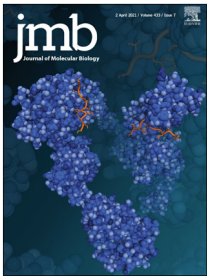

The Trimeric Major Capsid Protein of Mavirus is stabilized by its Interlocked N-termini Enabling Core Flexibility for Capsid Assembly

Alexander Christiansen [1], Marie Weiel [2], Andreas Winkler [3], Alexander Schug [4] and Jochen Reinstein [1]*

*1 - **Max Planck Institute for Medical Research,** Department of Biomolecular Mechanismsm Heidelberg, Germany*
*2 - **Karlsruhe Institute of Technology,** Steinbuch Centre for Computing and Department of Physics, Eggenstein-Leopoldshafen, Germany*
*3 - **Institute of Biochemistry,** Graz University of Technology. Graz, Austria*
*4 - **Institute for Advanced Simulation,** Jülich Supercomputing Center, Jülich, Germany*

***Correspondence to Jochen Reinstein:** jochen.reinstein@mpimf-heidelberg.mpg.de (J. Reinstein)*
https://doi.org/10.1016/j.jmb.2021.166859
Edited by Owen Pornillos

Abstract

Icosahedral viral capsids assemble with high fidelity from a large number of identical buildings blocks. The mechanisms that enable individual capsid proteins to form stable oligomeric units (capsomers) while affording structural adaptability required for further assembly into capsids are mostly unknown. Understanding these mechanisms requires knowledge of the capsomers' dynamics, especially for viruses where no additional helper proteins are needed during capsid assembly like for the Mavirus virophage that despite its complexity (triangulation number T = 27) can assemble from its major capsid protein (MCP) alone. This protein forms the basic building block of the capsid namely a trimer (MCP_3) of double-jelly roll protomers with highly intertwined N-terminal arms of each protomer wrapping around the other two at the base of the capsomer, secured by a clasp that is formed by part of the C-terminus. Probing the dynamics of the capsomer with HDX mass spectrometry we observed differences in conformational flexibility between functional elements of the MCP trimer. While the N-terminal arm and clasp regions show above average deuterium incorporation, the two jelly-roll units in each protomer also differ in their structural plasticity, which might be needed for efficient assembly. Assessing the role of the N-terminal arm in maintaining capsomer stability showed that its detachment is required for capsomer dissociation, constituting a barrier towards capsomer monomerisation. Surprisingly, capsomer dissociation was irreversible since it was followed by a global structural rearrangement of the protomers as indicated by computational studies showing a rearrangement of the N-terminus blocking part of the capsomer forming interface.

Introduction

Viral capsids with their ability to self-assemble reliably and with high fidelity are fascinating biological entities. In the case of icosahedral capsids, their faces and vertices are occupied by protein units in hexa-coordinate (hexons) and penta-coordinate (pentons) geometries respectively.[1,2] Despite their common architecture and organizational principles, their respective assembly paths differ.[3–5] The reaction path in each case presumably represents a thermodynamic compromise between enabling reliable assembly without too many kinetic traps and sufficient capsid stability to withstand the environment long enough to allow entry into a new host.[6–8] To meet this

challenge, some viruses employ additional tape measure proteins to guide and facilitate their assembly, while others use DNA, RNA, or lipids as scaffolds.[9–14] For viruses that do not employ additional proteins, the information for correct assembly must be contained in the hexons and pentons or their constituents. Knowledge of the energetics and dynamics of hexons and pentons is therefore crucial for insight into viral capsid assembly and could, in turn, be used to manipulate capsids and their assembly to develop tailor-made nanocontainers or, by identifying hotspots, to influence the viral life cycle for potential therapeutic purposes.[15,16]

The capsid of the marine Mavirus virophage is an ideal test system. Despite its size and complexity with a triangulation number of $T = 27$, its capsid assembly does not depend on any additional cofactors. The hexons alone are capable of forming stable capsid-like structures without pentons being present,[17] which makes their internal dynamics particularly important. The crystal structure of the major capsid protein (MCP) revealed a trimeric hexon (capsomer) consisting of three double jelly roll (DJR) proteins (protomers).[17] An interesting feature of the capsomer are the N-terminal tails forming extended strands. These strands wrap around their neighboring DJRs in the capsomer and interact with parts of the neighboring subunit's C-terminal domain (Figure 1). The whole arrangement is suggestive of a brace holding together the capsomer supported by clasps formed by parts of the C-terminal domain.[17] A similar feature has been described in other members of the adenoviral family, e.g., adenovirus type 2[18] or PRD1,[19] and it has been suggested as a general way to stabilize DJR capsomers.[20] However, no further biophysical backup for this proposed mechanism exists beyond this structural evidence.

Here, we address the crucial question about the interplay of structural dynamics and stability of the major capsid protein from mavirus by identifying flexible regions with H/D exchange coupled to mass spectrometry (HDX-MS) in conjunction with kinetic approaches and *in silico* simulations. We suggest that the Mavirus MCP capsomer (MCP_3) is stabilized in an assembly-competent conformation by a brace formed by entanglement of N-terminal residues that determines capsomer stability, but simultaneously allow for relatively high dynamics of the capsomers structure, a feature that is increasingly recognized to be crucial for successful capsid assembly.[21] These findings extend our understanding of the main building blocks not only of virophages but also of other related viruses. Similar N-terminal interactions are observed in other viral hexon structures, and our findings help in understanding viral stability and assembly in general.

Results

Dynamics of capsomer structure - high structural plasticity of the double jelly rolls

Using HDX coupled to mass spectrometry, we sought to detect accessible and mobile elements of MCP_3 to identify regions potentially relevant for capsomer stability and capsid assembly. H/D exchange rates in amide bonds were mapped for

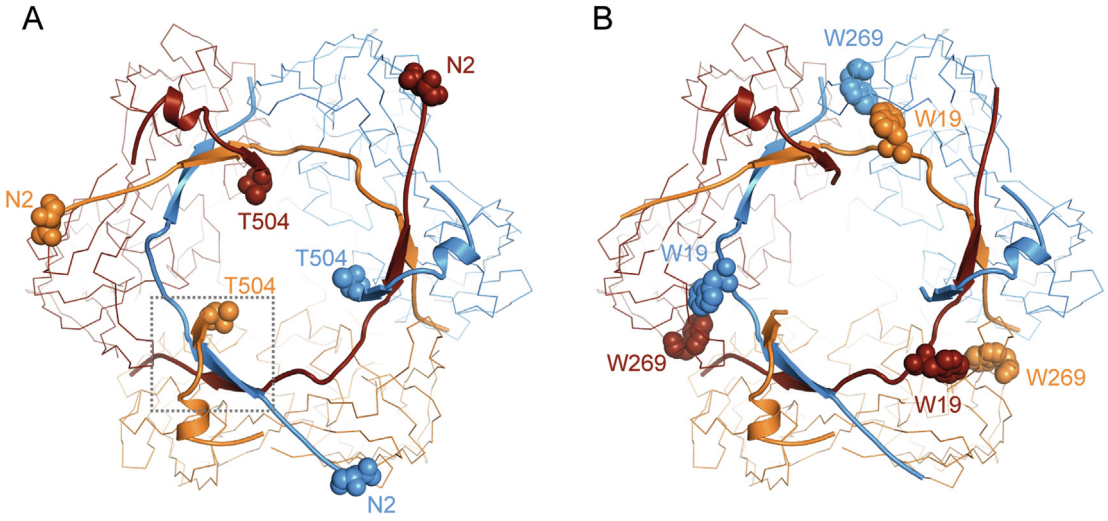

Figure 1. Structures and interactions at the basal interprotomer beta sheet of Mavirus MCP_3. Illustrated MCP_3 structure (PDB 6G45). MCP protomers are colored individually throughout the image (chain A orange, chain B blue, chain C brown). (A, B) The MCP capsomer is illustrated with the double jelly roll (DJR) core as ribbon and N- and C-terminal arms (residues 2 to 36 and 489 to 504, respectively) as cartoon. Spheres illustrate the N- and C-terminally resolved residues in the crystal structure in (A) and interacting tryptophan residues (W19 and W269) in (B). W19 is part of the N-terminal arm and in close proximity to W269 (positioned in DJR strand 1) from a neighboring protomer.

individual MCP peptide fragments. Above average conformational dynamics were observed for residues 1–36 and 495–505, comprising the N-terminal arm and the clasp-forming C-terminal part, respectively (Figures 1 and 2). Additionally, the C-terminal residues 505 – 606, not resolved in the crystal structure, exhibited very high exchange rates, suggesting that they are indeed unstructured as indicated from the missing density (Figure 2).[17]

As far as the jelly roll structures of the capsomer are concerned, the first DJR (DJR1) featured rather slow deuterium incorporation kinetics whereas the second DJR (DJR2, residues 300–

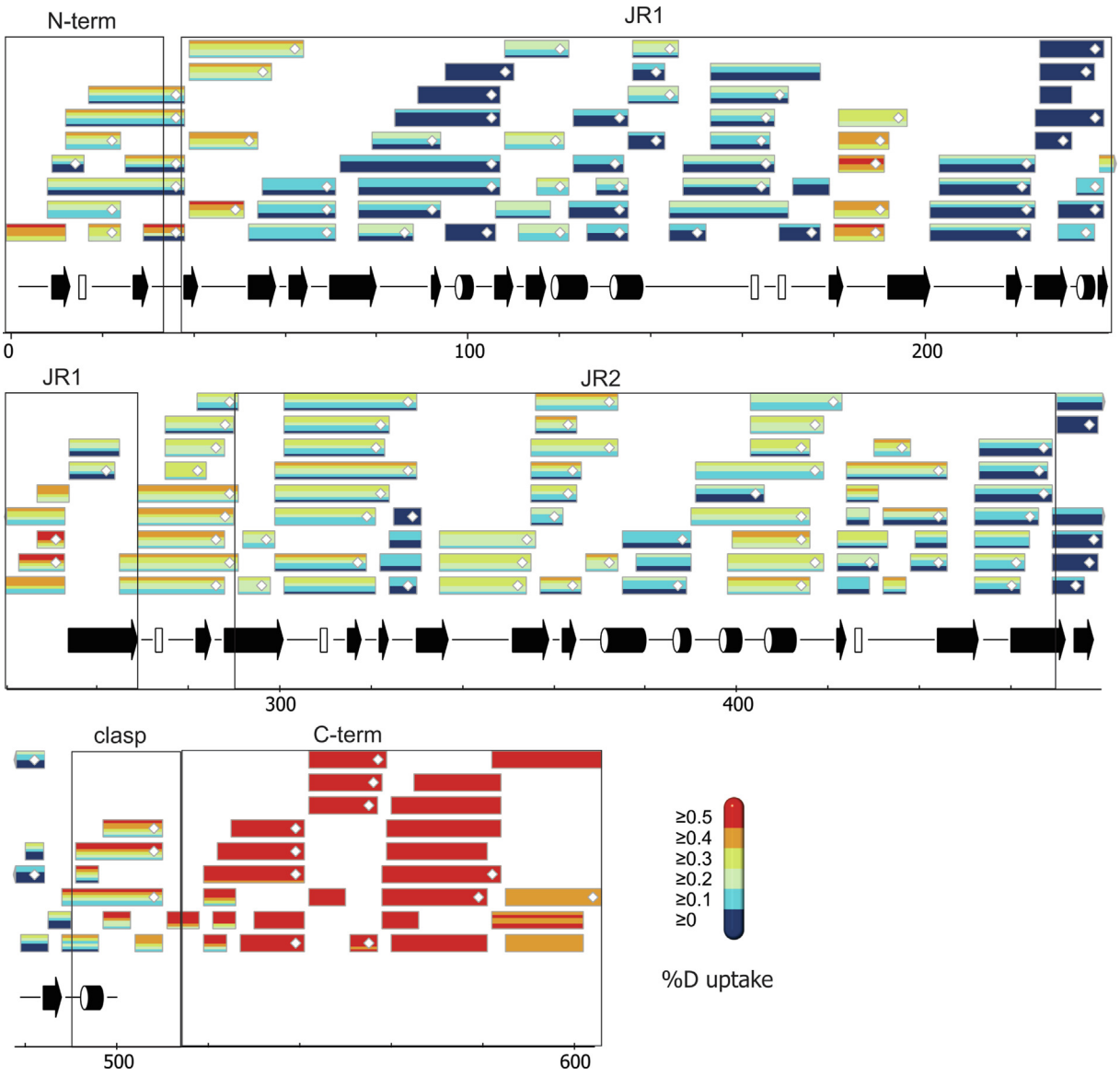

Figure 2. Overview of the HDX-MS analysis of full-length MCP. Individual peptides are represented as boxes with up to six different colors that, from bottom up, represent the deuteration times of 15 s, 45 s, 3 min, 15 min, 1 h, and 4 h with individual colors corresponding to the relative deuterium uptake normalized to the total number of exchangeable amide positions per peptide (see legend). Based on the fully exchanging C-terminal peptides, a back-exchange of approx. 40 % is estimated and, considering that the first two amide positions back-exchange quickly, the scale was adjusted to represent the accessible deuterium uptake range. The heat map consequently shows the secondary structure stability of individual peptides with red and blue indicating lowest and highest stability, respectively. MS^2 confirmed peptides are marked with diamonds. Terminal arrows at the end of a box indicate continuation of the peptide in the previous or following line. Secondary structure elements derive from DSSP analysis of the crystal structure (PDB 6G45, chain A). Numbering corresponds to the wt MCP (Uniprot A0A1L4BK98). Individual peptide deuteration plots can be found in Figure S6.

590) of each protomer showed relatively high conformational dynamics comparable to the N-terminal arm, especially in the CHEF beta sheet of DJR2 (Figures 2, 3 and S6). This is reminiscent of the observed plasticity in the CHEF beta sheet of DJR2 in the structure of the native Sputnik virophage. Here it was noted that the MCP structure has greater variability with displacements of the Cα atoms as much as 3.5 Å in specific regions of the second jelly roll that make contacts with neighboring capsomers.[22] The locations of these highly dynamic elements in the Mavirus Capsomer structure are indicated in Figure 3 which shows a heat map projection of the HDX exchange rates onto the structure. The majority of "hot" regions of the two jelly rolls lie on the outside surface that is involved in inter capsomer interactions in the capsid assembly, but also involves central parts of the DJRs. The correlation in dynamics of parts of the jelly roll structures within each protomer of the capsomer and that of the N-terminal part raises the question about the particular role of this structural element.

N-terminal arm detachment preceded unfolding

The crystal structure of the MCP capsomer (MCP_3) from Mavirus[17] is symmetric with intertwined N- and C-termini forming three interprotomer beta sheets at the capsomer's base. Each N-terminus (residues 1 to 36) spans both neighboring protomers, while the C-terminal residues 489 to 504 form a clasp around the N-termini from both neighboring protomers (Figure 1(A)). A tryptophan (Trp) residue in each N-terminus is located close to a second tryptophan residue in the adjacent protomer and part of the core structure (Figure 1(B)).

To probe capsomer stability, we recorded fluorescence emission spectra of MCP_3 in the

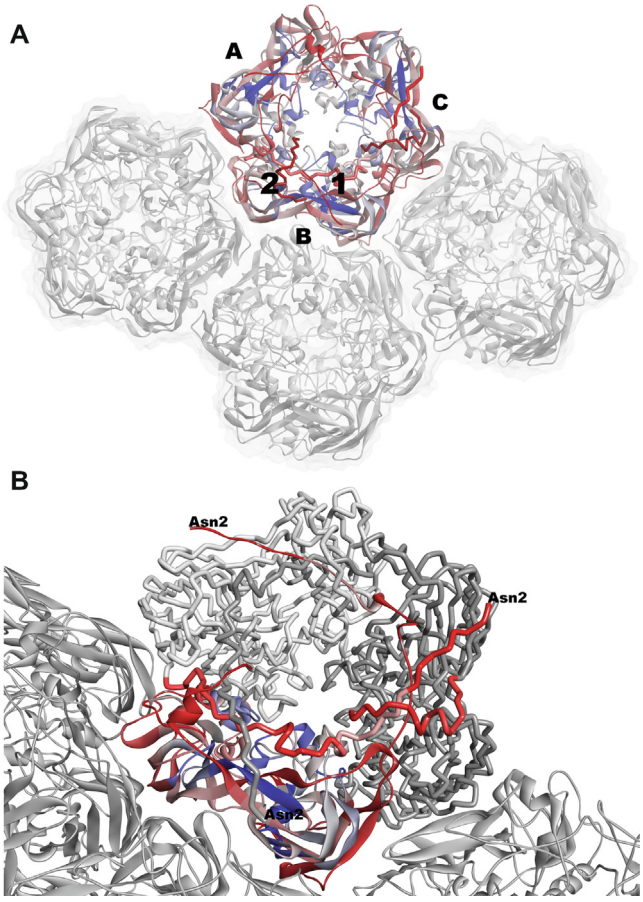

Figure 3. HDX-MS data at 240 min projected onto MCP3 structure as heat map. A: Mavirus MCP_3 structure[17] in the context of the capsid structure of Sputnik virophage.[22] Sputnik Capsomers are colored light grey and a Mavirus capsomer is shown as ribbon structure colored according to HDX exchange data at 240 min with blue-white-red scaling indicating increasing deuterium incorporation. The 3 protomers of the trimer are labelled A,B and C. The the two jelly-roll structures of protomer B are labelelled 1 and 2. The view is from inside the capsid. B: View from inside of capsid showing interface forming regions and dynamics of the two jelly-roll structures of protomer B colored according to the HDX heat map. Additionally, the N-terminal arm of protomer A is displayed as tube and colored according to HDX data as are the two clasp regions of protomers B and C. The N-termini of all three units are labelled with Asn2.

absence and presence of the chaotropic denaturant guanidinium chloride (GdmCl). Upon excitation of Trp residues at 297 nm, fluorescence emission spectra showed a nearly 3-fold increase and slight red shift in fluorescence intensity as the GdmCl concentration was raised to 1.0 M, and a decrease and further red shift on increasing further to 3.0 M GdmCl (Figure 4(A)). These data indicate two distinct transitions: a steep fluorescence increase in the range of 0.5 M to 1.0 M GdmCl, a plateau phase up to 1.5 M, and a decrease to 3.0 M GdmCl (Figure 4(B)). The second of these transitions originates from global protein unfolding, which leads to exposure of aromatic residues to a hydrophilic environment and causes a decrease in fluorescence intensity accompanied by a red shift. Still, even at 3.0 M GdmCl, the fluorescence intensity was higher than at 0.0 M, perhaps due to a substantial quenching of Trp residues in the native state. This initial de-quenching of Trp up to 1.0 M GdmCl is unusual for global protein unfolding and could reflect a different underlying process. Each MCP protomer contains only two Trp residues: at position 19 in the N-terminal arm and at position 269 in the core of its first double jelly roll (DJR1). W269 is spatially close (6 to 7 Å) to W19 of a neighboring protomer (Figure 1(B)). Close proximity of Trp to other aromatic residues is a source of unusual fluorescence behavior.[23,24] If this is the case for the Trp-Trp interaction here, it can be used to track the movement of the N-terminal arm relative to the capsomer core.

To validate the aromatic interaction between W19 and W269 as the cause of fluorescence de-quenching observed during the first transition, we introduced either of the point mutations W19I and W269I. Upon addition of GdmCl, the first transition phase observed for wildtype (wt) MCP_3 was lost in the Trp fluorescence signal for both MCP_3 W19I

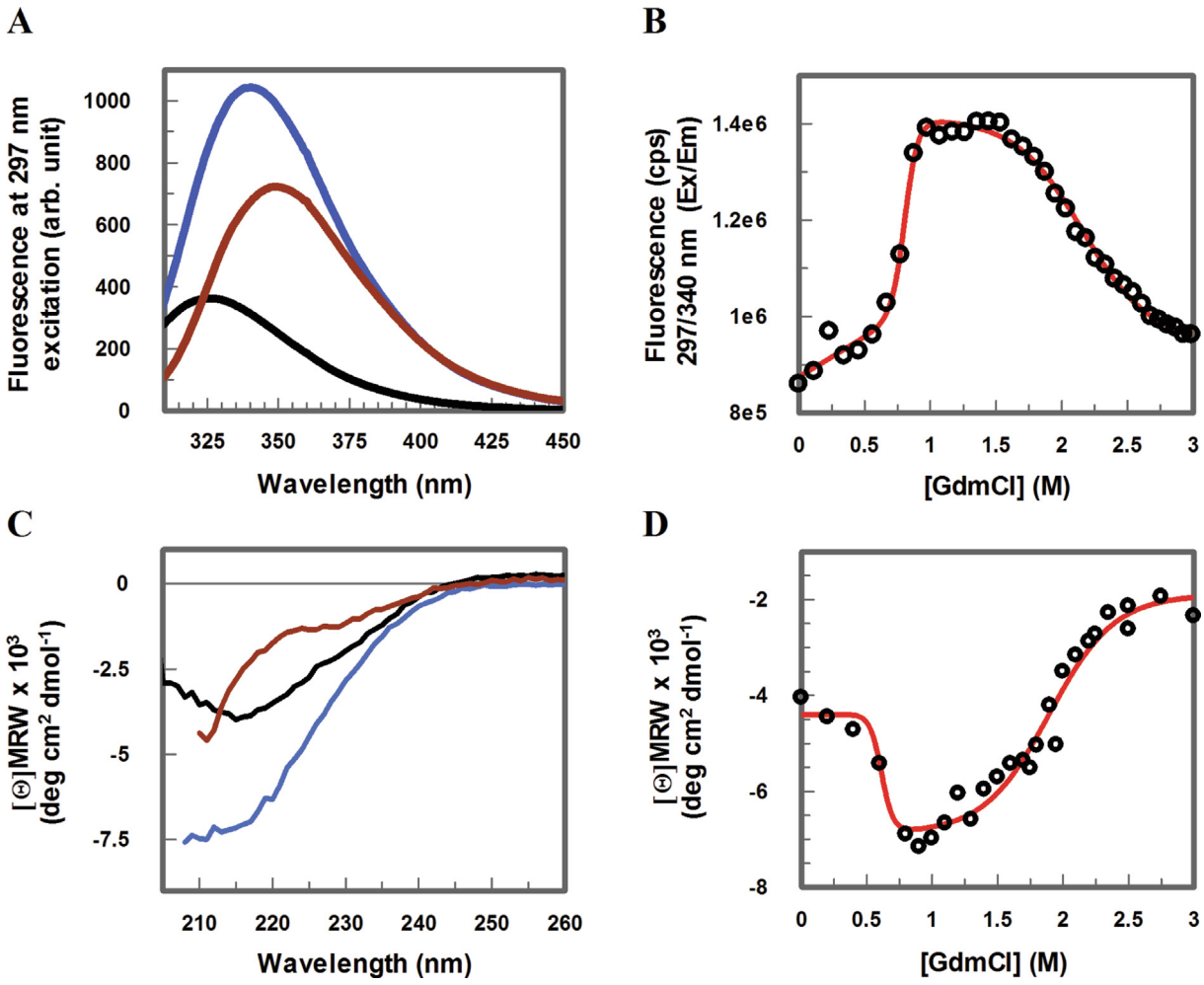

Figure 4. Capsomer structure and stability. CD and fluorescence spectra of MCP_3 at increasing GdmCl concentration [GdmCl] and GdmCl titration curves at fixed wavelengths. Equilibrium spectra at 0 M (black), 1 M (blue), and 3 M GdmCl (red) of 2.5 μM MCP measured with (A) fluorescence (297 nm excitation) and (C) CD. Chemical denaturation followed at fixed wavelength: (B) fluorescence at 340 nm for tryptophan fluorescence (297 nm excitation) and (D) CD at 217 nm. The red lines in (B) and (D) correspond to three-state fits to the experimental data.

and MCP_3 W269I. Only the second transition, with decreasing Trp fluorescence, was observed (Figure S1(B)). With increasing GdmCl concentration, fluorescence spectra of both MCP_3 W19I and MCP_3 W269I showed a red shift and a decrease in the emission peak reminiscent of protein unfolding (Figure S1(A)), comparable to the second transition for wt. Furthermore, near-UV circular dichroism (CD) measurements revealed a signal around 270 nm for wt MCP_3 indicating aromatic side chain interactions, which was absent for MCP_3 W19I and MCP_3 W269I (Figure S1(D)). For wt MCP_3, this signal disappeared at 1.0 M GdmCl (Figure S1(D)), coinciding with the first unfolding transition in the fluorescence data (Figure 4(B)). Thus, the dequenching observed during the first fluorescence transition upon addition of GdmCl was caused by Trp-Trp quenching between the N-terminal arm and the DJR core. It serves as an intrinsic reporter for structural changes affecting the relative locations of the N-terminal arms relative to the capsomer core.

Next, we probed the global structural integrity of MCP_3 using CD spectroscopy. The spectra of native MCP_3 displayed a shape consistent with the beta-strand rich DJR structure of the capsomer. Upon addition of GdmCl from 0.0 M to 1.0 M, we found an unexpected increase in CD intensity accompanied by a shift in the maximum wavelength, pointing to a structural rearrangement of the protein. Spectra at high GdmCl concentration exhibited a decrease in the CD signal and a further change in line with an unfolded random coil (Figure 4(C)). A denaturation curve measured at 217 nm displayed two transitions with inflection points comparable to those observed using fluorescence (Figure 4(D)). Far-UV CD spectra of the Trp-mutant proteins showed the same features as the wt, i.e., an increase in CD intensity from 0.0 M to 1.0 M GdmCl followed by a decrease towards 3.0 M GdmCl (Figure S1(C)). This indicates that the mutations do not affect the overall fold and unfolding mechanism of MCP_3.

Combining the results from fluorescence dequenching and CD spectroscopy, we conclude that MCP_3 exhibits a three-state unfolding mechanism, with the second transition resulting from global protein unfolding. The first transition probably reports on a conformational change which we assign to the separation of the two Trp residues in neighboring subunits as a consequence of the N-terminal arm moving away from the capsomer core.

Influence of intertwined N-termini and C-terminal clasp on capsomer dissociation

To assess whether the intertwined termini at the capsomer base stabilize the trimer, we investigated how GdmCl affects the capsomer's oligomeric state. From earlier experiments, purification following MCP expression in E. coli yield partly MCP_3 trimer and partly assembled capsid but not as the MCP monomer.[17] To permit FRET within MCP_3, a cysteine mutant (MCP D277C) was double-labeled with a mixture of Atto donor and acceptor dyes (Atto488 and Atto594, the location of D277 is indicated in Figure S2(A)). As expected, the emission spectrum of this molecule excited at the donor wavelength of 488 nm showed an additional FRET peak at the acceptor emission, thus confirming successful double-labeling. Addition of 1.0 M GdmCl resulted in a decrease in FRET as reflected in increasing donor and decreasing acceptor intensities (Figure S2 (B)). Strikingly, the FRET signal obtained as a function of GdmCl concentration revealed a single transition matching the first unfolding transition observed in CD and fluorescence spectroscopy (Figure S2(C)). This suggests that MCP_3 monomerizes upon detachment of the termini from the capsomer base. Time-resolved FRET measurements of donor/acceptor labeled MCP D277C showed a faster decrease of the FRET signal with increasing GdmCl concentration (Figure 5(A)). The resulting kinetic traces could be fitted with single exponentials. A replot of the derived rate constants revealed an initial increase up to 1.0 M GdmCl, after which the rate constants started to level off (Figure 5 (C)). The theoretical rate constant of disassembly under non-denaturing conditions was derived from linear extrapolation between 0.5 and 1.0 M GdmCl and was about $5{\cdot}10^{-7}$ s^{-1}, revealing the trimeric capsomers to be highly stable when the basal structure of N- and C-termini is intact.

To elucidate the sequence of events during capsomer disassembly as monitored by disappearance of fluorescence energy transfer FRET and separation of N-terminal arms from the core as shown by Trp fluorescence, we conducted stopped-flow measurements. Trp fluorescence displayed a fast phase accounting for the bulk of the amplitude, a minor second phase with a fluorescence increase up to 1.0 M GdmCl, and two clear phases at higher GdmCl concentrations (Figure 5(B)). Replotting the rate constants derived from exponential fits showed two well-separated processes. The first phase, associated with the intensity increase, stayed constant between 0.5 M and 1.0 M GdmCl (Figure 5(C), black dots) before increasing with GdmCl concentration. The second phase (Figure 5(C), black squares) started at 0.5 M and became more pronounced at higher concentrations. Rate constants at concentrations lower than 1.2 M were very hard to determine due to the low signal amplitude. Conversely, both W19I and W269I only exhibited a single phase at 297 nm excitation, coinciding with the slow phase of the wt protein (Figures S3(A) and (B)). Thus, the slow

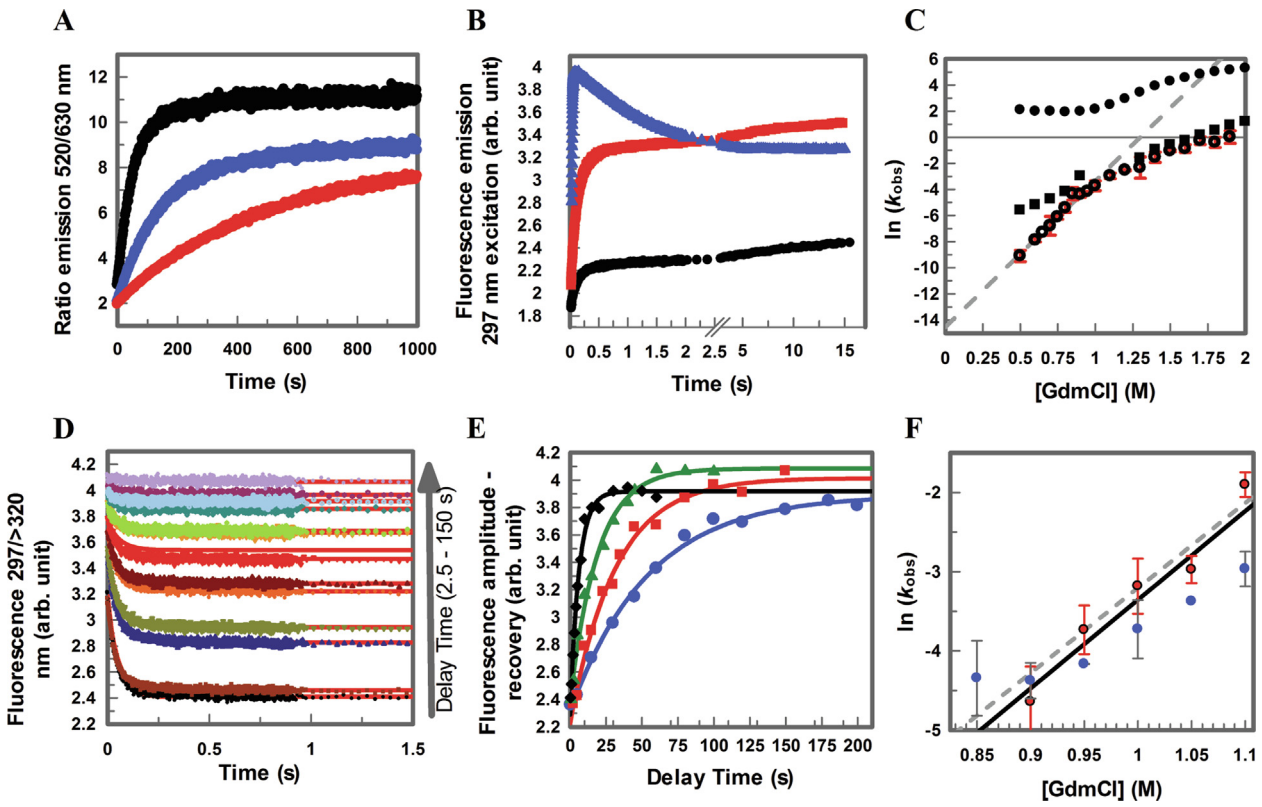

Figure 5. FRET, stopped-flow single and double-jump kinetic measurements for MCP_3. (A) Typical FRET kinetic traces upon a jump to 0.75, 0.85, and 0.95 M GdmCl from pure buffer (red, blue, and black circles). The FRET signal is indicated by the ratio of emission at 520 nm (donor) to 630 nm (acceptor). (B) Unfolding traces upon tryptophan excitation at 297 nm induced by jumps from 0 M GdmCl to 0.8, 1.0, and 1.6 M GdmCl, respectively (black, red and blue). Total fluorescence intensity was recorded with a 320 nm long pass filter. (C) Replot of rate constants derived from double-exponential fits to kinetic traces as shown in 5B (black circles: fast phase, black squares: slow phase). Rate constants from mono-exponential fits to FRET kinetics (open circles). Error bars were calculated from 2 to 4 repetitions at each [GdmCl]. (D) Kinetic traces for the second (backward) jump from 1 M to 0.45 M (back jump) after different delay times following an initial jump from 0.45 M GdmCl to 1 M GdmCl (forward jump). The fluorescence (297 nm excitation using a 320 nm long pass filter) was recorded after the second jump. Final protein concentration of MCP was 1 μM. (E) Plot of MCP_3 N-terminal arm irreversibility at varying [GdmCl]. For backward jumps to 0.45 M GdmCl from various [GdmCl], the final fluorescence value after 5 sec was recorded and plotted as a function of delay time. The solid lines represent mono-exponential fits to the resulting traces. Blue circles: 0.45 M to 0.95 M to 0.45 M; red squares: 0.45/1.0/0.45; green triangles: 0.45/1.05/0.45; black diamond: 0.45/1.1/0.45 (F) Replot of the rate constants obtained from (E) for MCP_3 N-terminal irreversibility as a function of [GdmCl]. The broken line is a linear fit to the data points and results in an extrapolated rate constant of 9×10^{-7} s^{-1} at 0 M GdmCl (ln = -13.9). The averages and error bars result from 2 to 3 repetitions. Red circles: rate constants derived from double-jump experiments; blue squares: rate constants derived from FRET experiments; solid black line: linear fit to the FRET dissociation data for comparison from (C).

phase apparently represents the unfolding of the DJR cores.

Regarding the reversibility of Trp de-quenching and capsomer dissociation, we found that refolding was only partially successful after chemical denaturation (Figure S4). Specifically, only the second unfolding transition (global unfolding) could be reversed on a minute timescale. To probe on a faster timescale whether recovery of the first unfolding transition is possible, we used double-jump kinetics measuring Trp fluorescence. The first jump was from 0.45 M GdmCl (on the brink of the first unfolding transition) to a higher concentration as indicated in Figure 5(D). This allowed the fast Trp de-quenching reaction to proceed for a defined delay time *t*1 (0.1 to 120 s), after which a second mixing reversed the GdmCl concentration to 0.45 M. An increase in *t*1 decreased recovery of the initial fluorescence signal after the second jump, i.e. it was increasingly difficult to reverse the de-quenching and thus the separation of the N-terminal arm from the DJR core (Figure 5(D)). Final fluorescence values after the double jump

versus different delay times yielded a trace showing the progressive irreversibility of the N-terminal separation and were fitted with a mono-exponential equation (Figure 5(E)). Repeated measurements with different intermediate GdmCl concentrations during the first mixing step showed that the time required to produce a reassembly-incompetent form was shorter at higher values (Figure 5(F)). Additionally, the rate constants derived from these traces increased linearly with denaturant concentration (Figure 5(F)). While the de-quenching caused by Trp separation could be reversed on a fast timescale, capsomer disassembly as observed by FRET was not reversible. Furthermore, the rate constants for the irreversible movement of the N-terminal arm as observed via the Trp-based signal and those for monomerization from FRET experiments were very similar (Figure 5(F)). Thus, we find that the inability to restore the N-terminal arm to its initial position is due to irreversible dissociation of the trimeric capsomers. The denaturation of individual protomers could be reversed to a certain extent. We therefore conclude that the intricate entanglement of the N-terminal arms and C-terminal clasps across the capsomer base has to be broken before capsomer dissociation can occur.

Due to the irreversible and fast dissociation step especially at higher GdmCl concentation it was not possible to extraxt reliable thermodynamic parameters from equilbirum chemical denaturation measurements for the 1st transition. Therefore we conducted double-jump unfolding-refolding experiments starting with MCP_3 at 0.4 M GdmCl, jumping to 1 M GdmCl for full arm detachment in the first mixing step t_1 and a second jump to return to a lower, final GdmCl concentration with online recording of kinetic traces as a function of t_2 (Figure S7(A)). The combination of unfolding rate constants (Figure 5(C)) and refolding rate constants allowed us to derive a chevron plot from which we estimated unfolding and refolding rate constants in buffer. Assuming a two-state reaction we could derive equilibrium and rate constants for N-terminal arm opening (Figure S7(B)).

Blockage of the protomer interface by the N-terminus prevents trimer re-association

To test our experimentally derived working model and uncover further structural dependencies and events, we investigated the role of the N-terminal arms in the capsomer structure *in silico* using coarse-grained structure-based models (SBMs). Based on energy landscape theory[25,26] and the principle of minimal frustration, SBMs probe dynamics arising from the system's native geometry. Such models are widely applied in protein folding studies[27,28] and can also be used to interpret experimental results more precisely.[29,30] Capsomer simulations indicate a high relative N-terminal mobility (Figure 6(A) and (B)) and suggest a distinct three-state dissociation mechanism (Figure 6(D)). Starting from the homotrimeric capsomer clamped by three structure-spanning terminal clasps, the N-terminal arms disentangle first while the capsomer core remains relatively stable. Subsequently, the complex separates into three structurally intact monomers before the individual subunits finally unfold completely. Originating from an intact capsomer structure with detached N-termini (Figure 6(E), grey), the arms could be folded back onto the main structure by gradually cooling the system (Figure 6(C)). However, starting from the fully dissociated capsomer, i.e., three separate but spatially close monomers (Figure 6(E), black), the system did not converge to the crystal structure when applying the same simulated annealing protocol.

To further investigate why the monomeric subunits cannot reversely reassemble into the native capsomer, we performed explicit-solvent molecular dynamics (MD) simulations. Simulations started from monomeric MCP with elongated N-termini as extracted from the capsomer crystal structure (Figure S5(F), red). The initially free N-terminus was observed to condense and finally attach to the main structure's DJR core (Figure S5(A) and (F), blue), overlapping with the interprotomer interface and thus interfering with reassembly (Figure S5(D)). After the initial collapse of the N-termini, the structure persistently stayed close to the compacted state (Figure S5(A) and (C)). We used the DSSP (Define Secondary Structure of Proteins) algorithm for estimating hydrogen bonds and assigning secondary structure to the amino acids of the MCP. The analysis predicted that the residues in the N-terminal arm with defined secondary structure leveled at maximum and average fractions of 47.2 % and approximately 20 %, respectively (Figure S5(B) and (E)).

Discussion

Since the Mavirus capsomers expressed in *E. coli*, suffice for assembly of the capsid, the assembly does not require specific co-factors provided by the native host, the giant virus or Mavirus itself.[17] The assembly process is thus entirely encoded in the capsomer properties. Both C- and N-terminal parts are crucial entities, as has been shown for assembly of other virus capsids as well. For phage P22, interactions of the capsomers' N-termini with adjacent capsomers are essential for capsid assembly.[31] Likewise, in the case of murine polyomavirus, the C-termini form contacts with neighboring capsomers to facilitate capsid assembly.[32] The bacteriophage T4 is similar to Mavirus in that the N-terminus is important for monomer-to-hexamer association of the MCP, although it is cleaved off in this process.[33] For Mavirus itself, the C-termini in MCP_3 have been shown to be important in supporting capsid assem-

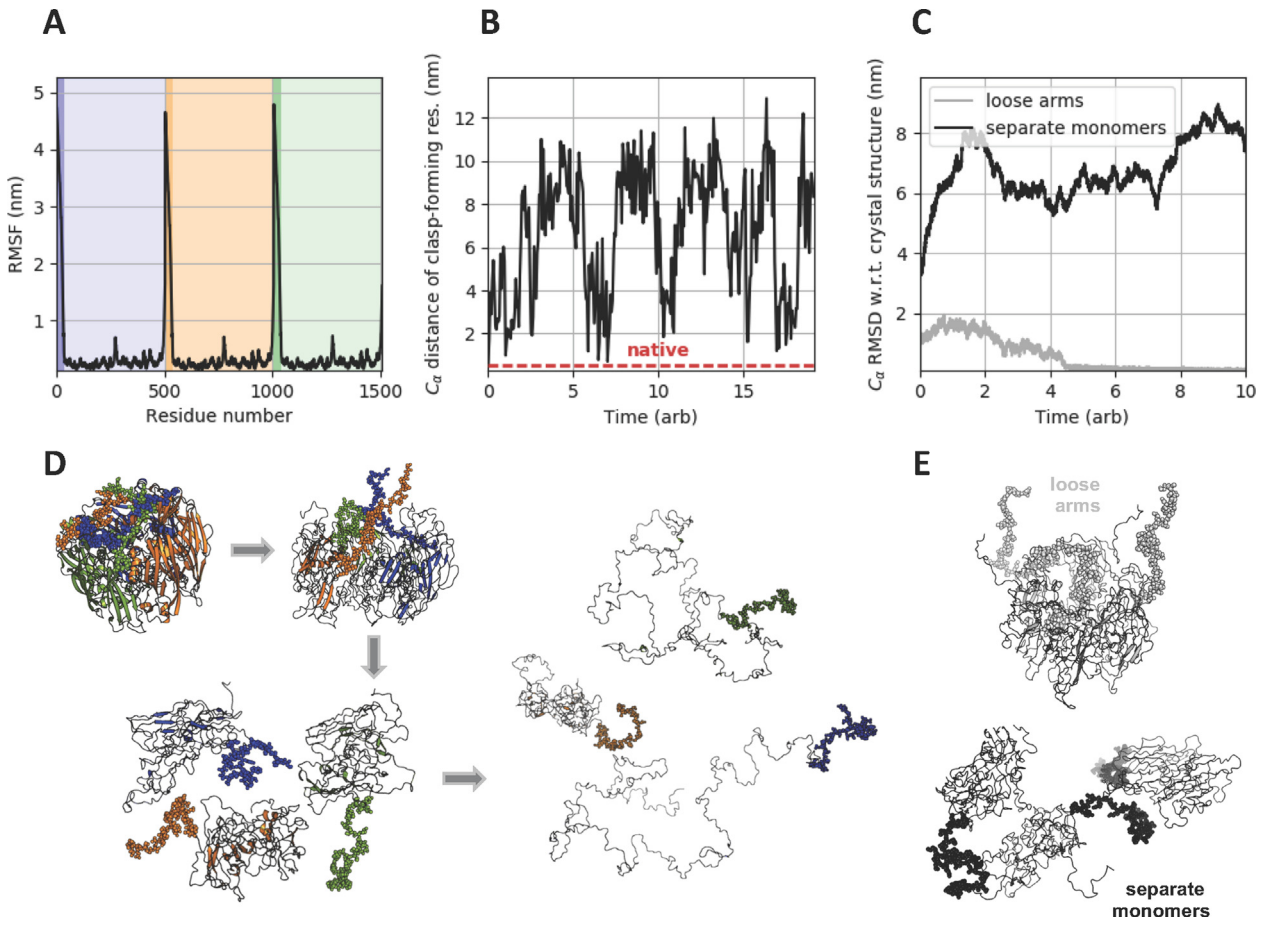

Figure 6. Capsomer structure-based-model (SBM) simulation results. (A) Average root-mean-square fluctuations of atomic positions per residue reveal a relatively high mobility of the N-terminus of MCP. The background is shaded according to the color code of individual monomeric subunits in the structure figure below with N-terminal regions depicted slightly darker. (B) C_α distance of clasp-forming residues as a function of simulated time. The C_α atoms of residues VAL9 and PRO1505 were considered. In the native clasp, these and neighboring residues provide the interaction between the N- and C-terminus of chain A and C, respectively. The calculated temporal development indicates repeated loosening and tightening of the clasp, i.e., opening and closing of the N-terminal arm. The simulation used a reduced GROMACS temperature of 90. (C) Time-dependent C_α RMSD with respect to the crystal structure. The simulations initiated from an intact capsomer structure with loose N-terminal arms (grey) and a fully dissociated state of three separate monomers (black) as displayed in (E) below. As evidenced by the minimizing grey curve, the loose arms could easily be folded back onto the structurally intact trunk. However, originating from separate monomers, the system did not re-converge to the crystal structure as indicated by the diverging black curve. (D) Representative cartoon structures illustrating a model of the MCP_3 dissociation mechanism with N-terminal arms depicted as spheres. Starting from the crystal structure, the N-terminal arms first detach from the stable core. This complex subsequently separates into three structurally intact monomers, before the individual subunits unfold. This simulation used a reduced GROMACS temperature of 106.

bly,[17] whereas the N-termini have not been reported to be directly involved in the assembly process.

Figure 7 summarizes our current view of the inherent dynamics of the MCP capsomer of Mavirus in a schematic model: In its trimeric form, the capsomer is stabilized by about 35 kJ/mol via intertwined N-terminal arms and C-terminal clasps. These two elements, but also parts of the two jelly roll core structures show regions of high dynamics as indicated by heat map coloring. Displacement of the N-terminal arms, as measured by Trp fluorescence de-quenching, allows the capsomer subunits to dissociate. As long as the trimer is intact, the N-terminal arms can revert to their initial state, enabling the clasps to close again (Figure 7 left). However, after dissociation of the protomeric units, reassembly to a trimer becomes impossible (Figure 7 middle). Computational studies suggest that this puzzling behavior results from the re-attachment of the detached N-terminus to the DJR core upon refolding, blocking the trimerization interface, and

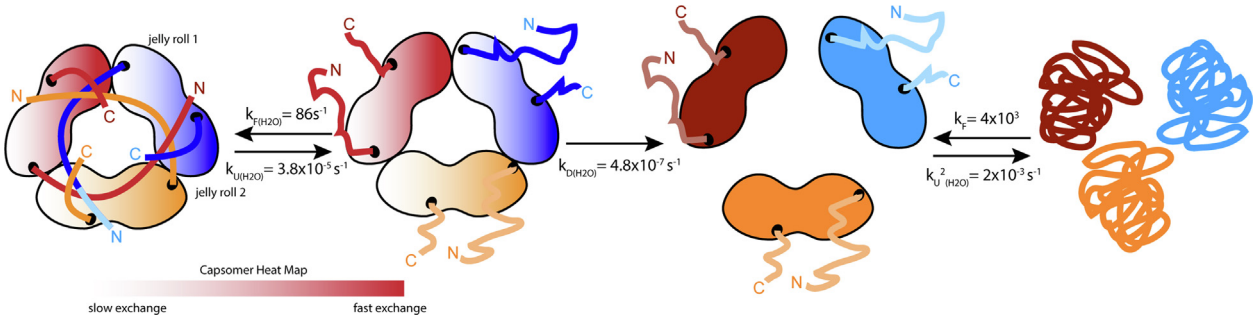

Figure 7. Model of Mavirus capsomer dynamics. Left: In the initial state, the monomers are associated to a trimer with the N-terminal arms forming a brace with neighboring protomers to stabilize the trimer. First reaction: Under denaturing conditions, the N-terminal arms are released while the trimer is still intact. Second reaction: Without the brace the protomers can dissociate in an irreversible step. Third reaction: The single monomers can reversibly fold and unfold, but are not capable of reforming a trimer on their own. Jelly roll 2 in each DJR shows a higher mobility compared to jelly roll 1 within the capsomer as indicated by the heat map. The rate constant for capsomer dissociation ($k_{D(H2O)}$) was determined from linear extrapolation of the FRET dissociation kinetics (Figure 5C). The rate constants for N-terminal arm detachment ($k_{F(H2O)}$ and $k_{U(H2O)}$) were derived from a chevron plot (Figure S7B) and the rate constants for protomer unfolding ($k^2_{U(H2O)}$) from linear extrapolation to 0 M GdmCl of the rate constants for the slower phase for unfolding (Figures 5C and S7B). The equilibrium constant for protomer folding (K_F) was calculated from chemical denaturation data for the second transition for CD and fluorescence (Figure 4B,D).

thus preventing re-association. Additionally, according to our CD spectroscopy data, the monomer's global structure changes as well, possibly also preventing re-association to a trimer and thus capsid assembly.

From a mechanistic perspective, the Mavirus capsomer stabilization shows some resemblance to domain swapping.[34,35] The closed monomeric state as observed in the MD simulation would have the exchanging N-terminal arm closely attached to the DJR core. It however remains unclear whether the approximately 36 N-terminal residues constitute a separate domain or folding unit. Neither in the crystal structure[17] nor in all-atom MD simulations did more than 20 % of the N-terminal residues adopt a discernible secondary structure. The question is why such an intricate interaction is necessary; after all, the domain exchange is associated with at least one energy barrier for a partial unfolding event.[36] Potential advantages of an oligomer with high internal symmetry, i.e., the lower chance of aggregation and higher stability, could also be realized without the sophisticated brace mechanism.[37] Mavirus itself provides a case in point with its minor capsid protein that forms a pentamer from five single jelly roll protomers but does not possess a brace or entanglement between the subunits. Nevertheless, it is found to be substantially more stable, at least against thermal denaturation.[17] In the light of our results, the penton apparently can afford to be more rigid as it plays a less prominent role in capsid assembly than the hexon.

In structural studies on other DJR-containing viruses such as PRD1,[19] Faustovirus,[20] or Adenovirus,[38] N-terminal interactions in capsomers were also observed and considered relevant for capsomer stability. Such interactions have been proposed as a general mechanism for stabilization of DJR capsomers.[20] Mavirus capsomers are considerably less stable than PRD1, which is very stable against denaturation,[39] suggesting that stabilization of the trimeric capsomer might not be the sole reason for its intertwined basal structure. Especially the change in overall structure upon capsomer dissociation as observed by CD spectroscopy points to different monomeric and oligomeric protomer structures. Calculations on the isolated protomer consistently predicted an alternative conformation specifically involving the released N-terminal arm that is not able to re-associate to trimers and thus most probably also not to capsids. As can be seen in, e.g., HK97[40] or PRD1,[41] capsid assembly often requires that hexons and pentons undergo conformational changes. A certain flexibility of the protein structure would be a prerequisite for this, which cannot be achieved with too rigid hexons. We suggest that the Mavirus capsomer balances the need for a pre-organized trimer with the necessity of switchable interface structures for efficient capsid assembly. For this purpose, it uses its specific brace mechanism to prevent the trimeric system from falling apart. This hypothesis is supported by HDX-MS data, showing that parts of the jelly roll units (especially in DJR2) of each protomer display higher conformational dynamics, comparable to that of the N-terminus. This could be a conserved mechanism in virophages. The second jelly-roll (DJR2) of the Sputnik virophage, for example, also is strongly involved in intercapsomer interactions and thus capsid formation, and shows high structural plasticity as well, while the intertwined arrangement of N- and C-termini at the capsomer's base is very similar to that in Mavirus.[22]

In summary, we showed that the main function of the brace and clasp mechanism of Mavirus is probably not only stabilization of the capsomer, as has been assumed from structural studies of related viruses. By preserving capsomer flexibility, this mechanism additionally allows for structural malleability of the capsomer required for capsid assembly as indicated by HDX-MS studies.[21] These findings extend our understanding of the assembly of not only Mavirus and the closely related Sputnik virophage but could potentially be extended to related viruses that share morphologies, such as poxo- or the large family of adenoviruses.[42] It further highlights the importance of studies on capsomer dynamics in addition to structural studies. Accordingly, a promising approach to hindering assembly and thus eventually virus propagation might be to rigidify capsomers, e.g. by binding or attaching specific ligands to decrease capsomer dynamics.

Material & methods

Material and buffer

The buffer used in these studies was a mixture of 20 mM Na-acetate, 20 mM Hepes, and 20 mM MES with addition of 150 mM NaCl and 1 mM EDTA (universal buffer UB). The advantage of this multicomponent buffer system is its good buffering capacity over a wide pH range (3–9).[43] All experiments were performed at pH 5 if not stated otherwise.

Protein purification

Protein purification for full-length MCP, MCP-D277C, MCP-W19L, and MCP-W269L was performed as reported earlier.[17] In brief, the gene of interest (cloned in a pETM-11) was expressed in *E. coli* BL21-CodonPlus(DE3)-RIL. After IPTG induction and harvesting, the cells were sonicated and applied to a Ni-column. The eluted protein was cleaved with TEV and applied to a second Ni-column. The flow through was dialyzed against low salt buffer (20 mM NaCl) and applied to a Resource S ion exchange column. The fractions containing MCP were collected and applied to a Superdex 200 16/60 (GE Healthcare, Little Chalfont, UK) size-exclusion column equilibrated with UB. The eluted fractions were analyzed for purity and size by SDS PAGE. Fractions containing MCP were pooled, concentrated by 50 kDa cut-off centrifugal filters (Amicron Ultra, Merck, Darmstadt, Germany), flash-frozen in liquid nitrogen, and stored at −80 °C. Protein concentration was determined using absorbance at 280 nm (ε_{280} 43780 mol^{-1} cm^{-1} for wild-type and MCP-277C, and 38280 mol^{-1} cm^{-1} for W19L and W269L). Concentrations stated in the text refer to monomer concentration.

CD spectroscopy

All CD spectra were collected on a JASCO J-810 spectropolarimeter (JASCO Germany GmbH) with a Peltier temperature-controlled cell holder (20 °C) using quartz cuvettes. Parameters used were a scanning speed of 20 nm/min, 2 sec integration time, and 1 nm bandwidth. The final spectra are an average of eight to twelve consecutive scans. Path length and final concentration depended on the spectral region of interest, 1 mm and 2.5 μM for far-UV and 1 cm and 1 μM for near-UV CD spectra. Chemical denaturation curves were recorded by mixing protein with GdmCl and subsequent equilibration for 30 min at room temperature before recording complete spectra.

Fluorescence spectroscopy

Intrinsic protein fluorescence spectra and FRET equilibrium measurements were recorded on a JASCO FP-8500 fluorescence spectrometer (JASCO Germany GmbH) with a Peltier-controlled cuvette holder (20 °C). Single spectra for intrinsic protein fluorescence were collected by 297 nm excitation with a final protein concentration of 2.5 μM. Measurements of FRET kinetics with Atto dyes and chemical denaturation were performed on a Fluorolog 3 (Horiba Scientific) with a Peltier-controlled cuvette holder (20 °C) and T-setup. An automatic titrator (Hamilton) was used for denaturation. Aliquots of 25 μl 6 M GdmCl (the exact concentration was determined with a refractometer) were added to an initial protein solution (1 μM, 800 μl). After equilibration (90–180 sec), a spectrum was recorded prior to addition of the next aliquot. The final spectra were corrected for chromophore dilution.

For FRET measurements, MCP D277C was labeled with Atto 488 dye as donor and Atto 594 dye as acceptor (Atto-Tec, Germany). In the capsomer, the C_α distance between the two introduced cysteine residues is approximately 4.0 nm and well within the chosen dye pair's Förster radius R_0 = 5.6 nm. Spectra for the dye-labeled protein were collected by exciting at either 488 nm (donor) or 594 nm (acceptor). For kinetic measurements (excitation 488 nm), MCP-277C was rapidly mixed with GdmCl and kinetic traces were recorded with two detectors at fixed wavelengths (520 nm and 630 nm, 0.1 sec integration time, 1 sec intervals). The FRET signal was calculated as the ratio of 520 nm to 630 nm signals and the resulting trace fitted with a mono-exponential curve. Final protein concentration was between 10 nM and 650 nM from different protein preparations with no effect of concentration on the observed kinetic rate constants. For equilibrium denaturation FRET spectra donor excitation at 488 nm was used with 45 nM final MCP concentration.

Stopped-flow experiments

All kinetic experiments were recorded on a BioLogic SFM-400 stopped-flow instrument in double or single mixing configuration (BioLogic Science Instruments, Claix, France) at 20 °C. The excitation wavelength was either 297 nm or 280 nm (2 mm slits). The fluorescence was recorded using two detectors simultaneously, one equipped with a 320 nm long pass (320FG03- 25, LOT Oriel Group), the other with a 310 nm band pass filter (10BPF10-310 Newport). The final protein concentration in the measurement cell after mixing was 1 μM.

For double-jump kinetics in direction of folding, the protein was incubated in 0.45 M GdmCl and during the first jump mixed with GdmCl and buffer to reach the desired higher GdmCl concentration. After a variable delay time (t1), the protein was mixed with buffer to dilute GdmCl back to 0.45 M with subsequent recording of the kinetic trace (t2). The mixing scheme was chosen to finally yield 1 μM protein in the measurement cell after the second mixing at all conditions.

For measurement of the dependence of the refolding rate of the first transition as function of GdmCl concentration the protein was incubated at 0.4 M GdmCl and during the first jump mixed with GdmCl to reach 1 M GdmCl. After an delay time of 1 second the GuHCl was diluted with buffer to the desired final GdmCl concentration and kinetic traces with a final protein concentration of 1 μM were recorded.

The acquired kinetic traces from single or double jumps were fitted with mono- or double-exponential equations to extract kinetic parameters using Origin 10.0 (OriginLab Corporation, USA) or Grafit 7 (Erithacus Software, UK).

Computational simulations

Starting from the capsomer crystal structure (6 g45.pdb), all-atom SBMs were set up with eSBMTools.[44,45] Native contacts were defined using the Shadow algorithm[46] and intra- and intermolecular contacts were weighted according to a 2:1 ratio. Simulations were performed with the MD package GROMACS[47] at different temperatures T in the range of 40 to 140 reduced GROMACS units (GU). Results are shown for $T = 90$ GU and $T = 106$ GU, for comparison, proteins are usually found to fold around $T = 100$ to 120 GU in SBMs. A leap-frog stochastic dynamics integrator was used with Verlet neighbor searching and periodic boundary conditions. Suitable partially unfolded and dissociated starting structures were manually extracted from the trajectories. Respective SBM simulations of these systems were set up and run analogously. The simulations additionally used a simulated-annealing protocol. Starting from $T = 90$ GU, the temperature was linearly decreased down to T = 60 GU after 90 % of the simulated time. Energy minimization of the dissociated starting structure was required before the actual run.

Explicit-solvent molecular dynamics simulations of MCP were set up and performed with GROMACS. The AMBER99SB-ILDN[48] force field was used. The first chain (residues 2 to 504) was extracted from the crystal structure 6 g45.pdb. The system was preprocessed by placing it in a box of TIP3P water with a minimum distance of 2 nm from the edges, charge-neutralized, energy-minimized, and equilibrated in the canonical and isothermal-isobaric ensemble. Simulations were run for 500 ns at temperatures $T = 300$, 314, 325, and 335 K at a pressure $p = 1$ bar. V-rescale temperature coupling and a Parrinello-Rahman barostat were used. Bonds were constrained using the LINCS algorithm. Electrostatics were treated with the Particle Mesh Ewald method. Verlet neighbor searching and periodic boundary conditions were applied.

Hydrogen-deuterium exchange coupled to mass spectrometry (HDX MS)

We performed hydrogen–deuterium exchange experiments to address the secondary-structure stability of full-length MCP with an emphasis on regions not visible in the available crystal structure. MCP was prepared at a final concentration of 93 μM and 5 μL aliquots were equilibrated for 1 min at 20 °C prior to starting the labeling reaction by adding 95 μL of deuterated UB at pD 7.0 (pH meter reading 6.6). Labeling reactions were prepared in triplicate and aliquots of 10 μL were removed after 10 s, 45 s, 3 min, 15 min, 1 h, and 4 h. Labeling was terminated by quenching with 80 μL of ice-cold 200 mM ammonium formic acid, pH 2.6, and 80 μL were injected into a cooled HPLC setup, as described previously.[49]

In brief, protein digestion was performed on a pepsin column (Poroszyme; Thermo Fisher) at 10 °C, and resulting peptides were washed on a C18 guard column (Discovery Bio C18, Sigma). Separation of peptides was achieved during a 7 min acetonitrile gradient (15–50 %) in the presence of 0.6 % (v/v) formic acid on a reversed phase column (XR ODS 75 mm × 3 mm, 2.2 μm; Shimadzu). Isotope distributions of individual peptides were recorded on a maXis electrospray ionization ultra-high-resolution time-of-flight mass spectrometer (Bruker) and deuterium incorporation was analyzed using the Hexicon 2 software package (for the complete analysis see also Figure S6)[50] (http://hx2.mpimf-heidelberg.mpg.de). The experimental conditions and analysis parameters are summarized in table S1 and cover recommended values to be reported.[51]

CRediT authorship contribution statement

Alexander Christiansen: Writing - original draft, Conceptualization, Methodology, Investigation.

Marie Weiel: Data curation, Writing - original draft, Investigation. **Andreas Winkler:** Writing - original draft, Software, Visualization, Investigation, Writing - review & editing. **Alexander Schug:** Supervision, Resources, Writing - review & editing. **Jochen Reinstein:** Conceptualization, Writing - review & editing, Project administration, Investigation.

DATA AVAILABILITY STATEMENT

Experimental data of this manuscript have been contained within this manuscript or are available upon request from J.R. (Jochen.Reinstein@mpimf-heidelberg.mpg.de) or A.W. concerning MS-HDX data (andreas.winkler@tugraz.at). Computational data are deposited on GitHub and available via the following link: https://github.com/mcw92/mavirus-capso_sims.

Acknowledgments

We thank Sabine Zimmermann, Susanne Eisel, and Ulrike Mersdorf for expert technical assistance, Diana Born for important insights and help with preparation of Figures 1 and 7 and Thomas Barends for help with the heat maps. We would also like to acknowledge continuous support from and interesting discussions with Ilme Schlichting and helpful comments on the manuscript by John Wray.

Funding

This work was funded by the Max Planck Society and the Austrian Science Fund (FWF) P32022 (to A. W.). The present contribution is supported by the Helmholtz Association Initiative and Networking Fund under project number ZT-I-003. The authors gratefully acknowledge support by the DFG under the GRK 2450.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2021.166859.

Received 19 November 2020;
Accepted 26 January 2021;
Available online 1 February 2021

Keywords:
virus protein;
HDX-MS;
structural dynamics;
protein folding and assembly;
MD simulation

Abbreviations used:
MCP, major capsid protein; DJR, double jelly roll: HDX-MS, hydrogen/deuterium exchange coupled to mass spectrometry; GdmCl, guanidinium chloride; CD, circular dichroism; FRET, Förster resonance energy transfer

References

1. Caspar, D.L., (1980). Movement and self-control in protein assemblies. Quasi-equivalence revisited. *Biophys. J.*, **32**, 103–138.
2. Caspar, D.L., Klug, A., (1962). Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.*, **27**, 1–24.
3. Zlotnick, A., (2005). Theoretical aspects of virus capsid assembly. *J. Mol. Recognit.*, **18**, 479–490.
4. Mannige, R.V., Brooks 3rd., C.L., (2010). Periodic table of virus capsids: implications for natural selection and design. *PLoS ONE*, **5**, e9423
5. Li, S., Roy, P., Travesset, A., Zandi, R., (2018). Why large icosahedral viruses need scaffolding proteins. *Proc. Natl. Acad. Sci. USA*, **115**, 10971–10976.
6. Parent, K.N., Zlotnick, A., Teschke, C.M., (2006). Quantitative analysis of multi-component spherical virus assembly: scaffolding protein contributes to the global stability of phage P22 procapsids. *J. Mol. Biol.*, **359**, 1097–1106.
7. Perlmutter, J.D., Hagan, M.F., (2015). Mechanisms of virus assembly. *Annu. Rev. Phys. Chem.*, **66**, 217–239.
8. Zlotnick, A., (2003). Are weak protein-protein interactions the general rule in capsid assembly?. *Virology*, **315**, 269–274.
9. Newcomb, W.W., Homa, F.L., Thomsen, D.R., Booy, F.P., Trus, B.L., Steven, A.C., et al., (1996). Assembly of the herpes simplex virus capsid: characterization of intermediates observed during cell-free capsid formation. *J. Mol. Biol.*, **263**, 432–446.
10. Wang, S., Chang, J.R., Dokland, T., (2006). Assembly of bacteriophage P2 and P4 procapsids with internal scaffolding protein. *Virology*, **348**, 133–140.
11. Basnak, G., Morton, V.L., Rolfsson, O., Stonehouse, N.J., Ashcroft, A.E., Stockley, P.G., (2010). Viral genomic single-stranded RNA directs the pathway toward a T=3 capsid. *J. Mol. Biol.*, **395**, 924–936.
12. Bancroft, J.B., Hiebert, E., Bracker, C.E., (1969). The effects of various polyanions on shell formation of some spherical viruses. *Virology*, **39**, 924–930.
13. Takimoto, T., Portner, A., (2004). Molecular mechanism of paramyxovirus budding. *Virus Res.*, **106**, 133–145.
14. Vennema, H., Godeke, G.J., Rossen, J.W., Voorhout, W. F., Horzinek, M.C., Opstelten, D.J., et al., (1996). Nucleocapsid-independent assembly of coronavirus-like

particles by co-expression of viral envelope protein genes. *EMBO J.*, **15**, 2020–2028.
15. Fukuhara, H., Ino, Y., Todo, T., (2016). Oncolytic virus therapy: a new era of cancer treatment at dawn. *Cancer Sci.*, **107**, 1373–1379.
16. Zlotnick, A., Lee, A., Bourne, C.R., Johnson, J.M., Domanico, P.L., Stray, S.J., (2007). In vitro screening for molecules that affect virus capsid assembly (and other protein association reactions). *Nature Protoc.*, **2**, 490–498.
17. Born, D., Reuter, L., Mersdorf, U., Mueller, M., Fischer, M. G., Meinhart, A., et al., (2018). Capsid protein structure, self-assembly, and processing reveal morphogenesis of the marine virophage mavirus. *Proc. Natl. Acad. Sci. USA*, **115**, 7332–7337.
18. Roberts, M.M., White, J.L., Grutter, M.G., Burnett, R.M., (1986). Three-dimensional structure of the adenovirus major coat protein hexon. *Science*, **232**, 1148–1151.
19. Benson, S.D., Bamford, J.K., Bamford, D.H., Burnett, R.M., (1999). Viral evolution revealed by bacteriophage PRD1 and human adenovirus coat protein structures. *Cell*, **98**, 825–833.
20. Klose, T., Reteno, D.G., Benamar, S., Hollerbach, A., Colson, P., La Scola, B., et al., (2016). Structure of faustovirus, a large dsDNA virus. *Proc. Natl. Acad. Sci. USA*, **113**, 6206–6211.
21. Patterson, A., Zhao, Z., Waymire, E., Zlotnick, A., Bothner, B., (2020). Dynamics of hepatitis B virus capsid protein dimer regulate assembly through an allosteric network. *ACS Chem. Biol.*, **15**, 2273–2280.
22. Zhang, X., Sun, S., Xiang, Y., Wong, J., Klose, T., Raoult, D., et al., (2012). Structure of Sputnik, a virophage, at 3.5-A resolution. *Proc. Natl. Acad. Sci. USA*, **109**, 18431–18436.
23. Vivian, J.T., Callis, P.R., (2001). Mechanisms of tryptophan fluorescence shifts in proteins. *Biophys. J.*, **80**, 2093–2109.
24. Nanda, V., Brand, L., (2000). Aromatic interactions in homeodomains contribute to the low quantum yield of a conserved, buried tryptophan. *Proteins*, **40**, 112–125.
25. Bryngelson, J.D., Onuchic, J.N., Socci, N.D., Wolynes, P. G., (1995). Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins*, **21**, 167–195.
26. Onuchic, J.N., Wolynes, P.G., (2004). Theory of protein folding. *Curr. Opin. Struct. Biol.*, **14**, 70–75.
27. Schug, A., Onuchic, J.N., (2010). From protein folding to protein function and biomolecular binding by energy landscape theory. *Curr. Opin. Pharmacol.*, **10**, 709–714.
28. Rey-Stolle, M.F., Enciso, M., Rey, A., (2009). Topology-based models and NMR structures in protein folding simulations. *J. Comput. Chem.*, **30**, 1212–1219.
29. Weiel, M., Reinartz, I., Schug, A., (2019). Rapid interpretation of small-angle X-ray scattering data. *PLoS Comput. Biol.*, **15**, e1006900
30. Reinartz, I., Sinner, C., Nettels, D., Stucki-Buchli, B., Stockmar, F., Panek, P.T., et al., (2018). Simulation of FRET dyes allows quantitative comparison against experimental data. *J. Chem. Phys.*, **148**, 123321
31. Parent, K.N., Khayat, R., Tu, L.H., Suhanovsky, M.M., Cortines, J.R., Teschke, C.M., et al., (2010). P22 coat protein structures reveal a novel mechanism for capsid maturation: stability without auxiliary proteins or chemical crosslinks. *Structure*, **18**, 390–401.
32. Chuan, Y.P., Fan, Y.Y., Lua, L.H., Middelberg, A.P., (2010). Virus assembly occurs following a pH- or Ca2+-triggered switch in the thermodynamic attraction between structural protein capsomeres. *J. R. Soc. Interface*, **7**, 409–421.
33. Stortelder, A., Hendriks, J., Buijs, J.B., Bulthuis, J., Gooijer, C., van der Vies, S.M., et al., (2006). Hexamerization of the bacteriophage T4 capsid protein gp23 and Its W13V mutant studied by time-resolved tryptophan fluorescence. *J. Phys. Chem. B*, **110**, 25050–25058.
34. Bennett, M.J., Choe, S., Eisenberg, D., (1994). Domain swapping: entangling alliances between proteins. *Proc. Natl. Acad. Sci. USA*, **91**, 3127–3131.
35. Bennett, M.J., Schlunegger, M.P., Eisenberg, D., (1995). 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci.*, **4**, 2455–2468.
36. Yang, S., Cho, S.S., Levy, Y., Cheung, M.S., Levine, H., Wolynes, P.G., et al., (2004). Domain swapping is a consequence of minimal frustration. *Proc. Natl. Acad. Sci. USA*, **101**, 13786–13791.
37. Goodsell, D.S., Olson, A.J., (2000). Structural symmetry and protein function. *Annu. Rev. Biophys. Biomol. Struct.*, **29**, 105–153.
38. Rux, J.J., Kuser, P.R., Burnett, R.M., (2003). Structural and phylogenetic analysis of adenovirus hexons by use of high-resolution x-ray crystallographic, molecular modeling, and sequence-based methods. *J. Virol.*, **77**, 9553–9566.
39. Mindich, L., Bamford, D., McGraw, T., Mackenzie, G., (1982). Assembly of bacteriophage PRD1: particle formation with wild-type and mutant viruses. *J. Virol.*, **44**, 1021–1030.
40. Gertsman, I., Gan, L., Guttman, M., Lee, K., Speir, J.A., Duda, R.L., et al., (2009). An unexpected twist in viral capsid maturation. *Nature*, **458**, 646.
41. Abrescia, N.G., Cockburn, J.J., Grimes, J.M., Sutton, G.C., Diprose, J.M., Butcher, S.J., et al., (2004). Insights into assembly from structural analysis of bacteriophage PRD1. *Nature*, **432**, 68–74.
42. San Martin, C., van Raaij, M.J., (2018). The so far farthest reaches of the double jelly roll capsid protein fold. *Virol. J.*, **15**, 181.
43. Brooke, D., Movahed, N., Bothner, B., (2015). Universal buffers for use in biochemistry and biophysical experiments. *AIMS Biophys.*, **2**, 336–342.
44. Lutz, B., Sinner, C., Bozic, S., Kondov, I., Schug, A., (2014). Native structure-based modeling and simulation of biomolecular systems per mouse click. *BMC Bioinf.*, **15**, 292.
45. Lutz, B., Sinner, C., Heuermann, G., Verma, A., Schug, A., (2013). eSBMTools 1.0: enhanced native structure-based modeling tools. *Bioinformatics*, **29**, 2795–2796.
46. Noel, J.K., Whitford, P.C., Onuchic, J.N., (2012). The shadow map: a general contact definition for capturing the dynamics of biomolecular folding and function. *J. Phys. Chem. B*, **116**, 8692–8702.
47. Abraham, M.J., (2011). Performance enhancements for GROMACS nonbonded interactions on BlueGene. *J. Comput. Chem.*, **32**, 2041–2046.
48. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J.L., Dror, R.O., et al., (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins-Struct. Funct. Bioinform.*, **78**, 1950–1958.
49. Winkler, A., Barends, T.R., Udvarhelyi, A., Lenherr-Frey, D., Lomb, L., Menzel, A., et al., (2015). Structural details of light activation of the LOV2-based photoswitch PA-Rac1. *ACS Chem. Biol.*, **10**, 502–509.

50. Lindner, R., Lou, X., Reinstein, J., Shoeman, R.L., Hamprecht, F.A., Winkler, A., (2014). Hexicon 2: automated processing of hydrogen-deuterium exchange mass spectrometry data with improved deuteration distribution estimation. *J. Am. Soc. Mass Spectrom.*, **25**, 1018–1028.

51. Masson, G.R., Burke, J.E., Ahn, N.G., Anand, G.S., Borchers, C., Brier, S., et al., (2019). Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nature Methods*, **16**, 595–602.