M.G. (1994). Analysis of the single-stranded DNA bacteriophage ϕX174, refined at a resolution of 3.0 Å. J. Mol. Biol. *237*, 517–543.

Miller, S.T., Hogle, J.M., and Filman, D.J. (1996). A genetic algorithm for the ab initio phasing of icosahedral viruses. Acta Crystallogr. D Biol. Crystallogr. *52*, 235–251.

Natarajan, P., Lander, G.C., Shepherd, C.M., Reddy, V.S., Brooks, C.L., and Johnson, J.E. (2005). Exploring icosahedral virus structures with VIPER. Nat. Rev. Microbiol. *3*, 809–817.

Nibert, M.L., Schiff, L.A., and Fields, B.N. (1996). Reoviruses and their replication. In Fields Virology, B.N. Fields, D.M. Knipe, and P.M. Howley, eds. (Philadelphia: Lippincott-Raven), pp. 1557–1596.

Perryman, A.L., Lin, J.H., and McCammon, J.A. (2004). HIV-1 protease molecular dynamics of a wild-type and of the v82f/i84v mutant: possible contributions to drug resistance and a potential new target site for drugs. Protein Sci. *13*, 1108–1123.

Pfeiffer, P., and Hirth, L. (1974). Aggregation states of brome mosaic virus protein. Virology *61*, 160–167.

Phelps, D.K., Speelman, B., and Post, C.B. (2000). Theoretical studies of viral capsid proteins. Curr. Opin. Struct. Biol. *10*, 170–173.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J. Comput. Chem. *26*, 1781–1802.

Rader, A.J., Vlad, D.H., and Bahar, I. (2005). Maturation dynamics of bacteriophage hk97 capsid. Structure *13*, 413–421.

Rapaport, D.C. (2004). Self-assembly of polyhedral shells: a molecular dynamics study. Phys. Rev. E Stat. Nonlin. Soft. Matter Phys. *70*, 051905.

Reddy, V.S., Giesing, H.A., Morton, R.T., Kumar, A., Post, C.B., Brooks, C.L., III, and Johnson, J.E. (1998). Energetics of quasiequivalence: computational analysis of proteinprotein interactions in icosahedral viruses. Biophys. J. *74*, 546–558.

Reinisch, K.M., Nibert, M.L., and Harrison, S.C. (2000). Structure of the reovirus core at 3.6 Å resolution. Nature *404*, 960–967.

Ritter, H., Martinetz, T., and Schulten, K. (1992). Neural Computation and Self-Organizing Maps: An Introduction, Revised English Edition (New York: Addison-Wesley).

Routh, G., Dodds, J.A., Fitzmaurice, L., and Mirkov, T.E. (1995). Characterization of deletion and frameshift mutants of satellite tobacco mosaic virus. Virology *212*, 121–127.

Shelley, J.C., Shelley, M.Y., Reeder, R.C., Bandyopadhyay, S., Moore, P.B., and Klein, M.L. (2001). Simulations of phospholipids using a coarse grain model. J. Phys. Chem. B *105*, 9785–9792.

Shepherd, C.M., Borelli, I.A., Lander, G., Natarajan, P., Siddavanahalli, V., Bajaj, C., Johnson, J.E., Brooks, C.L., III, and Reddy, V.S. (2006). VIPERdb: a relational database for structural virology. Nucleic Acids Res. *34*, D386–D389.

Shih, A.Y., Arkhipov, A., Freddolino, P.L., and Schulten, K. (2006a). Coarse grained proteinlipid model with application to lipoprotein particles. J. Phys. Chem. B *110*, 3674–3684.

Shih, A.Y., Freddolino, P.L., Arkhipov, A., and Schulten, K. (2006b). Assembly of lipoprotein particles revealed by coarse-grained molecular dynamics simulations. J. Struct. Biol., in press. Published online August 24, 2006. 10.1016/j.jsb.2006.08.006.

Speelman, B., Brooks, B.R., and Post, C.B. (2001). Molecular dynamics simulations of human rhinovirus and an antiviral compound. Biophys. J. *80*, 121–129.

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

Tama, F., Miyashita, O., and Brooks, C.L., III. (2004). Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network normal mode analysis. J. Mol. Biol. *337*, 985–999.

Tozzini, V. (2005). Coarse-grained models for proteins. Curr. Opin. Struct. Biol. *15*, 144–150.

Tozzini, V., and McCammon, A. (2005). A coarse grained model for the dynamics of flap opening in HIV-1 protease. Chem. Phys. Lett. *413*, 123–128.

Trylska, J., Tozzini, V., and McCammon, J.A. (2005). Exploring global motions and correlations in the ribosome. Biophys. J. *89*, 1455–1463.

Uchiyama, A., and Fane, B.A. (2005). Identification of an interacting coat-external scaffolding protein domain required for both the initiation of ϕX174 procapsid morphogenesis and the completion of DNA packaging. J. Virol. *79*, 6751–6756.

Valverde, R.A., Heick, J.A., and Dodds, J.A. (1991). Interactions between satelline tobacco mosaic virus, helper tobamoviruses, and their hosts. Phytopathology *81*, 99–104.

Wang, Y., Rader, A.J., Bahar, I., and Jernigan, R.L. (2004). Global ribosome motions revealed with elastic network model. J. Struct. Biol. *147*, 302–314.

Wien, M.W., Curry, S., Filman, D.J., and Hogle, J.M. (1997). Structural studies of poliovirus mutants that overcome receptor defects. Nat. Struct. Biol. *4*, 666–674.

Zhang, D., Konecny, R., Baker, N.A., and McCammon, J.A. (2004). Electrostatic interaction between RNA and protein capsid in cowpea chlorotic mottle virus simulated by a coarsegrain RNA model and a monte carlo approach. Biopolymers *75*, 325–337.