

# Breaking a virus: Identifying molecular level failure modes of a viral capsid by multiscale modeling

V. Krishnamani<sup>1</sup>, C. Globisch<sup>2</sup>, C. Peter<sup>2</sup>, and M. Deserno<sup>3,a</sup>

<sup>1</sup> Department of Molecular Physiology and Biophysics, University of Iowa, Carver College of Medicine, Iowa City, IA 52242, USA

<sup>2</sup> Department of Chemistry, University of Konstanz, Konstanz, Germany

Max-Planck-Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

<sup>3</sup> Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA

**Abstract.** We use coarse-grained (CG) simulations to study the deformation of empty Cowpea Chlorotic Mottle Virus (CCMV) capsids under uniaxial compression, from the initial elastic response up to capsid breakage. Our CG model is based on the MARTINI force field and has been amended by a stabilizing elastic network, acting only *within* individual proteins, that was tuned to capture the fluctuation spectrum of capsid protein dimers, obtained from all atom simulations. We have previously shown that this model predicts force-compression curves that match AFM indentation experiments on empty CCMV capsids. Here we investigate details of the actual breaking events when the CCMV capsid finally fails. We present a symmetry classification of all relevant protein contacts and show that they differ significantly in terms of stability. Specifically, we show that interfaces which break readily are precisely those which are believed to form last during assembly, even though some of them might share the same contacts as other non-breaking interfaces. In particular, the interfaces that form pentamers of dimers never break, while the virtually identical interfaces within hexamers of dimers readily do. Since these units differ in the large-scale geometry and, most noticeably, the cone-angle at the center of the 5- or 6-fold vertex, we propose that the hexameric unit fails because it is pre-stressed. This not only suggests that hexamers of dimers form less frequently during the early stages of assembly; it also offers a natural explanation for the well-known  $\beta$ -barrel motif at the hexameric center as a post-aggregation stabilization mechanism. Finally, we identify those amino acid contacts within all key protein interfaces that are most persistent during compressive deformation of the capsid, thereby providing potential targets for mutation studies aiming to elucidate the key contacts upon which overall stability rests.

---

<sup>a</sup> e-mail: [deserno@andrew.cmu.edu](mailto:deserno@andrew.cmu.edu)

## 1 Introduction

Viruses infect all domains of life: archaea, bacteria, and eukaryota. They are ubiquitous, extremely resilient to harsh environments, and protect their genetic information (RNA or DNA) within proteinaceous covers, which very frequently take the form of capsids—nanocontainers that enclose the viral genome. Such “viral capsids” are typically composed of smaller protein subunits, which spontaneously assemble in solution. As Watson and Crick pointed out, the genome of the virus must also code for its protective capsid. From this obvious fact alone they conclude that the resulting severe limitations in storable information compel viruses to encode not for a full capsid but instead for many copies of only a small number of different proteins, which then aggregate into highly stable isometric macromolecular structures—which is precisely what is observed [1].

The high symmetry and monodispersity of viral capsids implies that they crystallize exceptionally well, and so we are in the fortunate position to have atomic resolution X-ray and Cryo-EM structures for many of them [2]. These structures reveal that a very large number of spherical viruses share the same symmetry: that of an icosahedron (there are currently more than 440 redundant structures of viral capsids in the Protein Data Bank that are icosahedral [4]). Based on this fact (but with *much* fewer known examples to work with) Casper and Klug [3] have developed a classification scheme for icosahedral viral morphologies, which ends up systematically enumerating the different possibilities by which equilateral triangles can assemble into closed shells of icosahedral symmetry. The key parameter is the so-called  $T$ -number, which has the following meaning: the total number of triangles is  $20T$ ; moreover, each triangle is always composed of three protein subunits, and the  $T$ -number then also gives the total number of *symmetrically inequivalent locations* for these subunits. Many well-known viruses can be labeled by this simple classification scheme, for instance Rhinovirus ( $T = 3$ ), Hepatitis B ( $T = 4$ ), T7 bacteriophage ( $T = 7$ , the coincidence between name and  $T$ -number being just that: coincidental!) and Cowpea Chlorotic Mottle Virus (CCMV,  $T = 3$ ), to name but a few. Among all the known structures of spherical capsids, more than 70% of the capsids are *homomers*—meaning, they are constructed out of  $60T$  copies of a *single* protein subunit as their building block. But recall that these must form a structure which features  $T$  symmetrically inequivalent sites. Hence, these proteins of identical primary sequence generally exhibit slightly different secondary or tertiary structures at these different sites.

One of the fascinating aspects of viral capsid assembly is that the non-trivial global geometry and stability of the full structure is completely determined by the local chemical properties of the amino acids on the capsid protein’s exposed/buried surfaces and its elastic modes. In other words, the properties of an assembled viral capsid are fully encoded in the chemistry and geometry of a single building block. Morphology and elasticity of a viral capsid are hence beautiful examples of emergent properties which arise during a hierarchical self-assembly process. The underlying nucleation and growth mechanism of viral capsid was first proposed by Prevelige and coworkers [5]: empty capsids assemble by nucleation and subsequent growth: a critical nucleus forms, followed by sequential addition of subunits until the capsid is completed.

Unfortunately, studying this hierarchical assembly is extremely difficult. On the experimental side we run into the problem that capsid proteins are evolutionarily optimized to form capsids, and they do this so well (and so quickly) that it is very hard to capture stable intermediates. In addition, site-directed mutagenesis—one of the cornerstones of structural and functional studies in molecular biology—is a cumbersome way to study the hierarchical construction of large complexes. On the computational side, the central difficulty is to bridge the necessary length- and time scales involved

in this self-assembly process: even small capsids consist of a very substantial number of atoms, and capturing the very slow modes which large aggregates exhibit requires enormous computational resources. Fortunately, the hierarchical nature of the problem suggests an elegant solution to this challenge: multiscale modeling. This is the path we will follow in our work.

Gaining a better insight into this assembly process is not merely a matter of intellectual curiosity (even though it is that, too). A better understanding could for instance identify and predict those interactions between building blocks that are most crucial for assembly, and hence maybe more difficult for the virus to mutate in response to human drugs. Such knowledge could therefore be used to combat viruses by interfering with their capsid assembly process. Efforts along these lines have indeed been made to combat Hepatitis B viral infections [6–8]. Others have attempted to control the assembly of viral capsids around functionalized nanoparticles [9–13]. But even beyond biology, elucidating the nature of interactions that assemble viral capsids has practical implications in other fields, most notably nanotechnology. For instance, capsid proteins can be engineered to build “containers” for encapsulating chemicals and materials for industrial applications [14,15], and understanding the molecular cues that direct viral capsid assembly lies at the heart of “programmable materials” [16–18]. Specifically, the capsid of CCMV is being evaluated as a biotemplate for polymer encapsulation [19,20], 2D/3D array formation [21,22], enzyme nano-reactors [23], chemical conjugation [24], pathogen targeting [25,26] and drug delivery [27,28]. An in-depth understanding of the properties of viral capsids is central to developing such cutting edge applications.

Against this background, we here present a detailed residue-level description of the mechanical stability of a particular viral capsid, that of CCMV. This analysis is based on our previously developed multi-scale model of a CCMV capsid [29]. Specifically, we propose that key aspects of the assembly process—which is very challenging to study directly—can instead be inferred from the “inverse” process: the destruction of a capsid under an applied force. Accordingly, we mechanically compress the CCMV capsid by subjecting it to an external force, while monitoring the location, sequence of breakage events, and any change of interactions that “glue” the proteins together. This is of course very similar to experiments that “crush” capsid with, say, the tip of an atomic force microscope (AFM)—indeed, Michel et al. have done exactly this for CCMV [30]. The advantage of replicating this protocol in simulations is that we can monitor exactly where the capsid breaks at a much better resolution than what is experimentally possible.

A central result of our investigations is that we lend independent support to the proposed assembly order of a CCMV viral capsid [40,41]: protein dimers form the first stable substructures, which then associate to form a nucleation assembly, a pentamer of dimers (POD). Many such PODs are subsequently “glued” together (during the growth phase) by additional dimers. The two proteins forming the dimers that assemble into PODs will end up occupying two of the three symmetrically inequivalent sites (we label them A and B, and then speak of the “AB-dimer”), while the proteins that form the “glue-dimer” both belong to the third symmetry site (hence we label them C and speak of the “CC-dimer”). In addition, we propose a functional role for a well-known stabilization motif—a  $\beta$ -barrel—at the quasi-six-fold rotational symmetry site of CCMV capsid [42,43]. This structure, especially its *location*, has been very puzzling, because it is found at the hexameric site, not the pentameric one, and yet the pentamers are believed to be more stable. Is this barrel *destabilizing*? We instead argue that it is located at the weaker spot precisely *because* it is the weaker spot: it is a subtly cooperative structural motif that helps stabilize a region of the capsid which out of necessity cannot feature strong interactions.

## 2 Model system

### 2.1 The virus: CCMV

For the investigation presented in this paper, we built our model system based on the viral capsid structure of Cowpea Chlorotic Mottle Virus (CCMV) [44], which belongs to the Bromoviridae family. Its icosahedral ( $T = 3$ ) capsid spontaneously self-assembles in solution. As the  $T$ -number implies, the capsid consists of 180 chemically identical proteins, each of which comprises 190 amino acid residues. Structurally, each individual protein is folded around a core made out of 8  $\beta$ -strands. The N-terminus is rich in basic residues and extends into the interior of the capsid, where it helps to bind the genetic material (RNA in the case of CCMV). The C-terminal tail strongly binds to a second capsid protein and is hence key in forming very stable protein dimers. CCMV's  $T$ -number of 3 further implies that individual proteins distribute over three symmetrically inequivalent sites (A, B, and C), at which they adopt three marginally different structural conformations; as a consequence, dimers in the fully assembled capsid occur in two different types: AB and CC.

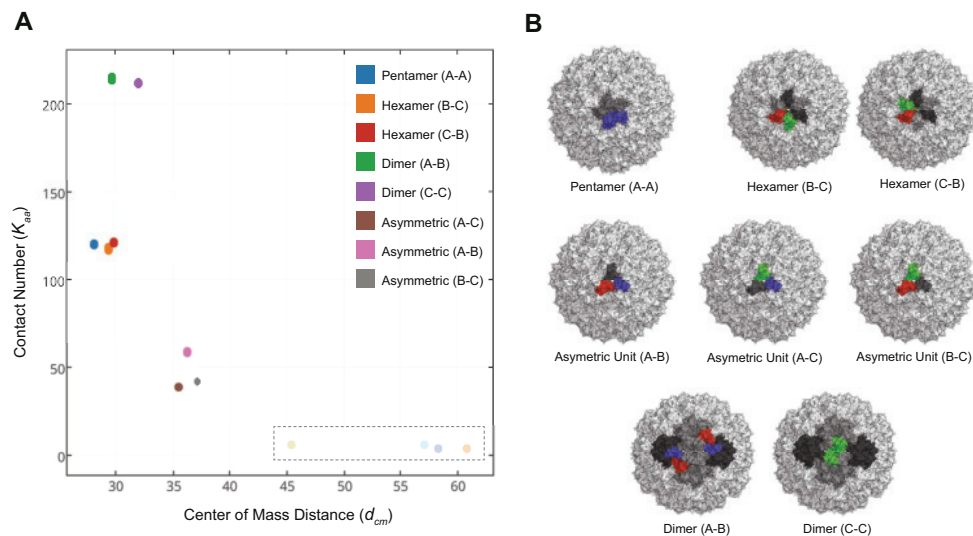
Further, two higher-order associations observed in the CCMV capsid are of significance: first, an assembly of five AB dimers, which we call the pentamer of dimers (POD); second, a hexamer of dimers (HOD), which in the fully assembled capsid contains three AB and three CC dimers, but which is not believed to be a distinct unit that is added to the capsid during assembly.

The AB and CC dimers naturally result from the assembly of individual monomers, but in a fully assembled capsid there are several other ways in which two individual proteins can end up touching, forming other types of interfaces that hold the capsid together. We can easily classify them based on the number of contacts between two protein units in the final crystal structure, and such an analysis leads to 12 distinct pairings (which includes the AB and CC dimers). Of these, only 8 are “strongly bound” (in terms of number of contacts at its interface), and so we restrict our subsequent discussion to those. Figure 1 shows the number of contacts which these interfaces have as a function of the center-of-mass distance between the subunits, illustrating the 8 pairings, which in turn can be subdivided into 3 groups that differ in the extent to which they touch. Figure 1B illustrates each interface embedded into the context of a full virus.

The self-assembly process of CCMV capsid is postulated to be nucleated by PODs which are subsequently pieced together (growth phase) to form the whole capsid by the addition of CC dimers [40]. The quasi-six-fold symmetry site (the hexamer of dimers, HOD) hence emerges during assembly but is not a unique entity being added to a capsid during assembly. While this assembly pathway is highly plausible, there is unfortunately only limited direct experimental or theoretical evidence for it, largely because the timescale of CCMV capsid self-assembly sits squarely in the “no-mans-land” of investigability: too fast for experiment, but too slow for simulation.

CCMV has an interesting stabilizing motif at the quasi-six-fold symmetry site, in the center of the HOD: the six N-terminal tails of the proteins which meet there form a short  $\beta$ -barrel. In contrast, no such barrel exists at the POD site, where the N-terminal tails of five proteins come together. Bereau et al. recently demonstrated in a computational study that this difference results from the subtly different local orientation and alignment of the five or six proteins [42]. Even if these authors artificially enhance the propensity of the force-field to form  $\beta$ -strands, such a structure never forms at the POD site, showing that its local geometry is not at all conducive to a barrel. Combined with our present studies of capsid stability under compression, we suggest a role for this barrel in the assembly process.

In-vitro experiments have shown that a deletion mutation of the N-terminus (residues 1–36, which also includes the barrel region) prevents the packaging of RNA



**Fig. 1. Classification of interaction categories:** (A) Pairwise correlation plot of the center of mass ( $d_{cm}$ ) distance versus the contact number ( $K_{aa}$ ) of protein pairs in the crystal structure of CCMV capsid. (B) Visualization of 8 categories of pairwise monomers with contact number higher than 10. The categories (9 to 12) are ignored (grayed out) in our analysis since an insignificant number of contacts are made by these protein pairs. Monomers are highlighted according to their types *A*, *B*, *C* in blue, red and green. The pentameric association site, hexameric association site, and asymmetric unit are highlighted in gray where appropriate.

but does *not* disrupt the formation of empty capsids [45,46]. Here we use a  $\Delta 1-36$  mutant of the Salt-Stable CCMV (Lysin 42 to Arginine) capsid protein in all our simulations. (Notice that throughout the paper we will use the residue numbering of the  $\Delta 1-36$  mutant, this means that residues 37–190 of the full-length protein will here be referred to as residues 1–154). The salt-stable mutant was also used in the AFM indentation studies of Michel et al. [30], whose force-indentation curves we previously used to validate our capsid model [29]. We note that the current understanding of viral capsid self-assembly, in particular the experimental and theoretical methods that have been used to study this process and their material properties, is detailed in recent reviews [47,48].

## 2.2 Computational model

The basis for our computational study is the MARTINI model, a coarse-grained (CG) forcefield for proteins [49,50], stabilized with one of two different types of elastic networks, ELNEDIN [51] and IDEN [29]. The MARTINI model provides a fairly detailed side chain resolution (up to four beads, for tryptophan) that can reasonably represent protein-protein contacts and packing. However, its low resolution on the backbone (one bead per amino acid) implies that it does not capture secondary structure formation. Consequently, it requires an additional “reinforcement” to preserve native secondary and tertiary structure contacts of a protein. The developers of the MARTINI forcefield have proposed a simple and straightforward procedure to construct an elastic network, termed “ELNEDIN” [51], where all residues within some cutoff distance (usually 9 Å) are connected by harmonic bonds. The equilibrium

lengths of these bonds are extracted from the crystal structure, while the spring constant  $K$  is uniformly fixed to be usually  $500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ .

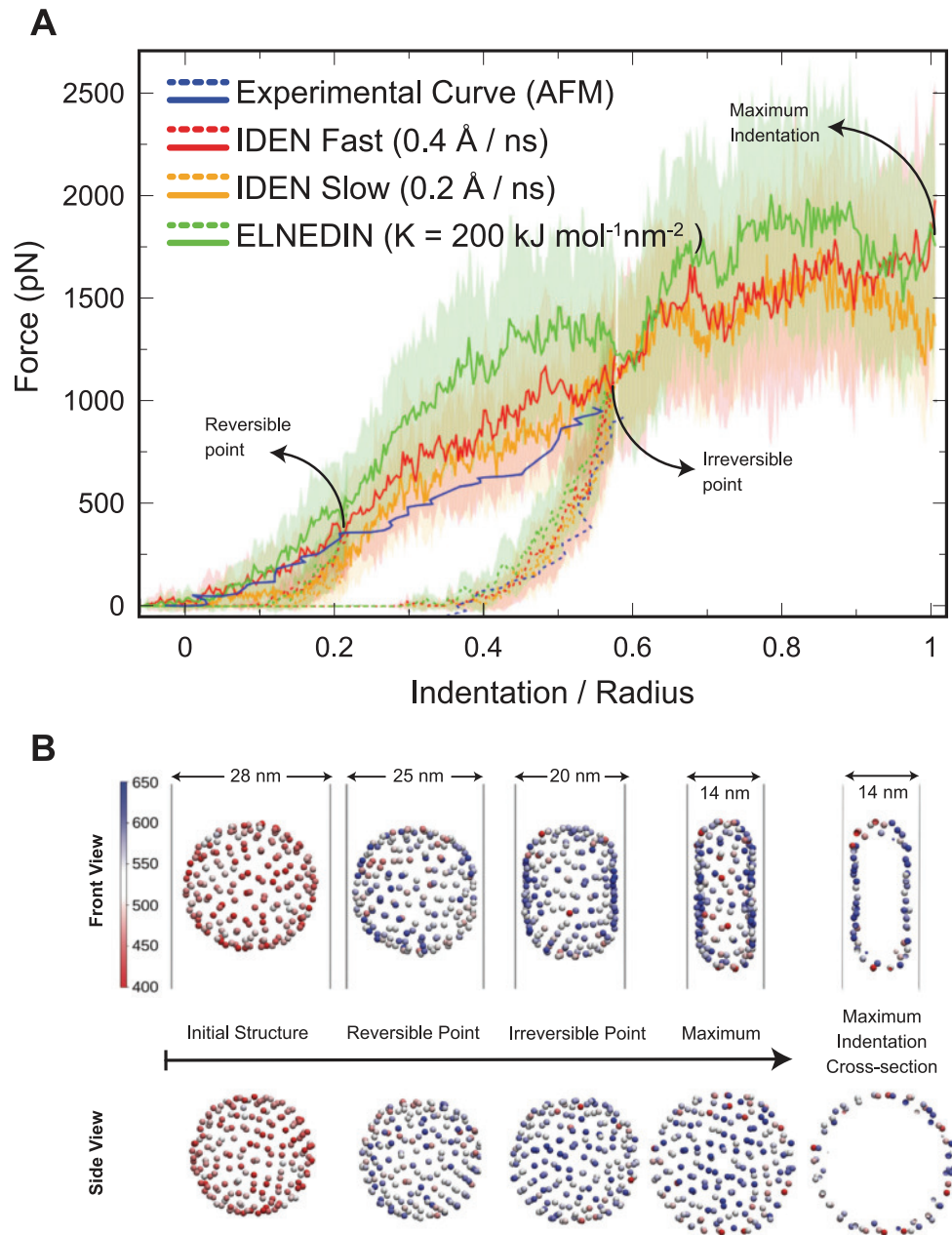
In our previous work on CCMV [29], we found that the MARTINI representation of CCMV dimers in which each protein is stabilized by an ELNEDIN network with a spring constant of  $500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  is too stiff, and a smaller spring constant of  $K = 200 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  more accurately replicates the large scale elastic properties of an equivalent all-atom dimer simulation. However, even with this “softer” network the local fluctuations were still under-represented. We showed that such differences at the dimer-level add up in a larger complex of dimers (such as a POD). When the CCMV capsid constructed out of 90 such dimers (ELNEDIN with  $K = 200 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ ) is compressed, our computed force-indentation relation [29] is noticeably too stiff (Fig. 2A) compared to the one measured in AFM compression experiments of CCMV viral capsids [30]. We hence proposed an iterative procedure to optimize the elastic network: tune the strengths of *individual* spring constants of the elastic network until they reproduce the fluctuation modes obtained from a reference all-atom simulation of a single dimer in solution [29]; we called this elastic network “IDEN”. We then demonstrated that MARTINI combined with the IDEN network performed better, resulting in quantitative agreement with the AFM experiments. Remarkably, the MARTINI+IDEN model even reproduced the *irreversible* regime of capsid breakage and the associated hysteresis of the force-indentation relationship [29].

In this paper, we will use both ELNEDIN (with the softer  $K = 200 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  spring constant) and IDEN elastic networks. We emphasize that the sole purpose of the elastic networks in our simulations is to maintain the structural integrity of a single capsid protein. All *inter*-protein interactions are governed by MARTINI forcefield parameters. Comparing ELNEDIN and IDEN elastic networks will permit us to assess whether our conclusions depend on these modeling details. Pleasingly, we will see: they don't.

## 3 Methods

### 3.1 Classifying interacting capsid protein pairs

As described in the previous section, we identified 12 types of interfaces where two proteins “make contact” in the capsid crystal structure (PDB ID: 1ZA7) [44]. A “contact” (in this case) is established if the distance separating the  $C_\alpha$  atoms of any two residues between any given protein pair is within  $10 \text{ \AA}$  of each other. For all these “interacting” pairs two observables were calculated: the Euclidean distance  $d_{\text{cm}}$  between the centers-of-mass of the proteins; and the number  $K_{\text{aa}}$  of amino-acid-contacts. Figure 1A shows  $K_{\text{aa}}$  versus  $d_{\text{cm}}$  for all interacting protein pairs in the capsid crystal structure. We found that the cut-off distance criterion of  $10 \text{ \AA}$  used here was well-suited to separate the various interface categories for easy classification. In the following we will focus on the 8 interfaces that exhibit a non-negligible number of contacts. These 8 interfaces can be further subdivided into three groups: the interfaces within the two dimers (A-B and C-C) are characterized by more than 200 amino-acid contacts. The interfaces between two monomers in the protein rings that form the centers of a POD (denoted as pentamer (A-A)) or a HOD (denoted as hexamer (B-C) and hexamer (C-B)) are made up of approximately 120 amino acid contacts. And the three interfaces within the asymmetric unit are characterized by 40 to 60 amino-acid contacts. Figure 1B illustrates these interface types and how they are situated in the capsid as well as the nomenclature used from now on. The nature of these interfaces will be characterized further below, and their fate upon compression of the capsid will be the central investigation in this paper.



**Fig. 2. Force vs. Indentation curves for compressive deformation of CCMV capsid:** (A) Comparison of experimental values reported for empty *SubE* mutants of CCMV [30] (blue) and computed values from CG simulations; 1) ELNEDIN elastic network with a uniform spring constant of  $K = 200 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  (green), 2) An IDEN elastic network with a fast compression speed ( $0.4 \text{ \AA ns}^{-1}$ ) (red) and 3) IDEN network with reduced compression speed ( $0.2 \text{ \AA ns}^{-1}$ ) (orange). Forces were calculated for forward indentations (solid lines) and the subsequent relaxation process (dashed lines). Values on the  $x$ -Axis are reported as indentation relative to the radius of the uncompressed capsid, i.e. the relevant stages of indentation are: uncompressed initial structure (wall distance of 28 nm,  $I = 0 \text{ nm}$ ,  $R = 14 \text{ nm}$ ),

### 3.2 Computational details

All simulations were performed using the GROMACS simulation package (versions 4.0.7 or 4.5.4) and the MARTINI force field [49, 50, 52, 53]. In two independent simulation setups, we used both the conventional ELNEDIN ( $K = 200 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ ) and our iteratively optimized IDEN elastic network to stabilize the conformation of each capsid protein. The construction and testing of the IDEN elastic network is described in detail by Globisch et al. [29]. As mentioned above, the elastic network “bonds” were only established between  $C_\alpha$  atoms within each monomer, while the protein-protein interactions was mediated by the MARTINI forcefield parameters.

The temperature (300 K) and pressure (1 bar) were maintained in the simulation box using the Berendsen method with coupling constants  $t_T = 0.5 \text{ ps}$  and  $t_P = 1.2 \text{ ps}$ , respectively. The non-bonded interactions were treated with a switch function—electrostatic interactions from 0.9 to 1.2 nm and Lennard-Jones interactions from 0.9 to 1.2 nm.

The time step was set to 10 fs and the neighbor list was updated every 5<sup>th</sup> time step. The box size exceeded the protein outer surface by a minimum of 2.25 nm in all directions. Note at this point that in any CG simulation an effective speed-up of the dynamics is expected due to the smoothening of the free energy landscape. Therefore, one typically determines a time-scaling factor that relates the accelerated CG dynamics to realistic timescales obtained from experiments or atomistic reference simulations. According to the authors of the MARTINI forcefield and the ELNEDIN elastic network, this time-scaling factor is approximately 4 (by analyzing the diffusion of CG water and lipids) [49, 54].

### 3.3 Compressing the virus

As described in Ref. [29] the empty CCMV capsid (without RNA; with the  $\Delta 1-36$  mutant) was mechanically compressed with the help of two semi-permeable repulsive walls (using a 10–4 Lennard Jones potential), that only interacted with the MARTINI protein beads but not with solvent particles or the ions. This mechanical compression was chosen to replicate experiments performed by Michel et al. [30]. The four-sided pyramidal AFM tip used in the experimental study can be approximated by a sphere with a 40 nm diameter, whereas the CCMV capsid itself has a diameter of 28 nm. An AFM tip of comparable or even larger size than the capsid rather probes global elastic properties instead of local ones. In the spirit of this observation, in the simulations we go to the limit of plane walls, thus removing one length scale, namely the tip size.

Initially we placed the walls just beyond either ends of the capsid with a separation of 29 nm, not yet touching the virus, which has a diameter of 28 nm. Subsequently, the separation between the walls was decreased stepwise in 0.04 nm

---

**Fig. 2.** (Continued). The beginning of reversible relaxation (wall distance of 25 nm,  $\frac{I}{R} = 0.22$ ,  $I = 3.08 \text{ nm}$ ), the beginning of irreversible relaxation (wall distance of 20 nm,  $\frac{I}{R} = 0.58$ ,  $I = 8.12 \text{ nm}$ ) and the final structure after breakdown (wall distance of 14 nm,  $\frac{I}{R} = 1.0$ ,  $I = 14 \text{ nm}$ ). Note that the simulation begins with a wall distance of 29 nm, leaving a 0.5 nm space between each side of the virus capsid and the wall, corresponding to the negative region on the  $x$ -axis. (B) Illustration of the compressive deformation of the capsid in the slow compression IDEN simulation. Individual capsid proteins are shown as spheres in this representation colored according to their contacts to the surrounding proteins (using a distance criterion of 0.6 nm). The numbers of these contacts vary between 400 or below (red; relaxed virus or ruptured contact) and 650 (blue; heavily compressed regions).



(fast compression) or 0.02 nm (slow compression) increments with a relaxation time of 1 ns per step. Compression simulations were carried out for the ELNEDIN network with  $K = 200 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  at the fast compression speed and for the optimized IDEN elastic network at the fast and the slow compression speed. The capsid orientation was not kept fixed during compression; instead, the capsid reoriented before in the first stages of the compression until it locked into a specific orientation at which it subsequently remained. We note that in the simulations with different compression speeds we found different capsid orientations.

Three types of compression were distinguished, with compression limits chosen based on the limits reported by Michel et al. in their AFM experiments [30]:

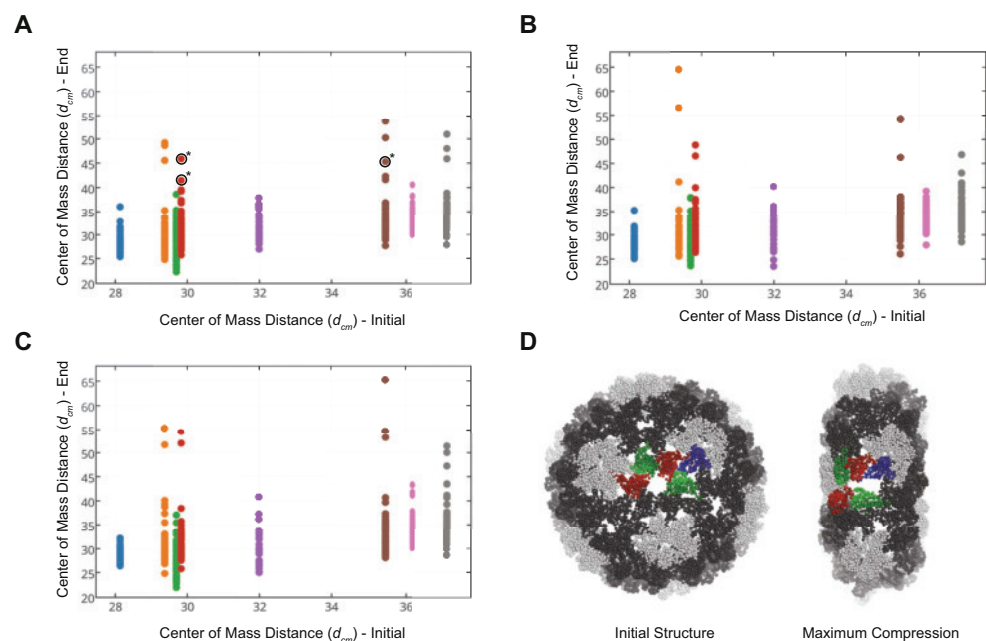
1. Reversible compression: compression was carried out to a (minimal) wall-to-wall separation of 25 nm. Then, the capsid can expand again elastically to its initial structure when the pressure by the walls is relaxed (by following the inverse procedure and receding the walls incrementally).
2. Irreversible compression with plastic deformation: upon further compression (down to a wall-to-wall separation of 20 nm) the capsid undergoes an irreversible plastic deformation and no longer recovers upon receding the walls.
3. Irreversible compression until failure: when compressed further to a wall-to-wall distance of 14 nm, the capsid visibly breaks at specific protein-protein interfaces.

Figure 2A shows the resulting force-indentation relation from simulation as well as the experimentally obtained data [30]. As discussed previously [29], all elastic networks provide a fairly reasonable description of the experimentally measured force-indentation relation. Compared to the earlier study, the compression here is carried out up to mechanical failure of the virus. Figure 2B shows depictions of characteristic capsid structures at the different indentation stages. Each monomeric protein is indicated by one spherical particle, and the coloring was made according to the number of contacts. Figure 2B shows that in general, compressing the capsid increases the number of contacts between the capsid proteins. Only at extreme compression of the capsid we observed a few critical breakage points in the capsid structure (indicated by red beads) that were predominantly found at the rim of the capsid and not at the flat regions that are in contact with the compressive walls. The view on the capsid along the compression axis (Fig. 2B, lowest panel) shows that during compression, the capsid circumference increases, leading to a positive tensile stress ( $\Sigma_+$ ) around its circumference. Since the number of contacts slightly increase, this must result from more contacts being formed along the direction of compression, along which the negative compressive stress ( $\Sigma_-$ ) acts. By a simple rotation of axes, one can see that this biaxial stress results in a shear stress  $(\Sigma_+ - \Sigma_-)/2$  along in-plane directions rotated by  $45^\circ$  with respect to the two principal directions. We observe the consequence of these stresses in rupture of the CCMV viral capsid at its weakest interactions along its circumference. In the following we will in more detail investigate, which protein/protein interactions are broken when the capsid ruptures.

## 4 Results

### 4.1 Identifying breaking interfaces

To identify the protein pairs/interfaces in a CCMV viral capsid that break or rupture upon compression, we plot the distance  $d_{\text{cm, end}}$  at the end of the simulation (i.e. at maximum indentation) versus the initial crystallographic distance  $d_{\text{cm, initial}}$  (Fig. 3A, Fig. 3B and Fig. 3C). Here  $d_{\text{cm}}$  is the distance between the *center of mass*



**Fig. 3. Center of mass ( $d_{\text{cm}}$ ) distance plots for the ELNEDIN and IDEN simulations.** The plot shows  $d_{\text{cm}}$  distance in the x-ray structure versus the  $d_{\text{cm}}$  at maximum compression of the CCMV capsid. The coloring pattern of each interface category is as established in Fig. 1. (A) IDEN elastic network with slower rate of compression ( $0.2 \text{ \AA ns}^{-1}$ ), (B) IDEN at  $0.4 \text{ \AA ns}^{-1}$  compression rate and (C) ELNEDIN elastic network with a compression rate of  $0.4 \text{ \AA ns}^{-1}$ . (D) Illustration of a locus of breakage in CCMV capsid at maximum compression with IDEN elastic network that was compressed slowly ( $0.2 \text{ \AA ns}^{-1}$ ). The colored monomers (A – blue, B – red, C – green) are involved in breakage and the corresponding protein pairs are marked with an *asterix* in A.

of each protein. Since the capsid loses its icosahedral symmetry as a result of compression,  $d_{\text{cm, end}}$  is not a unique number but depends on the location of a specific protein-protein interface and on how the compressive stress rearranges the two proteins with respect to each other. This results in a wide range of  $d_{\text{cm, end}}$  values at a given  $d_{\text{cm, initial}}$ . This representation is well suited to identify which interfaces consistently stay in contact and which interfaces don't. Generally, one finds that out of the 8 interface classes introduced above, only a few classes exhibit vast outliers of  $d_{\text{cm, end}}$ , measured at the end of the compressive deformation of the capsid. Visual inspection confirms that these are truly broken interfaces.

We consistently find that the rupture events involve either hexamer (B-C) or (C-B) interfaces or asymmetric (A-C) or (B-C) interfaces. This is observed in all simulations irrespective of the indentation speed or the details of the elastic network (4 independent simulations using IDEN and ELNEDIN networks. Data from ELNEDIN with  $K = 500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  are not shown). These interfaces also have the least number of contacts at maximum compression (see Table 1). The snapshots in Fig. 3D provides an illustration of one specific breakage event that involves the tearing apart of three interfaces—two hexamer (C-B) interfaces and one asymmetric (A-C) interface. Note that neither dimer (A-B) or (C-C) interfaces nor pentamer (A-A) interfaces ever break. Especially the latter observation is quite remarkable since the pentamer (A-A) interface falls into a similar category as the hexamer (B-C) and (C-B) interfaces both in terms of  $d_{\text{cm, initial}}$  and in terms of the number of amino-acid contacts

**Table 1.** List of residues that interact (contact clusters) with a neighboring protein are reported here from an average contact map of the last 10 ns of squeezing of CCMV viral capsid (see Fig. 3, maximum compression). A contact is defined if the backbone MARTINI bead (equivalent of a  $C_{\alpha}$  atom) of a residue in monomer 1 is within 6.0 Å of another residue's backbone bead in monomer 2 (see Fig. 4). The residue numbers correspond to the N-terminal truncated monomer as described in Sect. 2.1. The *electrostatic*, *hydrophobic* and *hydrophilic* contacts that are reported in this table were determined using the following convention: *Negative* amino acids are ASP and GLU; *Positive* amino acids are ARG, LYS and HIS; *Polar* amino acids are SER, THR, ASN, GLN, TYR, CYS and GLY; *Non-polar* amino acids are ALA, ILE, LEU, PHE, TRP, VAL, MET, PRO. An *electrostatic* interaction is identified if a negatively charged residue is within 6.0 Å of a positively charged amino acids or *vice versa*. Hydrophobic and hydrophilic interactions were established if two amino acids that belong to the same *group* are within 6.0 Å of each other. The numbers reported in the brackets are number of contacts in each category obtained from an average contact map of the first 10 ns of the squeezing simulation.

	Name	Monomer 1	Monomer 2	Hydrophobic	Hydrophilic	Electrostatic	Total
Pentamer	A-A	17, 19-25, 56-62, 89-97, 135-137, 146	29-31, 67-72, 79-97, 104, 122-125	55 (62)	22 (25)	3 (2)	203 (233)
		Dimer	A-B	1-17, 56-59, 97, 99, 111, 113-116, 135-153	1-20, 39-42, 44, 56, 58-60, 95-99, 138-154	63 (77)	35 (42)
C-C	1-17, 19, 54-60, 96-99, 114, 136-154			3-19, 41, 56, 58-60, 95-99, 101, 138-154	58 (74)	24 (42)	10 (11)
Hexamer	B-C	19-25, 57-61, 90-97, 135, 137	29, 68-70, 79-90, 92, 95, 123-125	34 (62)	13 (18)	2 (2)	133 (207)
	C-B	18-23, 58-66, 98, 130-136	30-32, 63, 65-77, 85-87, 124,126	31 (44)	21 (33)	2 (2)	152 (207)
Asymmetric Unit	A-C	51, 73-83, 85-90, 101-110, 113-114, 117-118	1-6, 9-10, 43-45, 104-117	30 (31)	23 (16)	3 (4)	172 (168)
	A-B	1-10, 104-117	51, 56, 73-92, 95-97, 99-110, 113-114, 117-119	43 (40)	20 (15)	4 (3)	222 (200)
	B-C	1-11, 44-45, 104-116	51, 73-80, 85-90, 100-110, 113-114, 117-118	29 (33)	21 (14)	2 (3)	177 (170)

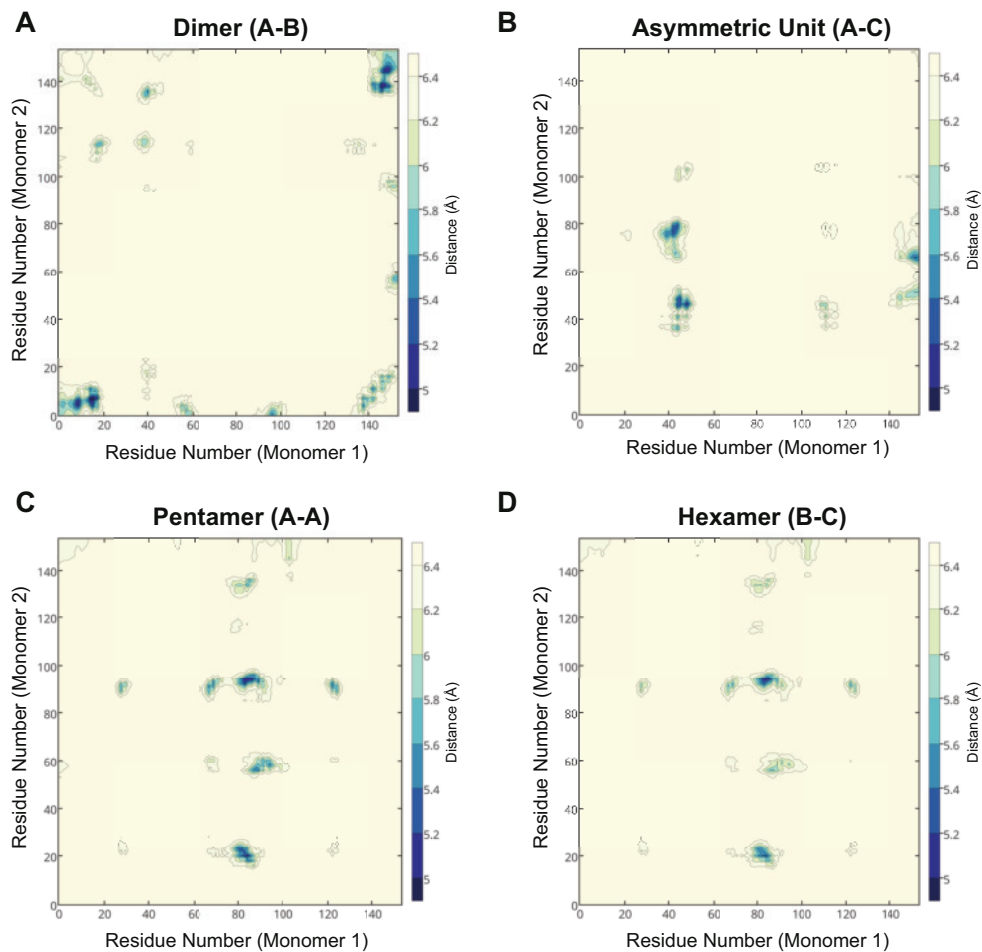
in the crystal structure. This aspect will be further investigated below. Note that the snapshot also illustrates the observation that the rupture events predominantly occur at the rim of the capsid (see also Fig. 2B). Given that upon compression the capsid locks into a particular orientation, one might worry that this would end up being one in which the high stress rim of the capsid is devoid of pentamers, and that this would explain why the pentamer interfaces do not break. This is not true, though. We have counted the number of pentameric and hexameric symmetry sites near the rim, and they occur essentially at the same ratio that also holds for the whole capsid, namely 3:5. In fact, pentameric symmetry sites are too densely spaced on the surface of the capsid to find any orientation that appreciably “depletes” the rim of pentamer interfaces.

## 4.2 Characterizing contact interfaces of CCMV capsid at maximum compression

For different types of protein-protein interfaces (as introduced in Sect. 3.1 and Fig. 1) we calculated two-dimensional contact maps between the beads that replaces the backbone atoms (equivalent to  $C_\alpha$  atoms) in the MARTINI forcefield (Fig. 4). These contact maps were averaged over the last 10 ns (just before maximum compression, see Fig. 2) of the compression over all protein-pairs (that do not break) within each interface category. Figure 4 illustrates a comparison of the average contact maps for four interfaces at the end of the indentation, namely, the dimer (A-B) interface, the asymmetric unit (A-C) interface, the pentamer (A-A) interface and the hexamer (B-C) interface (see Sect. 3.1 and Fig. 1).

Then we identified clusters of contacts in the contact map by image processing. We used binary thresholding to concentrate on the most prominent contact areas and subsequently applied a uniform filter to smooth out the boundaries and make them continuous to enable blob detection using an open source library OpenCV for the Python programming language. Note that even though the contact clusters are spread apart in the contact map, they are spatially close. Closer inspection reveals that the identified contact clusters show some variability. Not all clusters within any given interface category show up in each instance of a protein-protein contact; some are not even present in the initial crystal structure. However, for each interface category, a subset of contact clusters consistently appeared in each instance of an interface category, even after squeezing. These “persistent” clusters also fluctuated least over time (in terms of their median distance), and hence we posit that they are the most important ones for stability.

As observed earlier, the dimer (A-B) and (C-C) interfaces never broke. Structurally, these interfaces are formed by a strong contact between the C-termini of the two proteins, with the end of each tail fit well into a groove on the surface of the respective other protein. The contact map nicely visualizes this, with contact pattern of residue clusters at the edge of the map (see Fig. 4A). Often, tails are believed to be floppy, but here the tail of one protein closely attaches to its partner and stabilizes the dimer. As a consequence, the median distance of all contact clusters fluctuates less for the dimer (A-B) and (C-C) interfaces than for the pentamer (A-A) and the hexamer (B-C) and (C-B) interfaces. Already in Fig. 1 we had seen that these dimer interfaces are the ones with the highest number of contacts in the uncompressed virus’ crystal structure. In Table 1 we list the specific residue numbers that make up the contact clusters at each interface and the number of contacts that are either hydrophobic, hydrophilic or electrostatic. We list contact numbers for a 10 ns time window at the end of the squeezing simulation and for a 10 ns time window at beginning of the simulation (corresponding to the almost uncompressed capsid during the simulation, not the crystal structure; numbers are given in brackets in Table 1). As a criterion



**Fig. 4. Comparison of the contact maps at the endpoint of the indentation.** Average contact maps for the classes (A) Dimer (A-B), (B) Asymmetric Unit (A-C), (C) Pentamer (A-A) and (D) Hexamer (B-C) are shown. The contacts were averaged over the last 10 ns of all the non-breaking monomer pairs that belong to a certain category in our indentation simulations with IDEN elastic network with a compression speed of  $0.2 \text{ \AA ns}^{-1}$ . Table 1 lists all the “contacts” if the backbone beads (equivalent to  $C_{\alpha}$ ) of two residues are within  $6.0 \text{ \AA}$  of each other, classified according to interface categories. Note that the residue number shown in the contact maps refers to the  $\Delta 1-36$  mutant and would correspond to residues 37–190 in the wildtype structure 1ZA7 PDB.

to identify contacts we used the average distances between the  $C_{\alpha}$  beads of residues within each interface category. A contact was counted if the distance was smaller or equal to  $6 \text{ \AA}$ . The table highlights the large number of residues involved in contacts in the dimer (A-B and C-C) interfaces which explain their stability at maximum compression. About 10 contacts (nearly twice the percentage as any other interface) in each of the dimer (A-B) and (C-C) interfaces are made up of favorable electrostatic interactions between oppositely charged sidechains (salt bridges).

It had been observed before (Fig. 1A) that in the uncompressed virus (crystal structure) the hexamer (B-C) and (C-B) interfaces exhibit very similar contact

numbers compared to the pentamer (A-A) interface. Figure 4C and 4D now show that the contact patterns of the pentamer (A-A) and the hexamer (B-C) interface are almost identical (the same holds for the hexamer (B-C) interface which is not shown), and the average number of contacts of all such interfaces in the first 10 ns are also very similar (reported within brackets in Table 1). This is quite remarkable, since it means that these interfaces, which are located at completely different sites on the capsid surface, are constituted by the same types of protein-protein contacts. Note that these interfaces have the largest fraction of *hydrophobic* contacts among all interfaces (Table 1). This observation is even more remarkable, once it is being placed in the context of the stability of these interfaces upon compression of the virus and in the context of virus assembly. It has been suggested that viral capsid self-assembly is primarily driven by hydrophobic interactions, electrostatics and directional specificity imposed by polar and van der Waals interactions [55–57].

As discussed above, hexamer (B-C) and (C-B) interfaces frequently ruptured upon compression (regardless of the elastic network model, Fig. 3), while the pentamer (A-A) interface always remained intact. This fits to the observation one makes upon closer inspection of Fig. 4C and 4D, which shows that at the end of the compression the pentamer (A-A) interface has a larger number of contacts than the two hexameric interfaces. Given that pentamer (A-A), hexamer (B-C) and (C-B) interfaces start out with similar number of contacts in the first 10 ns, at maximum compression, hexameric interfaces lose a substantial number of its initial contacts (a majority of lost contacts are *hydrophobic* in nature) when compared to the pentamer (A-A) interface (Table 1), we propose that the geometry and the stress distribution at a symmetry site influences capsid stability rather than specific contacts between monomer pairs.

Curiously, even though there is nothing substantially different about the asymmetric unit interfaces (A-C) or (B-C), except that these interfaces have a slightly lower contact count—specifically, slightly lower hydrophobic contacts—when compared to the asymmetric unit (A-B) interface, they are the ones that break during compression in all our simulations. This is consistent with our previous observations that pentamer (A-A) and dimer interfaces never break, hence the monomers of the asymmetric (A-B) interface—that are part of PODs—do not rupture. This further supports our argument that the stress distribution at these symmetry sites influences capsid stability. Note that the number of contacts in monomer pairs of the asymmetric unit increase slightly during compression (compared to the initial 10 ns), while all the other interfaces lose their initial contacts (Table 1). We attribute this difference to the rearrangements of the asymmetric unit interfaces as a result of local stress distribution at this symmetry site upon squeezing the capsid. We find that all interfaces have a similar fraction of *electrostatic* contacts, except for the dimers (A-B and C-C). In contrast, the fraction of *hydrophobic* contacts in hexameric (B-C and C-B) and pentameric (A-A) interfaces are more than other interfaces.

## 5 Discussion and conclusions

AFM experiments on mechanical deformations of viral capsids provide valuable large-scale information, such as the indentation force required to deform and ultimately break the capsid [30]. However, these experiments do not have the resolution to actually identify the locus of breakage upon compression. In contrast, our CG simulations readily reveal where the capsid is most vulnerable to mechanical compression. Since the small-scale fluctuations of our protein dimers match those from atomistic simulations (by construction), while the large-scale emergent properties from our model—in particular the force-indentation relation—agree with experimental measurements, this

lends credence to our results. All the more so since our key findings do not depend on some details of the elastic network parametrization.

Our ability to identify the locus of breakage allows us to present a plausible order of hierarchical assembly of the CCMV capsid, in contrast to simulations using continuum models [58] where identifying the locus of breakage is impossible by construction. By correlating the stability of the interfaces with the thermodynamics of interfacial interactions, we argue that the kinetics of assembly follows the energetics of interfaces in the following manner.

1. The strongest contact interfaces are the dimer (A-B) and (C-C) interfaces, in terms of number of contacts that make up the interface (Fig. 1 and Table 1). We never see such dimers falling apart during mechanical compression. Indeed, it was proposed, based on experimental evidence, that these dimers assemble first [40].
2. The second strongest interaction (in terms of number of contacts) is that at pentamer (A-A) and hexamer (B-C) and (C-B) interfaces. Our contact maps show that these interfaces are formed by virtually the same amino acid contacts on the capsid protein surfaces. However, the large-scale geometry of association is different for these two interfaces, as illustrated for instance by the difference in  $d_{\text{cm,initial}}$  as well as by the number of contacts at high compression (Fig. 1 and Table 1). Intriguingly, upon compression only the hexamer (B-C) and (C-B) interfaces break, while the chemically identical pentamer (A-A) interfaces always stay intact. This observation suggests that the overall geometry of association significantly affects stability.
3. To rationalize this finding, we propose that the pentameric aggregate assembles readily, without much need to deform the individual proteins, while this is not so for the hexamer. A stable hexamer—whether freely formed or part of a full capsid—would hence be under internal *pre-stress*, which adds to the external stress imposed during compression, leading to an earlier failure of the hexamer. If so, an isolated hexamer should also have a higher energy than an isolated pentamer, and this would shift the balance during initial protein assembly towards pentamers, suppressing the frequency of free hexamers. This would be advantageous, because in the classical assembly path—pentamers of dimers being “glued together” by additional C-C dimers—hexameric units play no role. Additionally, recent experiments on CCMV capsid disassembly by Law-Hine et al. report the capsid breakage process to proceed through intermediate structures composed of aggregates of pentamer of dimers and free dimers [59].
4. In this classical assembly pathway, the interfaces that form last are the centers of asymmetric units, as well as the hexameric association site. Both of them ought to be weak (to ensure thermodynamically that they form last), and they indeed frequently break in our compression simulations (Fig. 3).

Our proposed model for hierarchical assembly order according to binding strength and mechanical/geometrical stability of the oligomers is not only thermodynamically plausible, but has advantageous consequences for the *fidelity* of assembly. Since several large pieces need to cooperatively move into place in order to “click in”, one might expect small misalignments to matter most in the viral capsid growth after nucleation. It is easy to imagine that this subtle alignment process would be greatly hindered if its underlying interactions were strong, since this could readily trap a conformation in metastable intermediates for which rearrangements become difficult.

This point is further supported by a well-known curious structural feature of the CCMV capsid [43]: a  $\beta$ -barrel formed at the center of the hexameric association site, where the C-termini of six capsid proteins come together; a similar feature is missing at the center of the pentameric site. We have recently confirmed in simulations [42] that such a barrel indeed only forms at a hexameric association site, because the

local geometry with which five C-terminal tails meet at the center of the pentameric site is not conducive to barrel formation, even if the  $\beta$ -propensity of the force field is significantly increased.

We propose that the presence of this  $\beta$ -barrel structural motif at the hexameric site is in fact a clever design feature of the capsid: the barrel is so sensitive to the precise positioning and orientation of its participating capsid proteins that it can only form once all six units have found their correct mutual placement [42]. Forming the  $\beta$ -barrel as an additional stabilizing motif will hence strengthen one of the weakest sites of the capsid not by simply increasing an interaction strength (which would interfere with assembly order), but instead by “closing a lock” when proteins at the hexameric site attain a specific geometry. Forming the barrel can hence also be viewed as an error check-point. This would allow the proteins at the hexameric site the ability to rearrange until they are in fact properly oriented. We should point out that in our simulation model the capsid protein was truncated at its N-terminal tail, such that the amino acids that would ultimately form the  $\beta$ -barrel are not part of the simulation. The breaking strength measured by us thus corresponds to  $\beta$ -barrel-free assemblies (i.e., before the lock is turned).

To summarize, we have presented an independent verification of a previously proposed mechanism of CCMV capsid self-assembly. We identify specific interfaces that are unstable during mechanical uniaxial compression of the capsid and propose intermediates in the assembly by studying the breakage pattern of the capsid, including the role of a stabilizing structural motif at the quasi-sixfold rotational symmetry site. Since most intermediates on assembly pathways remain challenging to characterize, our identification of specific residues forming contacts in the CCMV capsid is of significance for designing future experimental studies. Such combined efforts to identify and characterize factors that affect viral capsid assembly will influence the beneficial engineering of viruses and the design of antiviral drugs.

We dedicate this paper to Kurt Kremer, whose curiosity-driven approach to many soft matter questions has always been an inspiration for us, and to whom we are deeply grateful for his continuous support as a mentor and collaborator over many, many years. Kurt had and has a decisive influence on our view on soft matter physics, polymer theory, computer simulations and multiscale models. We are also grateful to the Volkswagen Foundation for supporting this project within the framework of the program “New Conceptual Approaches to Modeling and Simulation of Complex Systems” ([www.volkswagenstiftung.de](http://www.volkswagenstiftung.de)).

## References

1. F.H.C. Crick, J.D. Watson, *Nature* **177**, 473 (1956)
2. V.S. Reddy, P. Natarajan, B. Okerberg, K. Li, K.V. Damodaran, R.T. Morton, C.L. Brooks, J.E. Johnson, *J. Virol.* **24**, 11943 (2001)
3. D.L.D. Casper, A. Klug, *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1 (1962)
4. C.B. Frances, F.K. Thomas, J.B.W. Grahame, F.M. Edgar, D.B. Michael, R.R. John, K. Olga, S. Takehiko, T. Mitsuo, *Eur. J. Biochem.* **80**, 319 (1977)
5. P.E. Prevelige Jr., D. Thomas, J. King, *Biophys. J.* **3**, 824 (1993)
6. J.S. Stephen, R.B. Christina, P. Sreenivas, G.L. Warren, M.G. Finn, Z. Adam, *Proc. Natl. Acad. Sci.* **102**, 8138 (2005)
7. J.S. Stephen, Z. Adam, *J. Mol. Recognit.* **19**, 542 (2006)
8. W. Guo-Yi, Z. Xiao-Jing, Y. Chang-Cheng, J. Dong, Z. Ling, L. Yan, W. Lai, W. Yu, C. Hong-Song, *J. Chemother.* **20**, 458 (2008)
9. S. Jingchuan, D. Chris, D. Marie-Christine, M. Ayaluru, C. Chao, G. Kodetham, S. Barry, D. Mrinmoy, M.R. Vincent, H. Andreas, *Proc. Natl. Acad. Sci.* **104**, 1354 (2007)



10. C. Chao, D. Marie-Christine, T.Q. Zachary, D. Mrinmoy, S. Barry, D.B. Valorie, R.C. Paul, M.R. Vincent, C.K. Cheng, D. Bogdan, *Nano Lett.* **6**, 611 (2006)
11. J.Y. Pil, T.N. Ki, Q. Jifa, L. Soo-Kwan, P. Juhyun, M.B. Angela, H.T. Paula, *Nat. Mater.* **5**, 234 (2006)
12. Y. Ibrahim, S. Sourabh, F.S. Nicole, *Curr. Opin. Biotechnol.* **22**, 901 (2011)
13. L. LiNa, G.H. Richard, R.B. Veronica, A.L. Steven, F. Stefan, *J. Am. Chem. Soc.* **128**, 4502 (2006)
14. Y. Mark, W. Debbie, U. Masaki, D. Trevor, *Annu. Rev. Phytopathol.* **46**, 361 (2008)
15. A.A. Elizabeth, I. Steven, S.P. David, Y.W. Edwin, W.C. James, K. Kent, *Nano Lett.* **6**, 1160 (2006)
16. L. Andrew, N. Zhongwei, W. Qian, *Nano Res.* **2**, 349 (2009)
17. E.F. Christine, L. Seung-Wuk, R.P. Beau, M.B. Angela, *Acta Mater.* **51**, 5867 (2003)
18. P.P. Dustin, E.P. Peter, D. Trevor, *ACS Nano* **6**, 5000 (2012)
19. F.D. Sikkema, M. Comellas-Aragones, R.G. Fokkink, B.J.M. Verduin, J.J.L.M. Cornelissen, *Org. Biomol. Chem.* **5**, 54 (2007)
20. C.B. Chang, C.M. Knobler, W.M. Gelbart, T.G. Mason, *ACS Nano* **2**, 281 (2008)
21. M.T. Klem, D. Willits, M. Young, T. Douglas, *J. Am. Chem. Soc.* **125**, 10806 (2003)
22. P.A. Suci, M.T. Klem, F.T. Arce, T. Douglas, M. Young, *Langmuir* **22**, 8891 (2006)
23. M. Comellas-Aragones, H. Engelkamp, V.I. Claessen, N.A.J.M. Sommerdijk, A.E. Rowan, *Nat. Nanotechnol.* **2**, 635 (2007)
24. E. Gillitzer, P. Suci, M. Young, T. Douglas, *Small* **2**, 962 (2006)
25. P.A. Suci, D.L. Berglund, L. Liepold, S. Brumfield, B. Pitts, *Chem. Bio.* **14**, 387 (2007)
26. P.A. Suci, Z. Varpness, E. Gillitzer, T. Douglas, M. Young, *Langmuir* **23**, 12280 (2007)
27. C.R. Kaiser, M.L. Flenniken, E. Gillitzer, A.L. Harmsen, A.G. Harmsen, *Int. J. Nanomed.* **2**, 715 (2007)
28. Y. Ma, R.J.M. Nolte, J.J.L.M. Cornelissen, *Adv. Drug Delivery Rev.* **64**, 811 (2012)
29. G. Christoph, K. Venkatramanan, D. Markus, P. Christine, *PloS one* **8**, e60582 (2013)
30. J.P. Michel, I.L. Ivanovska, M.M. Gibbons, W.S. Klug, C.M. Knobler, G.J.L. Wuite, C.F. Schmidt, *Proc. Natl. Acad. Sci.* **103**, 6184 (2006)
31. A.J. Rader, D.H. Vlad, I. Bahar, *Structure* **13**, 413 (2005)
32. F. Tama, O. Miyashita, C.L. Brooks, *3rd J. Mol. Biol.* **337**, 985 (2004)
33. R. Konecny, J. Trylska, F. Tama, D. Zhang, N.A. Baker, *Biopolymers* **82**, 106 (2006)
34. D. Zhang, R. Konecny, N.A. Baker, J.A. McCammon, *Biopolymers* **75**, 325 (2004)
35. A. Arkhipov, P.L. Freddolino, K. Schulten, *Structure* **14**, 1767 (2006)
36. A. Arkhipov, W.H. Roos, G.J.L. Wuite, K. Schulten, *Biophys. J.* **97**, 2061 (2009)
37. C. Marek, O.R. Mark, *J. Chem. Phys.* **132**, 015101 (2010)
38. C. Marek, O.R. Mark, *PloS one* **8**, e63640 (2013)
39. M. Zink, H. Grubmüller, *Biophys. J.* **94**, 1350 (2009)
40. Z. Adam, A. Ryan, M. Jennifer, P.C. Johnson, J.Y. Mark, *Virology* **277**, 450 (2000)
41. E.B. Johanna, C.R.K. Heinrich, S.S. Ulrich, *BMC Biophys.* **5**, 22 (2012)
42. B. Tristan, G. Christoph, D. Markus, P. Christine, *J. Chem. Theory Comput.* **8**, 3750 (2012)
43. A.S. Jeffrey, B. Brian, Q. Chunxu, A.W. Deborah, J.Y. Mark, E.J. John, *J. Virol.* **80**, 3582 (2006)
44. J.A. Speir, S. Munshi, G. Wang, T.S. Baker, J.E. Johnson, *Structure* **3**, 63 (1995)
45. X. Zhao, J.M. Fox, N.H. Olson, T.S. Baker, M.J. Young, *Virology* **205**, 486 (1995)
46. J. Tang, J.M. Johnson, K.A. Dryden, M.J. Young, A. Zlotnick, *J. Struct. Biol.* **154**, 5967 (2006)
47. R.F.K. Bruinsma, S. William, *Annu. Rev. Condens. Matter Phys.* **6**, 245 (2015)
48. J.D.H. Perlmutter, F. Michael, *Annu. Rev. Phys. Chem.* **66**, 217 (2015)
49. S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, A.H. de Vries, *J. Phys. Chem. B* **111**, 7812 (2007)
50. L. Monticelli, S.K. Kandasamy, X. Periole, R.G. Larson, D.P. Tieleman, *J. Chem. Theory Comput.* **4**, 819 (2008)
51. X.C.M. Periole, S.J. Marrink, M.A. Ceruso, *J. Chem. Theory Comput.* **5**, 2531 (2009)

52. M. Seo, S. Rauscher, R. Pomes, D.P. Tieleman, *J. Chem. Theory Comput.* **8**, 1774 (2012)
53. B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **4**, 435 (2008)
54. M. Baaden, S.J. Marrink, *Curr. Opin. Struct. Biol.* **23**, 878 (2013)
55. M. del Alamo, M.G. Mateu, *J. Mol. Biol.* **345**, 893 (2005)
56. D. Chandler, *Nature* **437**, 640 (2005)
57. W.K. Kegel, P. van der Schoot, *Biophys. J.* **91**, 1501 (2006)
58. W.H. Roos, M.M. Gibbons, A. Arkhipov, C. Uetrecht, N.R. Watts, P.T. Wingfield, A.C. Steven, A.J.R. Heck, K. Schulten, W.S. Klug, *Biophys. J.* **99**, 1175 (2010)
59. D. Law-Hine, A.K. Sahoo, V. Bailleux, M. Zeghal, S. Prevost, P.K. Maiti, S. Bressanelli, D. Constantin, G. Tresset, *J. Phys. Chem. Lett.* **6**, 3471 (2015)