



OPEN

Dynamic stability of salt stable cowpea chlorotic mottle virus capsid protein dimers and pentamers of dimers

Janos Szoverfi^{1,2} & Szilard N. Fejer^{2,3}✉

Intermediates of the self-assembly process of the salt stable cowpea chlorotic mottle virus (ss-CCMV) capsid can be modelled atomistically on realistic computational timescales either by studying oligomers in equilibrium or by focusing on their dissociation instead of their association. Our previous studies showed that among the three possible dimer interfaces in the icosahedral capsid, two are thermodynamically relevant for capsid formation. The aim of the current study is to evaluate the relative structural stabilities of the three different ss-CCMV dimers and to find and understand the conditions that lead to their dissociation. Long timescale molecular dynamics simulations at 300 K of the various dimers and of the pentamer of dimers underscore the importance of large contact surfaces on stabilizing the capsid subunits within an oligomer. Simulations in implicit solvent show that at higher temperature (350 K), the N-terminal tails of the protein units act as tethers, delaying dissociation for all but the most stable interface. The pentamer of dimers is also found to be stable on long timescales at 300 K, with an inherent flexibility of the outer protein chains.

Virus-like particles (VLPs) are becoming an increasingly important tool for a variety of applications in biotechnology and medicine. In order to be able to create efficiently self-assembling VLPs, however, it is necessary to understand the processes that stabilize these particles and guide their self-assembly. We do not have such an understanding at the atomistic level yet, and the number of publications about modelling association of peptides^{1,2} or formation of protein–protein complexes^{3–5} on an atomistic scale is surprisingly small. The formation of protein–protein dimers can be explained as a diffusive association succeeded by the formation of a stereospecific complex⁶. The cowpea chlorotic mottle virus (CCMV) capsid is an ideal system to study atomistically, as multiple copies of the capsid protein are able to efficiently self-assemble into VLPs even in the absence of genetic material, while the self-assembly process itself is hierarchical and well-studied experimentally^{7–9}. It has been shown previously that for the whole CCMV capsid, the disassembly pathway is not a mirror of the assembly pathway¹⁰. Coarse-grained Monte Carlo simulations coupled with experiments show that in case of CCMV, the thermal dissociation is analogous to a two-dimensional phase transition, during which the dimer subunits on the capsid are in equilibrium with free dimers in solution¹¹. Conformational stabilities of monomers in pentameric and hexameric environments in the salt stable CCMV icosahedral capsid have been modelled previously as well¹². However, a better understanding of the dissociation process of CCMV dimers can give new insight into the dynamic stability of various dimers.

Several research groups revealed the importance of conformational fluctuations and solvation changes in the process of association and dissociation of proteins^{13,14}. All-atom simulations could reveal significant moments in the process of complex formation or dissociation, but due to the time and resource cost of such calculations just a few publications have been released in this topic. A study by Zhang et al.¹⁵ seeks to understand the dissociation of protein complexes through long time all-atom molecular simulations of two mutants, these are compared to previous simulations of the wild-type protein by the same group¹⁶. The majority of simulations showed that protein dissociation occurred within 2.4 μ s. After the analysis of the solvation energies Zhang et al. observed that the dissociation correlates with loss of protein–protein contacts in the complex. Thermodynamics of the separation is influenced significantly by the internal protein dynamics. The final conclusion of the research shows that

¹Faculty of Chemical Engineering and Biotechnologies, University Politehnica of Bucharest, 1-7 Gheorghe Polizu Street, 011061 Bucharest, Romania. ²Provitam Foundation, 16 Caisului Street, 400487 Cluj-Napoca, Romania. ³Institute of Chemistry, University of Pécs, 6 Ifjúság Street, Pécs, Hungary. ✉email: szilard.fejer@cantab.net

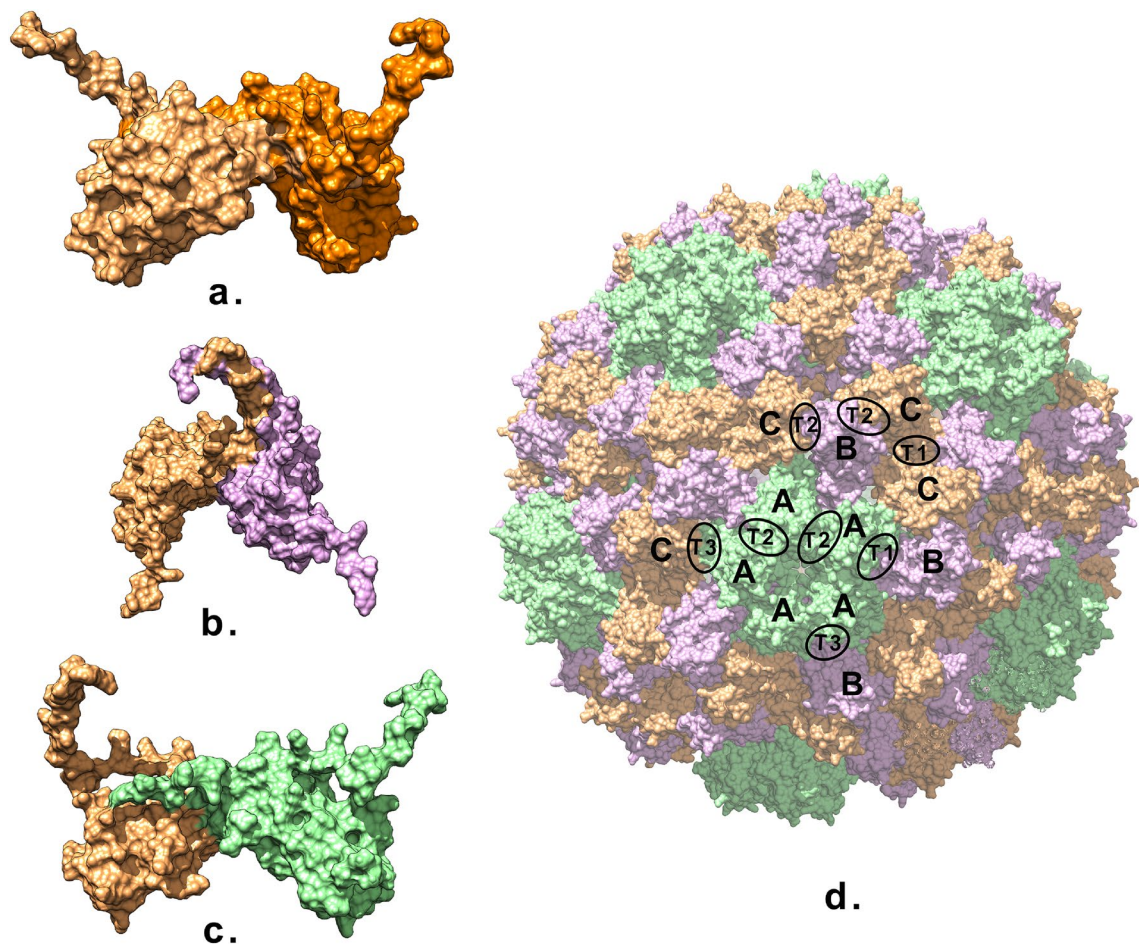


Figure 1. (a) T1, (b) T2 and (c) T3 dimers from the original capsid; (d) the whole icosahedral ss-CCMV capsid, coloured by chains, showing the positions of the three interface types between the protein monomers.

the process of dissociation usually follows multiple pathways and involves transitions to less favorable interface contacts. Banerjee et al. analyzed the contribution of the dynamics of the water in the dissociation or association process of insulin monomer¹⁷. They achieved the separation of insulin dimers and showed that water molecules play an important role in the separation or association of proteins.

CCMV is an icosahedral plant virus with 12 pentameric and 20 hexameric capsomers, with a triangulation number of 3¹⁸. The coat protein subunits are comprised of an eight stranded, antiparallel, β -barrel core. The capsomers are probably stabilized by the linkage through the extensions of the C-terminal from coat proteins¹⁸. The N-terminal region from one protein clamps the adjacent protein's C-terminal arm. It is thought that a hexameric tubular structure, called β -hexamer, plays an important role in the stability of the capsid through the the N- and C-terminal extensions of the coat proteins contributing to the formation of the quaternary structure, while the hexameric tubular structure allows for tight clustering of neighbouring monomers into capsomer structures¹⁸. Chen et al. studied the interactions between the components of the CCMV capsid¹⁹ and found that flexible N-terminal tails are not only involved in the formation of protein dimers, but responsible for kinetic traps as well. The observed stability of the pentamer and hexamer of dimers suggests that the dimers themselves serve as the basic building blocks throughout the capsid formation process. A point mutation (K42R) in the N-terminal tail increases the capsid stability under high ionic strength conditions (salt stable CCMV)^{20,21}.

Our aim was to investigate the dynamic stability of oligomers that can act as building blocks for the salt stable CCMV (ss-CCMV) capsid²¹, namely the changes happening on the protein–protein interface. We performed *in silico* simulations under different conditions to achieve the dissociation or destabilization of protein dimers in various configurations compatible with the icosahedral ss-CCMV capsid architecture. All-atom molecular dynamics simulations were performed in implicit and explicit solvent at different temperatures to understand in detail the processes that could lead to the dissociation of a dimer, and, implicitly, give us insight into the relative dynamical stabilities of the various dimers. The simulated dimers were analyzed with respect to changes in the binding interface of the proteins (Fig. 1). A pentamer of dimers was selected from the icosahedral virus capsid and simulated in similar conditions as the dimers, in order to gain information about the behaviour of interfaces in a more restricted environment.

Results and discussion

Long timescale MD simulation for T1, T2 and T3 interfaces. Three dimers with different types of interfaces were selected from the ss-CCMV protein capsid (PDB id 1za7, 165 residues) to study their behaviour during all-atom MD simulation in explicit water. Dimers from the original, icosahedral virus capsid were used to characterize the initial interfaces using the Pymol InterFaceResidue plugin and the CoCoMaps webserver²², with interaction distance cutoff set at 5 Å²³. Note that this is an inter-residue cutoff distance, so two residues are defined to be in contact whenever any two atoms belonging to the different residues are closer than this cutoff. The types of interactions are defined by CoCoMaps according to the polar nature of the residues in contact: hydrophobic residues are non-polar (e.g. Ala, Val, Leu, Ile, Thr), hydrophilic residues have polar sidechains (e.g. Glu, Asp, Ser, Thr, Asn).

As the first 25 residues are missing from the 1za7 pdb structure for chains B and C, in order to have continuous numbering for the dimer structures, we restarted numbering of the chains starting from residue 26. Our numbering scheme therefore goes from 1 to 330 for the dimers. Note that for chain A, 39 residues could not be resolved in the crystal structure, so we constructed the three possible dimer interfaces from chains B and C.

The T1 dimer has the largest interface between the three types, with an interface surface area of 2957.6 Å², 14% of the total surface area of two monomers, and it involves 40 residues out of 330. The interactions stabilizing the connection of the monomers are mainly hydrophilic–hydrophilic interactions (36), hydrophobic–hydrophobic interactions (40) and hydrogen bonds (18). The stabilizing hydrophilic–hydrophilic interactions have a polar character, an electronegative atom (oxygen or nitrogen) being in contact with a hydrogen atom in a position that does not satisfy criteria for hydrogen bonds. The dimer has C₂ rotational symmetry, the angle between the two monomers is 72°. The interface surface area of T2 dimer is 2034.7 Å², 9.8% of the total surface area of two monomers. 31 residues are involved in this interface. The angle between the two monomers is 34.6°. There are 39 hydrophobic–hydrophobic interactions, 12 stabilizing hydrophilic–hydrophilic interactions and 12 hydrogen bonds between the monomers.

The T3 dimer has 13 residues on the interface. The buried area upon complex formation is 838 Å², 4.05% of the total surface area of two monomers. 16 stabilizing hydrophilic–hydrophilic interactions, 4 hydrophobic–hydrophobic interactions and 3 hydrogen bonds are identified between the monomers.

2 μs NPT ensemble molecular dynamics simulations were carried out at 300 K for T1, T2 and T3 dimers (see Supplementary Movies 1–3). The simulations for T1 and T2 were repeated with different initial velocities, but as the results were similar for the first microsecond, the process was stopped to save computational time. The complete trajectories for all explicit solvent simulations of T1 and T2 are shown in Supplementary Fig. 7. RMSD calculations were carried out for the C_α backbone of the proteins and for the binding interface residues identified in the 1za7 PDB structure.

C_α RMSD values for T1 show a bigger fluctuation during the simulation and reach a value of about 1.1 nm in the first 300 ns, stabilizing after that at around 0.9 nm (average 0.94 ± 0.09 nm). C_α RMSD of T2 grows to an average value of 0.82 ± 0.1 nm. The average C_α RMSD of the T3 dimer is 1.1 ± 0.73 nm, having an accentuated fluctuation compared to that of T1 or T2 (Fig. 2a).

The tail regions of the dimers have specific roles in the stabilization of the dimers and the formation of the capsid²⁴. In both the T1 and T2 dimers the regions with the highest movement are the terminal regions, but their contribution to the C_α RMSD is different for T1 and T2, due to their position. In case of the T1 dimer, the C-terminal residues are located mostly on the interface between the monomers, while the N-terminal residues can move more freely. The terminal regions of the T2 dimers are located in the outer regions of the protein dimer. The average C_α RMSD value calculated for the dimers without the flexible tail regions is 0.71 ± 0.1 nm for T1 and 0.58 ± 0.08 nm for T2 (Fig. 2a). The C_α RMSD value of 0.71 nm for the main body of the T1 dimer is remarkably similar to that of previous 500 ns long simulation results of wild type CCMV dimers¹⁹.

The interface RMSD (iRMSD) values calculated for each trajectory show the movement of atoms from the residues present on the contact area between the two proteins in each dimer (Fig. 2b). The average iRMSD for T1 (0.6 ± 0.08 nm) is much lower than that for T2 (1.14 ± 0.15 nm). This indicates that the T2 interface is structurally less stable than T1, this being in a good agreement with our previous study²⁵. Within the first 50 ns of the simulation the values quickly increase to over 0.7 nm. Interface RMSD values for the T1 dimer show a deviation from the average between 1500 and 1950 ns, afterwards returning to the average values. The change is caused by a temporary large-scale movement of the hairpin between residues 10 and 24, away from the body of the protein. We continued the simulation only for this trajectory to a total length above 2.4 μs and found that the average iRMSD of 0.6 nm does not change within this timeframe for the T1 interface (see Supplementary Fig. 3).

The N-terminal tail regions of the monomers are in close contact in the T2 dimer and are located away from the body of the monomers, therefore the tail regions are quite mobile. After 200 ns simulation this tail gets glued to the main body of the dimer and remains there for the time of the simulation, while the interface between the two tails is also completely reorganised. Due to these changes, the iRMSD of the T2 dimer reaches very high values (avg 1.14 ± 0.16 nm), compared to that of the starting structure, which wrongly suggests an accentuated instability of the interface. For this reason, the iRMSD was recalculated without the first 14 residues from the N-terminal (residues missing from the A chain of the protein asymmetric unit in the PDB structure), and we found that the contribution of the N-terminal tail to the iRMSD is around 30%.

Per-residue fluctuations were calculated to monitor the motion of the residues (Fig. 2c). Interface residues show a higher mobility for the T1 dimer due to the fact that several exposed residues are part of the terminal regions. Regions containing β-sheets from the core have a lower mobility, although the outer turns present bigger motions. Tail regions show much higher values both on the N- and C-terminal parts of the monomers (see Supplementary Fig. 1). Several residues of the β-barrel are on the protein–protein interface. The high iRMSD values observed for T2 are due to the accentuated movement of tail regions, as more such residues are located on the

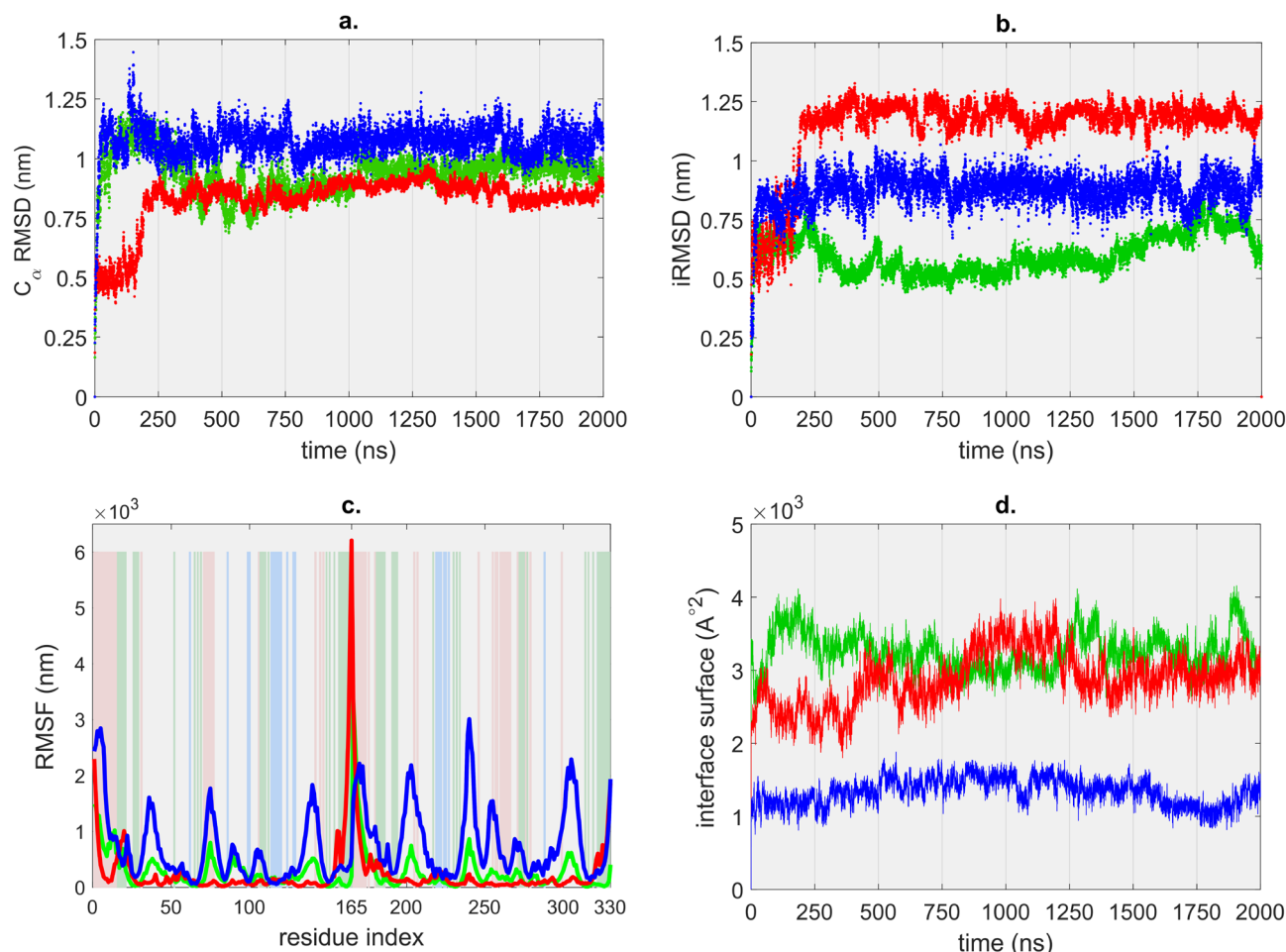


Figure 2. (a) C_{α} RMSD, (b) interface RMSD, (c) per-residue fluctuation (background colour bars represent the interface residues for T1 (green), T2 (red), T3 (blue)), (d) change of the interface surface area during the trajectory for T1 (green), T2 (red) and T3 (blue) dimers.

interface. In case of the T3 dimer the overall fluctuation is bigger than for the others, as the relative orientation of the chains and the small interface allow for bigger motions. The N-terminal tails are located away from the body of the dimer, and are not involved in the forming of the interface, so they can move freely, shown by the high fluctuation values. The peaks shown in Fig. 2c between those for the terminal regions represent motions of the turns from the β -barrels in the outer region of the dimer, these are most pronounced for T3.

The solvent accessible surface area (SASA) of the dimers is relatively constant throughout the simulation for each dimer. In our case the examination of the binding surface is more relevant. The interface surface was calculated as the difference of the dimer SASA and that of the standalone monomers and is shown in Fig. 2d. In all cases the interface surface change through each trajectory shows a slow increase with high fluctuation. The interface surface of the T1 dimer is initially higher than that of the T2 dimer, increasing from the initial 14% of the total surface to 19%. The large oscillation of the values shows that the solvent accessibility of the interface changes continuously due to small-scale movements. The surface of the T2 interface is smaller but after 1000 ns of simulation increases to that observed for T1 with an accentuated increase between 800 and 1200 ns. The interface surface of T3 stays relatively constant with low oscillation of the values.

The behaviour of the monomers relative to each other was investigated by tracking the change in distances, number of contacts (polar and non-polar interactions), number of hydrogen bonds, and the angles defined by the principal axis of inertia of the monomers.

The distance between the center of mass of the monomers in case of T1 is constant (3.8 ± 0.13 nm) with a temporary decrease between 400 and 700 ns. Both the number of contacts (293 ± 34) and the hydrogen bonds (25 ± 4) shows a slow increase. However, we observe a significant variation during the first 1000 ns in the angle between the two monomers, decreasing to 33° and subsequently increasing to 90° , eventually oscillating around the original 70° for the rest of the trajectory. This fan-like motion was previously described by Globisch et al.²⁶

As we did not observe drastic changes on the interfaces of three types of dimers during the long simulations, we chose another, putative dimer from the docking results of our previous work²⁵ that had the biggest RMSD score (3.09 nm) when compared to the original T3 dimer in order to start the simulation from a dimer that is not present in the icosahedral capsid and is therefore far from any native protein-protein interface for ss-CCMV. The chains in this dimer (type X, TX) are bound together through the first 15 residues of the N-terminal tail.

Molecular dynamics simulations were carried out under the same conditions than for the other dimers (see Supplementary Movie 4). As only the N-terminal tails are in contact in the TX dimer, the flexibility of the dimer is rather high (see Supplementary Fig. 4). The fluctuation of the residues is substantially larger than for the other 3 (native) dimers. After 100 ns of simulation the C_{α} RMSD reaches 1.4 nm and stabilizes around this value. The monomers are coming closer to each other and at a moment they are forming contacts between the turn regions of the β -barrel cores. The N-terminal tails remain in close contacts and stick to the body of each monomer. The number of contacts and hydrogen bonds increase constantly through the simulation. At the starting point of the simulation only 8 hydrogen bonds are found, mainly between the N-terminal tails. During the simulation an average of 16 ± 3 bonds are present, with new connections being formed between the loop regions of β -barrels and the tails and the body of dimers. The distance between the monomer chains is reduced from 4.5 to 3.9 nm and the angle between them from 90° to 10° . The structural stability of the TX dimer is therefore much lower than that of any native (T1, T2, T3) dimers.

In order to assess the possible mechanistic reasons for the enhanced stability of ss-CCMV compared to the wild type CCMV capsid, we analyzed the contacts and hydrogen bonds of the arginine residue located at the point mutation site of ss-CCMV (ARG17—chain C, ARG182—chain B). Note that ARG17 and ARG182 are at position 42 of chains C and B of the ss-CCMV protein when using the original numbering convention (see Supplementary Fig. 2). Throughout the 2 μ s NPT simulations, the arginine residues of the T1 dimer have the most contacts: ARG17 is in contact with PRO328 (for 0.52 μ s in total along the trajectory), has a hydrogen bond with ASP324 (0.14 μ s); ARG182 contacts with PRO163 (1.98 μ s), VAL164 (1.82 μ s), TYR165 (1.86 μ s), has a hydrogen bond with TYR165 (0.2 μ s). In the T2 dimer, ARG17 is in contact with PRO279 (1.0 μ s), GLU280 (0.6 μ s), has a hydrogen bond with GLU316 (0.4 μ s); ARG182 has a hydrogen bond with GLU9 (0.4 μ s). For the T3 dimer, ARG17 has no contacts, while ARG182 contacts only with TYR165 (1.8 μ s). It is therefore possible that the point mutation stabilizes the T1 dimer the most, compared to the T2 and T3 dimers. However, in order to better assess the possible reason for structure stabilization, contact environments should be directly compared between wt-CCMV and ss-CCMV. Unfortunately, the lower resolution of the wt-CCMV crystal structure (PDB ID 1cwp) and the difference in missing residues complicates such attempts. A more straightforward way to compare the two environments would be to use the optimised ss-CCMV dimer structures as templates, reverting the mutation to wild type, reoptimising the structures and redoing the long simulations.

Behaviour of T1, T2, T3 and TX interfaces at 350 K. A remarkable property of the wild type CCMV capsid is that it is stable at 75°C at pH 4.5 and ionic strength of 0.1 M^{27} . We chose a temperature slightly above this temperature to see if dissociation is possible in the three dimer configurations, as during capsid disassembly some protein–protein interfaces must disappear. No experimental data is available for the melting temperature of the ss-CCMV capsid. However, all three brome mosaic virions (a virus related to CCMV) have remarkably similar melting temperatures to that of wild type CCMV (about 70°C)²⁸. We expect similar dissociation/melting temperatures for ss-CCMV as well. Note that capsid melting does not necessarily mean complete dissociation into monomers in experiments, so melting and dissociation temperatures are not necessarily equivalent. As these temperatures depend strongly on pH as well, simulations have to aim higher than the highest experimental temperature to achieve protein–protein dissociation that can destabilize a fully formed capsid. We therefore performed 4 parallel molecular dynamics simulations at 350 K in implicit solvent for each of the T1, T2, T3 and TX dimers to see which interface can be most destabilized in the absence of explicit water molecules. The simulations had the same starting structures, and only differed in the initial velocities, while other parameters were similar to those in the simulations with explicit water. The T1 dimer proved to be relatively stable at this temperature (Fig. 3a), with a bigger fluctuation of the residues (C_{α} RMSD of 0.99 ± 0.11 nm) and minor changes on the interface (iRMSD of 0.61 ± 0.1 nm). The biggest contribution to the overall movement is given by the N-terminal tail that can move freely at the beginning of the simulation.

In case of the other three dimers (T2, T3 and TX), dissociation events occurred in at least one of the four parallel runs (Fig. 3b–d, see also Supplementary Movie 5). We observed two different mechanisms for dissociation of the dimers at 350 K. Complete dissociation took place swiftly in one simulation, while in the three other trajectories the dimer reached a near-dissociated state during the first 5–10 ns of the simulation. The interaction between the N-terminal tails still maintains a small interface, while the monomers keep moving away but stay hooked together at the same time through this small interface (see Supplementary Fig. 5). By 75 ns the dissociation of the dimer is complete in these trajectories as well. While both the SASA and the interface surface is slowly increasing for T1 throughout the simulation, we observe an increase of SASA and a large decrease of interface surface for three of the four simulations just before complete dissociation (Supplementary Movie 5) or the formation of the tether between the two tails. In case of the remaining 8 trajectories in total for T2, T3 and TX, the N-terminal tails of the dimers remain attached and prevent the complete separation of the two chains throughout the simulations. Supplementary Figure 6e shows the last frames of each high-temperature trajectory in which dissociation is not complete.

The N-terminal tails act as a tether and prevent dissociation in three different dimer configurations (T2, T3 and TX), suggesting that this large, mobile part of the protein can in fact play an important role during the self-assembling process, increasing the effective concentration of interacting building blocks by forming weakly bound flexible intermediate states and allowing the chains to find the most favourable native contacts. As the RNA binding domain (residues 1–25) is missing from the ss-CCMV structure used here, the real N-terminal region is considerably longer, and therefore the tethering effect aiding assembly of empty capsids in vitro might be important on double the length scale observed in our simulation. As the N-terminal domain is essential for RNA binding, such a tethering effect is likely important for the assembly of empty capsids only, in solutions

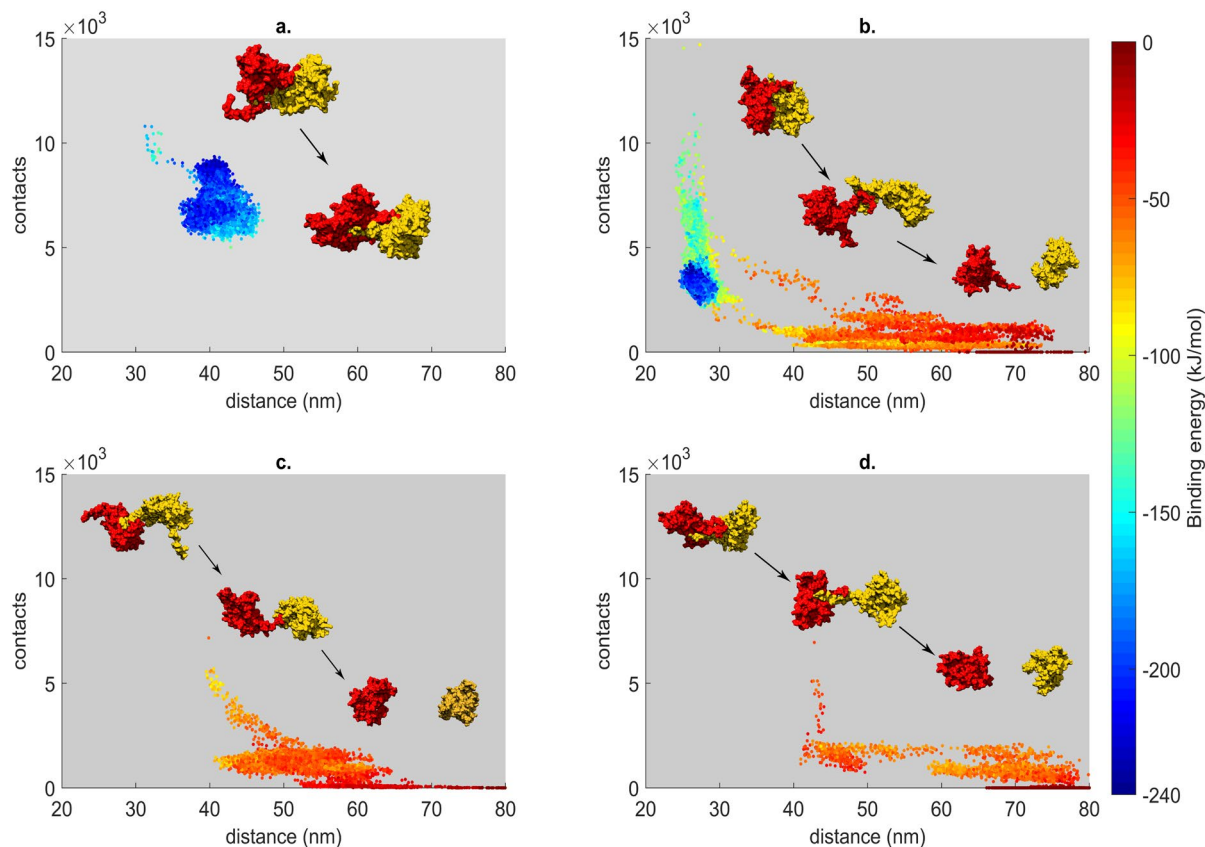


Figure 3. Binding energy landscape as a function of monomer contacts and center of mass distances, with initial and final structures presented, of the implicit water MD simulations for T1 (a), T2 (b), T3 (c) and TX (d) at 350 K. All frames from the four trajectories per dimer were used to generate these graphs.

with no RNA present, speeding up the association of T1 dimers into the pentamer of dimer seed structure, and helping the subsequent addition of T1 dimers to the seed.

After we observed that in certain conditions the dissociation of dimers is possible, we decided to inspect the behaviour of the dimers at 350 K in explicit water and constant pressure. However, during the 600 ns simulations all dimers remained associated (Fig. 4). The T2 dimer presents C_{α} RMSD value of 0.64 nm, lower than that observed at 300 K. In contrast, T3 and TX have high C_{α} RMSD values at 350 K. These two dimers are attached mostly through their N-terminal tails with relatively small binding interfaces, therefore an accentuated motion of the protein chains is possible. The interface surface area is growing in all simulations of the T2, T3 and TX dimers. We performed normal mode analysis for the four simulations (Fig. 4d–g). The low-frequency mode of the T1 dimer shows the fan-like motion described previously with minor loss in secondary structure. The T2 interface presents a twisted relative motion of the subunits, with an accentuated motion of the terminal residues. The monomers of the T3 dimer are moving more independently to each other, the lowest mode being similar to a rotatory motion of one of the chains. The character of motions is similar in case of the TX and T3 interfaces, the TX interface changes continuously during the simulation.

Although the different behaviour of high temperature simulations in explicit and implicit solvent is likely due to multiple factors (simulations are being done on essentially different energy landscapes due to the different solvent models), the inherent viscosity of explicit solvent models is an important factor hindering dissociation. The dimer in implicit solvent simulations does not have to ‘push away’ water molecules, and that can explain why the T2, T3 and TX dimers readily dissociate. Explicit and implicit models have been compared previously by Anandkrishnan et al.²⁹ and it was found that the implicit model can be up to 100-fold faster than the explicit model for reproducing large-scale conformational changes.

Long timescale MD simulation for the pentamer of dimers. Trajectories of the two parallel 2 μ s simulations (10000 frames, 0.2 ns/frame) were analyzed, and are shown in Fig. 5a (see also Supplementary Movie 6). C_{α} RMSD was calculated for both of the trajectories with reference to the first frame. The RMSD values grew rapidly in the first 30 ns of simulation, afterwards a fluctuation between 0.1 and 1.4 nm was observed with an average of 1.10 ± 0.12 nm for the first run and 1.21 ± 0.12 nm for the second run, respectively. Figure 5b shows the residue fluctuations of the inner and outer chains in the PD. Hydrogen bond and angle variations during the simulations are illustrated in Fig. 5c,d, respectively.

The pentamer of dimers (PD) is stable throughout the simulation, which is in good agreement with previous observations suggesting that the PD is a stable subunit during the assembly of the full icosahedral capsid³⁰. For

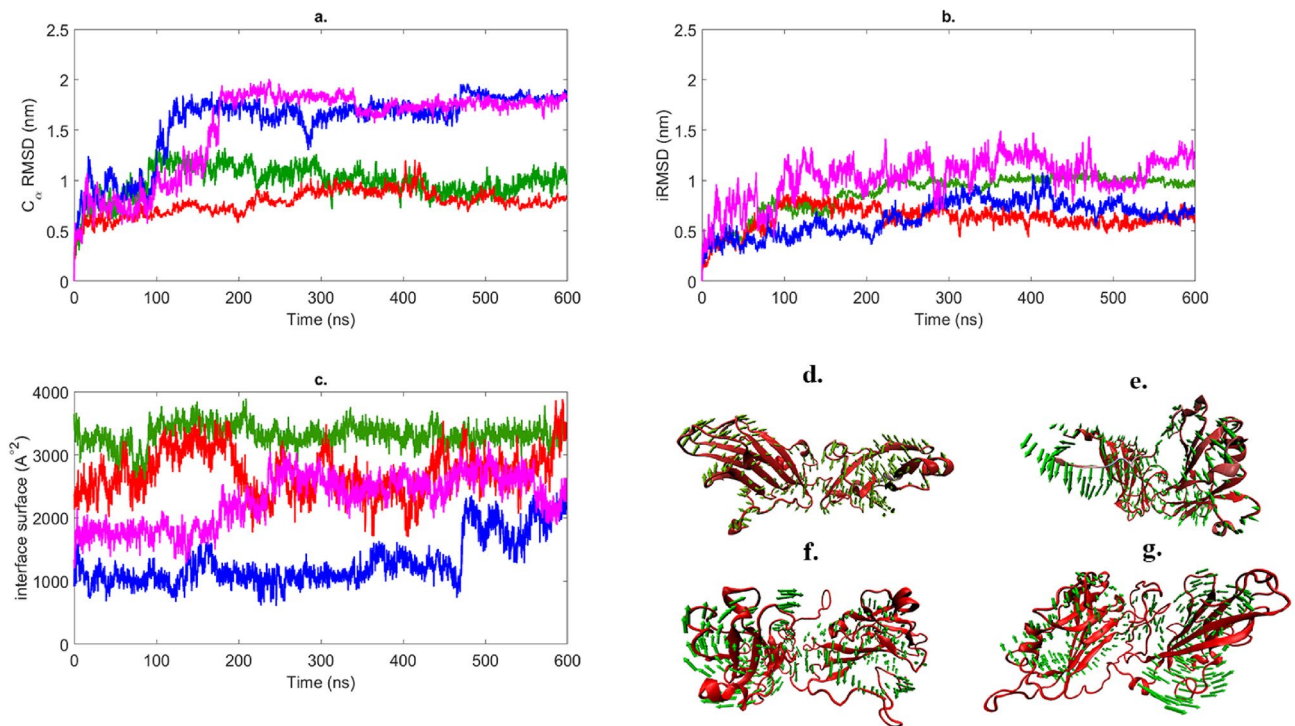


Figure 4. (a) C_{α} RMSD, (b) iRMSD and (c) interface surface for T1 (green), T2 (red), T3 (blue) and TX (purple) for 600 ns NPT ensemble in explicit water on 350 K. Structures (d–g) show the lowest frequency motions for T1 (d), T2 (e), T3 (f) and TX (g).

a better view of the changes during the simulation, we analysed the dimers contained in the PD by the interface type, as there are 5 dimers of T1 interface (AB–IJ) and 5 dimers of T2 interface (AC–IA) in this structure. The average C_{α} RMSD and average interface RMSD values for the dimers within the PD are summarized in Table 1.

The large differences between the two types of interface is due to the fact that the T2 interfaces are located in the inner circle of the PD, thus their mobility is reduced. T1 interfaces are between a protein from the inner chains and another one from the outer circle. The outer chains have overall less contacts, therefore they are more mobile. Clearly, this mobility is present only in the PD configuration, and the whole PD structure must be as rigid as the inner part of the pentamer of dimers, when incorporated in the icosahedral capsid. The flexibility of the outer chains of the standalone PD might be useful to enable the subsequent addition of T1 dimers during the capsid assembly.

Among the T1 dimers, the one labelled EF behaves differently from the others. After 500 ns simulation an increase of 0.4 nm both for C_{α} RMSD and iRMSD can be observed. The outer monomer moves out of the plane of the PD (see Supplementary Fig. 6). The movement is similar to the fan-like motion of the standalone T1 dimer. The N-terminal regions of the inner monomers are arranged in a specific conformation, forming hydrogen bonds with the tail regions of two neighboring chains on one side. The arrangement is conserved during the simulation, thus the role of the N-terminal region in the PD formation and stabilization is emphasized²⁴. The N-terminal tails of the outer chains are sticking to the body of the monomer in the initial phase of the simulation.

To understand the movement of the protein chains, we calculated the angles between adjacent chains using as reference an axis for each of the monomer bodies (monomers without the terminal regions). The average angles for two dimers (AB, EF) were similar to those of the original T1 dimer, while the rest showed lower values. Our results suggest that the EF dimer from the PD is behaving like a standalone T1 dimer.

A small decrease in the number of hydrogen bonds can be observed for two dimers (AB, EF), while the other T1 dimers had similar values as for the standalone dimers. The average angles between the chains in T2 dimers are similar to those in the initial T2 dimer. The fluctuation of the values is different for the 5 dimers. The EG and GJ dimers behave similarly to the standalone T2 dimer.

Conclusions

Our long molecular dynamics simulation results demonstrate the increased dynamic stability of the T1 interface compared to the other two native interfaces that exist in the icosahedral capsid configuration. The difference in stabilities becomes more evident at high temperature (350 K), where dimers connected by less stable, T2 or T3 interfaces, dissociate or remain only loosely associated via the flexible N-terminal regions being hooked together. Although proper dissociation is achieved only using implicit solvent models, the high-temperature explicit solvent simulations also point to the enhanced stability of the T1 interface compared to the other two.

We have also shown that an important assembly intermediate, the pentamer of dimers is a stable structure on the timescale of the simulation, with more flexible outer regions and possible fan-like motions localised on

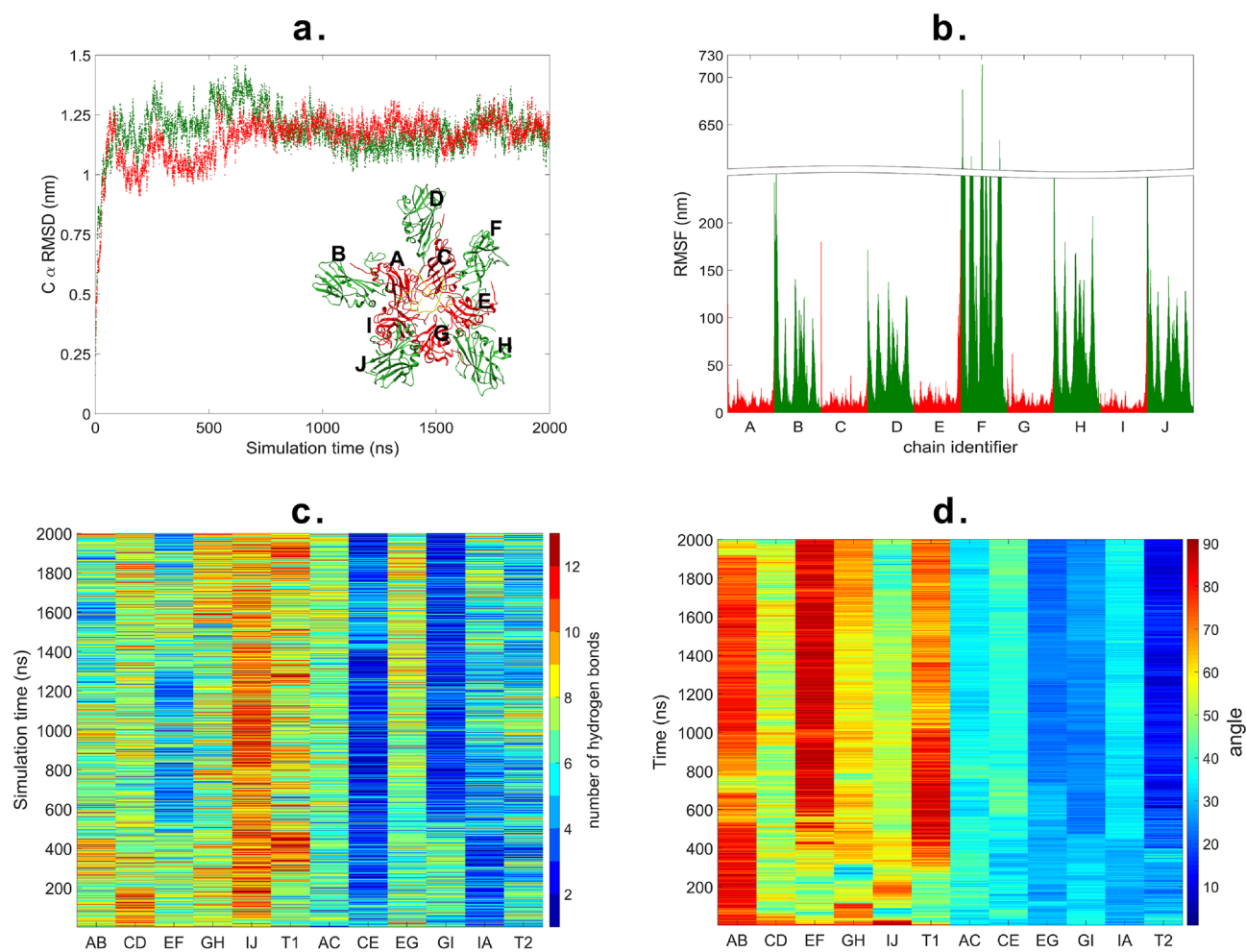


Figure 5. (a) C_{α} RMSD of the PD from two parallel MD simulations starting from different random velocities from the same structure for 2 μ s; (b) residue fluctuation in the PD for the inner chains (red) and outer chains (green); (c) variation in the number of hydrogen bonds for the separated T1 dimers (AB ... IJ), standalone T1 dimer (T1) and separated T2 dimers (AC ... IA), standalone T2 dimer (T2); (d) variation in the angle between the body of monomers in the separated T1 dimers (AB ... IJ), standalone T1 dimer (T1) and separated T2 dimers (AC ... IA), standalone T2 dimer (T2).

	C_{α} RMSD (nm)	iRMSD (nm)
T1 AB	0.77 ± 0.07	0.29 ± 0.04
T1 CD	0.79 ± 0.08	0.38 ± 0.05
T1 EF	1.09 ± 0.19	0.69 ± 0.18
T1 GH	0.89 ± 0.07	0.4 ± 0.06
T1 IJ	0.82 ± 0.07	0.34 ± 0.05
T2 AC	0.3 ± 0.03	0.21 ± 0.02
T2 CE	0.36 ± 0.04	0.26 ± 0.05
T2 EG	0.35 ± 0.04	0.21 ± 0.02
T2 GI	0.42 ± 0.06	0.32 ± 0.05
T2 IA	0.28 ± 0.03	0.21 ± 0.02

Table 1. Average C_{α} RMSD and iRMSD for each dimer of the pentamer of dimers.

T1 dimers. The inner part of the PD is more rigid, while the flexibility of the outer part might be important in the next step of capsid self-assembly, the subsequent addition of T1 dimers to the PD seed.

Methods

Studied systems. Dimers of ss-CCMV capsid protein were selected according to our previous study²⁵. The T1 and T2 interfaces play a significant role in the capsid formation, while T3 seemed less relevant. Type X is the dimer that had the highest RMSD compared to the original T3 dimer, selected from the 2000 docking results of two monomers.

Every dimer in the study has 4996 atoms in 330 residues. Identification of interfacial residues and characterization of the protein-protein interfaces in the three dimers was made using the InterfaceResidues Pymol script³¹. All structures were prepared with LEaP³². Solvation was done with an octahedral box of TIP3P³³, where water molecules were added to a distance of 10 Å from the protein. The overall negative electrostatic charge was redistributed among all atoms in explicit water simulations. The systems contained the same number of waters and the total charge was neutralized with 6 added Na⁺ ions.

A pentamer of dimers (PD) was selected from the capsid of ss-CCMV. The structure contains 24,980 atoms in 1650 residues. The inner part of the PD consists of 5 chains connected with the outer chains through T1 interfaces, while T2 interfaces are present between them. For a detailed analysis of the simulations, the PD was segmented into 10 dimers by the interface between the monomers: T1 interfaces (AB-IJ) and T2 interfaces (AC-IA). The PD was solvated with an octahedral box of TIP3P³³ with water molecules to a distance of 10 Å from the protein. The system was neutralized with 30 Na⁺ ions.

Molecular dynamics simulations. The same MD protocol, adapted from Mafucci et al.³⁴ was used for all NPT simulations.

- Hydrogen atoms were minimized for 1000 cycles of steepest descent and 5000 cycles of conjugated gradient, while the rest of the atoms were restrained.
- Water molecules and ions were relaxed using 2000 cycles of steepest descent and 5000 cycles of conjugated gradient.
- The solvent box was equilibrated at 300 K by 100 ps of NVT and 100 ps of NPT simulation using a Langevin thermostat and Berendsen barostat, respectively, with restraints on every atoms, excluding waters.
- The energy of side chains and water molecules was then minimized with backbone restraints of 25 kcal/mol.
- Complete minimization with backbone restraints of 10 kcal/mol (2500 cycles of steepest descent and 5000 cycles of conjugated gradient).
- Heating up the system to 300 K in 6 steps of 5 ps each ($\Delta T = 50$ K), where backbone restraints were reduced from 10 kcal/mol to 5 kcal/mol.
- Full equilibration in the NVT ensemble (100 ps, backbone restraints of 5 kcal/mol) and in the NPT ensemble (1 step of 200 ps, backbone restraints of 5 kcal/mol; 3 steps of 100 ps each, reducing the backbone restraints from 5 to 1 kcal/mol, and 1 step of 1 ns with 1 kcal/mol of backbone restraints).

Molecular dynamics simulations were performed with the pmemd.cuda module of the Amber14 software package on NVIDIA Tesla K40 GPU. The force field used was ff03³⁵. Pressure equilibration was made on manually edited systems by running NPT simulation several times on high pressure (1000 bars) for 20 ps until a density of ≈ 1.00 kg/dm³, all atoms were restrained, excluding waters. An electrostatic and nonbonded cutoff of 8 Å, a Berendsen barostat, PME for long-range electrostatic interactions, and the SHAKE algorithm were applied to all calculations. Long timescale all-atom simulations were carried out in explicit water for the PD and analyzed for a better view of conformational changes in the CCMV capsid. Two simulations of NPT molecular dynamics were conducted for 2 ms in explicit water with different initial velocities, on constant temperature of 300K with Langevin thermostat and constant pressure of 1 bar with Berendsen barostat. The timestep used was 2 fs during the 2 ms simulation.

Production runs were performed in NPT ensemble with pressure set to 1 bar, as follows:

- 2 μ s on 300 K for T1, T2 and T3.
- 1.5 μ s on 300 K for TX.
- 600 ns on 350 K for T2, T3 and TX.

Implicit solvent simulations were performed with the OBC variant of the Generalized Born model (igb = 5)³⁶, with no electrostatic cutoff, after heating the system to 350 K in 7 steps, increasing the temperature by 50 K in each step.

Trajectory analysis. Trajectory analyses were carried out using the *cpptraj* software from the Amber14 package³⁷, visualization was made with VMD³⁸ and UCSF Chimera³⁹.

RMSDs were calculated based on the input topology file to a reference frame of the coordinates both for C α and for the interface residues (iRMSD). Mean RMSD values were calculated for the whole trajectories in each case. Atomic positional fluctuations (RMSF), Connolly surface area⁴⁰, secondary structures⁴¹, distance between the center of mass of monomers were determined. The interface surfaces for the dimers were calculated as the

difference between the sum of the molecular surface areas for the separate chains and the molecular surface of the complex. Contacts between the atoms of the dimers were calculated with a distance cutoff of 7 Å. Binding energy was calculated with the MM-PBSA module of Amber14⁴².

Principal component analysis (PCA)⁴³ was calculated for alpha-carbon atoms on trajectories. The first two principal components were computed and a covariance matrix was generated. The first two eigenvectors of the matrix correspond to PC1 and PC2. PCA values were linked to every atom and visualized with the NMWiz plugin in VMD⁴⁴ where relative motions of the atoms are shown as arrows proportional with their values.

Data availability

The datasets used and/or analysed during this study are available from the corresponding author on reasonable request.

Received: 18 March 2022; Accepted: 3 August 2022

Published online: 22 August 2022

References

- Blöchliger, N., Xu, M. & Caffisch, A. Peptide binding to a pdz domain by electrostatic steering via nonnative salt bridges. *Biophys. J.* **108**, 38. <https://doi.org/10.1016/j.bpj.2015.03.038> (2015).
- Saglam, A. S., Wang, D. W., Zwier, M. C. & Chong, L. T. Flexibility vs preorganization: Direct comparison of binding kinetics for a disordered peptide and its exact preorganized analogues. *J. Phys. Chem. B* **121**(43), 10046–10054 (2017).
- Plattner, N., Doerr, S., Fabritiis, G. & Noé, F. Complete protein–protein association kinetics in atomic detail revealed by molecular dynamics simulations and Markov modelling. *Nat. Chem.* **9**, 2785. <https://doi.org/10.1038/nchem.2785> (2017).
- Paul, F. *et al.* Protein-peptide association kinetics beyond the seconds timescale from atomistic simulations. *Nat. Commun.* **8**, 1095 (2017).
- Pan, A. C. *et al.* Atomic-level characterization of protein–protein association. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4244–4249 (2019).
- Schreiber, G. Kinetic studies of protein–protein interactions. *Curr. Opin. Struct. Biol.* **12**, 41. [https://doi.org/10.1016/S0959-440X\(02\)00287-7](https://doi.org/10.1016/S0959-440X(02)00287-7) (2002).
- Lavelle, L., Michel, J. P. & Gingery, M. The disassembly, reassembly and stability of CCMV protein capsids. *J. Virol. Methods* **146**, 311–316. <https://doi.org/10.1016/j.jviromet.2007.07.020> (2007).
- Díaz-Valle, A., García-Salcedo, Y. M., Chávez-Calvillo, G., Silva-Rosales, L. & Carrillo-Tripp, M. Highly efficient strategy for the heterologous expression and purification of soluble cowpea chlorotic mottle virus capsid protein and in vitro pH-dependent assembly of virus-like particles. *J. Virol. Methods* **225**, 23–29 (2015).
- Hassani-Mehraban, A., Creutzburg, S., van Heereveld, L. & Kormelink, R. Feasibility of cowpea chlorotic mottle virus-like particles as scaffold for epitope presentations. *BMC Biotechnol.* **15**, 80. <https://doi.org/10.1186/s12896-015-0180-6> (2015).
- Law-Hine, D. *et al.* Reconstruction of the disassembly pathway of an icosahedral viral capsid and shape determination of two successive intermediates. *J. Phys. Chem. Lett.* **6**, 3471–3476. <https://doi.org/10.1021/acs.jpcclett.5b01478> (2015).
- Tresset, G. *et al.* Two-dimensional phase transition of viral capsid gives insights into subunit interactions. *Phys. Rev. Appl.* **7**, 014005 (2017).
- Bereau, T., Globisch, C., Deserno, M. & Peter, C. Coarse-grained and atomistic simulations of the salt-stable cowpea chlorotic mottle virus (ss-ccmv) subunit 26–49: β -barrel stability of the hexamer and pentamer geometries. *J. Chem. Theory Comput.* **8**, 3750–3758 (2012).
- Kortemme, T. & Baker, D. A simple physical model for binding energy hot spots in protein–protein complexes. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14116–21. <https://doi.org/10.1073/pnas.202485799> (2002).
- Davis, F. P. & Sali, A. PIBASE: A comprehensive database of structurally defined protein interfaces. *Bioinformatics* **21**, 1901–1907. <https://doi.org/10.1093/bioinformatics/bti277> (2005).
- Zhang, L., Borthakur, S. & Buck, M. Dissociation of a dynamic protein complex studied by all-atom molecular simulations. *Biophys. J.* **110**, 877–886. <https://doi.org/10.1016/j.bpj.2015.12.036> (2016).
- Zhang, L. & Buck, M. Molecular simulations of a dynamic protein complex: Role of Salt-bridges and polar interactions in configurational transitions. *Biophys. J.* **105**, 2412–2417. <https://doi.org/10.1016/j.bpj.2013.09.052> (2013).
- Banerjee, P. & Bagchi, B. Dynamical control by water at a molecular level in protein dimer association and dissociation. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 2302–2308. <https://doi.org/10.1073/pnas.1908379117> (2020).
- Speir, J. A., Munshi, S., Wang, G., Baker, T. S. & Johnson, J. E. Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure* **3**, 63–78. [https://doi.org/10.1016/S0969-2126\(01\)00135-6](https://doi.org/10.1016/S0969-2126(01)00135-6) (1995).
- Chen, J., Lansac, Y. & Tresset, G. Interactions between the molecular components of the cowpea chlorotic mottle virus investigated by molecular dynamics simulations. *J. Phys. Chem. B* **122**, 9490–9498. <https://doi.org/10.1021/acs.jpcc.8b08026> (2018).
- Bancroft, J. B., Rees, M. W., Johnson, M. W. & Dawson, J. R. O. A salt-stable mutant of cowpea chlorotic mottle virus. *J. Gen. Virol.* **21**, 507–513. <https://doi.org/10.1099/0022-1317-21-3-507> (1973).
- Fox, J. M., Zhao, X., Speir, J. A. & Young, M. J. Analysis of a salt stable mutant of cowpea chlorotic mottle virus. *Virology* **222**, 115–122 (1996).
- Vangone, A., Spinelli, R., Scarano, V., Cavallo, L. & Oliva, R. COCOMAPS: A web application to analyze and visualize contacts at the interface of biomolecular complexes. *Bioinformatics* **27**, 2915–2916. <https://doi.org/10.1093/bioinformatics/btr484> (2011).
- Salamanca Vilorio, J., Allega, M. F., Lambrugh, M. & Papaleo, E. An optimal distance cutoff for contact-based protein structure networks using side-chain centers of mass. *Sci. Rep.* **7**, 6. <https://doi.org/10.1038/s41598-017-01498-6> (2017).
- Speir, J. A. *et al.* Enhanced local symmetry interactions globally stabilize a mutant virus capsid that maintains infectivity and capsid dynamics. *J. Virol.* **80**, 3582–3591. <https://doi.org/10.1128/JVI.80.7.3582-3591.2006> (2006).
- Antal, Z., Szoverfi, J. & Fejer, S. Predicting the initial steps of salt-stable cowpea chlorotic mottle virus capsid assembly with atomistic force fields. *J. Chem. Inf. Model.* **57**, 78. <https://doi.org/10.1021/acs.jcim.7b00078> (2017).
- Globisch, C., Krishnamani, V., Deserno, M. & Peter, C. Optimization of an elastic network augmented coarse grained model to study CCMV capsid deformation. *PLoS ONE* **8**, e60582. <https://doi.org/10.1371/journal.pone.0060582> (2013).
- Garmann, R. F., Comas-García, M., Gopal, A., Knobler, C. M. & Gelbart, W. M. The assembly pathway of an icosahedral single-stranded rna virus depends on the strength of inter-subunit attractions. *J. Mol. Biol.* **426**, 1050–1060. <https://doi.org/10.1016/j.jmb.2013.10.017> (2014).
- Chakravarty, A., Reddy, V. S. & Rao, A. Unravelling the stability and capsid dynamics of the three virions of bromo mosaic virus assembled autonomously in vivo. *J. Virol.* **94**, e01794 (2020).
- Anandakrishnan, R., Drozdetski, A., Walker, R. C. & Onufriev, A. V. Speed of conformational change: Comparing explicit and implicit solvent molecular dynamics simulations. *Biophys. J.* **108**, 1153–1164 (2015).

30. Zlotnick, A., Aldrich, R., Johnson, J. M., Ceres, P. & Young, M. J. Mechanism of capsid assembly for an icosahedral plant virus. *Virology* **277**, 450–456. <https://doi.org/10.1006/viro.2000.0619> (2000).
31. Schrödinger, LLC. *PyMOL—The PyMOL Molecular Graphics System, Version 1.8* (Schrödinger, LLC, 2015).
32. Schafmeister, C., Ross, W. & Romanovski, V. *Leap* (University of California, San Francisco, 1995).
33. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926. <https://doi.org/10.1063/1.445869> (1983).
34. Maffucci, I. & Contini, A. Improved computation of protein–protein relative binding energies with the Nwat-MMGBSA method. *J. Chem. Inf. Model.* **56**, 1692–1704. <https://doi.org/10.1021/acs.jcim.6b00196> (2016).
35. Duan, Y. *et al.* A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**, 1999–2012. <https://doi.org/10.1002/jcc.10349> (2003).
36. Onufriev, A., Bashford, D. & Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins Struct. Funct. Bioinform.* **55**, 383–394. <https://doi.org/10.1002/prot.20033> (2004).
37. Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095. <https://doi.org/10.1021/ct400341p> (2013).
38. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) (1996).
39. Pettersen, E. F. *et al.* UCSF chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612. <https://doi.org/10.1002/jcc.20084> (2004).
40. Connolly, M. L. Analytical molecular surface calculation. *J. Appl. Crystallogr.* **16**, 548–558. <https://doi.org/10.1107/S0021889883010985> (1983).
41. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen bonded and geometrical features. *Biopolymers* **22**, 2577–2637. <https://doi.org/10.1002/bip.360221211> (1983).
42. Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **10**, 449. <https://doi.org/10.1517/17460441.2015.1032936> (2015).
43. David, C. C. & Jacobs, D. J. Principal component analysis: A method for determining the essential dynamics of proteins. *Methods Mol. Biol. (Clifton, N.J.)* **1084**, 193–226. https://doi.org/10.1007/978-1-62703-658-0_11 (2014).
44. Bakan, A., Meireles, L. M. & Bahar, I. ProDy: Protein dynamics inferred from theory and experiments. *Bioinformatics* **27**, 1575–1577. <https://doi.org/10.1093/bioinformatics/btr168> (2011).

Acknowledgements

This work was supported by a grant of the Romanian Ministry of Research and Innovation, CNCS-UEFISCDI, Project Number PN-III-P1-1.1-TE-2019-1018, within PNCDI III.

Author contributions

S.N.F. and J.S. designed the simulations, J.S. conducted the simulations, J.S. and S.N.F. analysed the results. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-18019-9>.

Correspondence and requests for materials should be addressed to S.N.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022